ADULT AGE DIFFERENCES IN MEMORY:
THE ROLES OF FACILITATION AND INTERFERENCE

by

Scott Charles Brown

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

0-612-49984-7

ADULT AGE DIFFERENCES IN MEMORY:

THE ROLES OF FACILITATION AND INTERFERENCE

Scott Charles Brown

Department of Psychology, University of Toronto

Ph.D. 2000

Abstract

Five experiments examined whether age differences in a memory task were greater when successful responding was opposed by prior learning (i.e., interference) than when it was consistent with prior learning (i.e., facilitation). Older and younger adults' performance under facilitation and interference conditions was examined in a direct memory test (Experiments 1 and 2), as well as an indirect memory test (Experiments 3, 4, and 5). The experimental results were used to distinguish between two alternative accounts of aging and memory: impaired recollection, versus impaired inhibition.

Experiments 1 and 2 employed a variant of Jacoby's (1991) process dissociation procedure, in which participants studied sentence completions that had been either the dominant or the non-dominant completion for that sentence in a prior learning phase. Both experiments revealed age declines in recollection of repeated targets and in inhibition of repeated distractors; however, these age decrements were reliable only in Experiment 2. Supportive conditions in the first two experiments (i.e., reinstatement of perceptual context and of prior knowledge) did not reliably reduce age differences in recall.

Experiments 3, 4, and 5 employed Reingold's (1995) letter deletion task to examine age differences in facilitation and interference priming. Facilitation priming appeared age-invariant when absolute reaction time (RT) differences were considered, but showed age declines when proportionate RT differences were considered (Experiment 4). Interference priming was age-invariant in one experiment (Experiment 3), but showed age-related increases in two other experiments employing greater numbers of trials (Experiments 4 and 5). These latter two experiments employed manipulations (i.e., presentation frequency and working memory load) that were expected to differentially increase interference for older adults. However, the latter two experiments revealed an age-related increase in interference across *all* conditions that reflected an increased number of very long RTs in older adults.

In summary, the five experiments provided evidence for age declines in both recollection and inhibition. These results support the view that age differences in memory are due to general failures of cognitive control, rather than a specific inhibitory failure. However, the form and degree of age-related deficits appear to depend on task and subject parameters.

Acknowledgments

I am very grateful to Fergus Craik for his exceptional support, guidance, and encouragement during my six years as a graduate student at the University of Toronto. I thank my thesis committee members, Eyal Reingold and Morris Moscovitch, for sharing their time and expertise throughout all stages of the dissertation research. I also thank Morris Moscovitch for generously allowing me to test his senior volunteers on the Erindale College campus of the University of Toronto. I thank the remaining members of my final oral examination committee – Lynn Hasher, Ken Dion, and Jay Pratt – for their considerable efforts on my behalf. I also thank Denise Park and the members of the Craik lab for their many helpful suggestions.

Special thanks go to Elizabeth Olszewska, who assisted in scheduling research participants, and to Marilyne Ziegler, who provided computer programming assistance. I thank Ann Lang and Jennie Sawula for their unflagging administrative support. I thank Jim Perretta, Carol Okamoto, and Jennifer Rivkin for their help in testing research participants. I also thank Jim Perretta for his help in stimulus construction. In addition, I appreciate the intangible supports provided by Carol, Jim, Brenda, and my other friends in the psychology department. I appreciate the funding provided to me by the Department of Psychology, the School of Graduate Studies, and the Institute for Human Development, Life Course, and Aging. I am also grateful to the many younger and older volunteers who took part in this research.

Lastly, I am indebted to my parents, Charles and Charlene Brown, for their constant emotional and financial support. This dissertation is dedicated to my two late grandmothers, Bernadette Brown and Angeline Mooney, who spurred my interest in cognitive aging research.

Table of Contents

List of Tables

List of Figures

List of Appendices

CHAPTER ONE:
GENERAL INTRODUCTION

The goal of the present research is to better account for the appearance and disappearance of age differences in memory across the literature. The fact that there are age declines in memory has been well documented: A recent literature review (Craik & Jennings, 1992) indicated that age differences are most pronounced on explicit/direct memory tests, which instruct participants to refer back to a study episode. In contrast, age differences are smaller or nonexistent on implicit/indirect tests, which require no conscious awareness of the study episode (e.g., Light & La Voie, 1993). In addition, some researchers have found sizable age decrements when participants must inhibit irrelevant information (e.g., Hasher & Zacks, 1988; Zacks & Hasher, 1997); however, other researchers have found no age differences in susceptibility to distraction (see, e.g., Burke, 1997, and McDowd, 1997, for some contradictory findings). These disparate patterns of age differences might be best addressed by an analysis based on environmental support (e.g., Craik, 1983, 1986) and on facilitatory versus interfering encoding/retrieval conditions (e.g., Jacoby, 1991; Reingold, 1995).

Environmental support. Craik (1983, 1986) argued that the activity of remembering is a complex interaction between internal mental processes and external information (see also Jenkins, 1979). He therefore proposed that tasks differ in the extent to which their external context induces or supports the required mental operations. A further suggestion was that older adults depend heavily on environmental support, relative to younger adults, and should perform relatively well when such support is present and poorly when it is absent. In cases where individuals cannot utilize external support, they must rely instead on "self-initiated" or internally-driven operations, which are resource demanding and hence especially difficult for older adults (Craik & Byrd, 1982). This account explains why age differences are typically smaller in recognition than free recall (Craik & McDowd, 1987): Recognition is strong in environmental support (i.e., studied items reappear at test), whereas free recall provides little support (i.e., no cues appear at test), thus increasing the need for self-initiated processing.

One implication of the environmental support hypothesis is that improvements in encoding and retrieval conditions should produce a pattern of compensation in which older adults show greater performance gains than younger adults. This prediction has received some empirical validation: For example, age differences in memory have been reduced when the same general semantic cues appear at study and test (e.g., Rabinowitz, Craik, & Ackerman, 1982; Shaw & Craik, 1989). However, Light (1991) has indicated that the compensation pattern appears infrequently in the literature: The majority of studies reported that young and old benefit equally from greater environmental support, whereas some studies reported a compensation pattern and still others reported greater benefit to younger adults. In response to Light's criticism, Craik and Jennings (1992) implied that the pattern of age-related benefits from increasing environmental support depends on the initial level of support and that most tasks have been only moderately supportive. That is, greater benefits to the young occur when conditions are highly open-ended (e.g., few study and/or test cues); that equal benefits occur with somewhat greater guidance and constraint (e.g., pictures); and finally that compensation occurs when conditions are highly supportive at both encoding and retrieval (e.g., semantic study/test cues). Moreover, Craik and his colleagues argue that the degree to which older adults benefit from supportive conditions depends partly on subject parameters, such as education and levels of social and physical activity. More specifically, it appears that older adults who are well-educated, healthy, and neurologically intact benefit more from environmental support than do older adults who are "challenged" along one or more of these dimensions (see also, e.g., Craik, Byrd, & Swanson, 1987).

However, Craik and Jennings' (1992) account is unsatisfactory, given the tendency for any set of findings to be fitted post hoc to one or other of the theoretical patterns. It is therefore argued that an analysis of tasks in terms of facilitation and interference provides a better means of accounting for the various patterns of age differences.

Facilitation versus interference in memory. The use of both facilitation and interference conditions within the same memory task may be useful for understanding the nature of age differences in memory. Facilitation may be defined simply as the reinstatement at retrieval of previously learned associations and contexts (e.g., an A-B, A-B situation in associational learning), whereas interference may be regarded as the provision of different associations and/or contexts than those employed in prior learning (e.g., an A-B, A-C situation). These two types of conditions provide a means of accounting for the differing magnitude of age differences in the literature. For instance, it is plausible that on direct memory tests, age differences should be larger in interference conditions (i.e., when prior learning opposes successful responding) than in facilitation conditions (i.e., when prior learning contributes to successful responding). However, to date, there have been relatively few aging studies that have employed both facilitation and interference conditions within the same memory task.

One exception to this general rule was an early study by Ruch (1934): This was the first study of intentional paired-associate learning that directly compared younger versus elderly adults (Kausler, 1994). Ruch examined age differences in the facilitatory versus interfering effects of pre-experimental exposure to studied stimuli. Specifically, Ruch found that age differences in learning were small when the to-be-associated elements were logically related word pairs (e.g., the words brown and green) but that age differences were large when the stimulus and response were the elements of false arithmetic problems (e.g., the response "5" for the stimulus "2x2"). Ruch correctly predicted that the former situation would produce relatively small age differences, given the familiarity of the studied material. In contrast, Ruch correctly predicted that the latter situation would generate interference from past habits (e.g., the habit of responding "4" to "2x2") and that such interference would increase with increasing age. This finding of large age differences in an interference condition, but not in a facilitation condition, is consistent with an age-related decline in the ability to inhibit prepotent influences in memory (Hasher & Zacks, 1988).

More recently, Hay and Jacoby (1999) have examined the beneficial versus harmful effects of experimentally-acquired habits on age differences in memory. More specifically, these researchers employed an initial learning phase in which an associative cue (e.g., "knee-b_n_") was repeatedly paired with one word (e.g., "bone") on the majority of trials and with another word (e.g., "bend") on the remaining trials. The associative cue then reappeared on a subsequent study list, with either the typical completion or the atypical completion from the previous training phase. At test, the older and younger participants were instructed to retrieve the studied completion for each cue word. In Hay and Jacoby's procedure, the reappearance of the typically trained completion at study was a facilitation or "in-concert" condition, in which both habitual responding and recollection of the specific study item should contribute to correct responding. In contrast, the reappearance of the atypical completion at study was an interference or "opposition" condition, in which recollection but not habitual responding should yield a correct response. These two types of condition served as a basis for estimating the contributions of recollection and habitual influences to memory performance; for this reason, the paradigm is called the *process dissociation procedure* (see also, e.g., Hay & Jacoby, 1996; and Jacoby, 1991). In line with much previous work (e.g., Ruch, 1934; Winocur & Moscovitch, 1983), Hay and Jacoby (1999) found an age-related increase in interference: That is, the older adults were more likely than the younger adults to erroneously produce the typically-trained completion when the atypically-trained completion reappeared at study. However, in addition, Hay and Jacoby found reliable age decrements in facilitation. In other words, the younger adults continued to outperform the older adults even when the studied completion had appeared with its associative cue multiple times at training. Jacoby and his colleagues regard this type of finding as evidence for age-declines in conscious recollection but age-invariance in automatic influences (see also, e.g., Jacoby, Yonelinas, & Jennings, 1997, and below). Unfortunately, however, Hay and Jacoby did not directly compare the magnitude of age differences in recall in their two types of condition (facilitation versus interference). Instead, these researchers reported participants' correct performance for the facilitation conditions and their error rates for the interference conditions, which served as the basis for calculating recollection and habitual influences (see also "The impaired recollection hypothesis" and Chapter 2, below). Consequently, it is not possible to definitively assess whether age differences in recall were greater in the interference conditions of Hay and Jacoby's (1999) study, as compared to their facilitation conditions.

The two studies described above (i.e., Ruch, 1934; and Hay & Jacoby, 1999) examined age differences in facilitation and interference conditions of a direct test, where participants deliberately refer back to a prior study episode. Unfortunately, there appear to be even fewer aging studies comparing facilitation and interference in an indirect test, where no conscious recollection is required (see, e.g., Roediger & McDermott, 1993, for a review of indirect memory tests). One possible example of such an indirect test is the Stroop (1935) task, in which participants are asked to name the ink color in which color words appear. Participants are typically faster at color-naming when the color and word are congruent (e.g., the word red printed in red ink) than when the color and word are incongruent (e.g., the word red printed in blue ink). The former condition is an example of facilitation where both word-reading (a highly practiced or automatic process) and color-naming (a less practiced or consciously controlled process) act to produce the desired response. In contrast, the latter condition is an example of interference because color-naming produces the desired response, whereas word-reading does not. The task can therefore be considered a type of indirect memory test because, although no study list is presented, the participants' previous learning (i.e., reading words in everyday life) is indirectly assessed. With regard to age differences on the Stroop task, researchers have generally found that older adults show greater interference than younger adults, but are reasonably equivalent to younger adults in facilitation (see, e.g., Spieler, Balota, & Faust, 1996). This finding is once again consistent with the claim (e.g., Hay & Jacoby, 1999) that aging impairs consciously controlled processes such as color naming, but does not affect automatic processes such as word reading. However, this paradigm has one of the same interpretative problems as the Ruch (1934) study: Namely, the degree of facilitation and interference shown by participants depends largely on pre-experimental learning, which is beyond experimenter control and therefore may differ substantially both within and across age groups.

There do not appear to be many other aging studies which have incorporated facilitation and interference conditions within the same indirect memory test. The vast majority of indirect tests measure facilitation, or improvement in speed and accuracy as a result of prior experience (also called *priming*; see, e.g., Roediger & McDermott, 1993; and Schacter, 1987, for reviews of the implicit/indirect memory literature). The one major exception to this rule is the *negative priming* procedure (e.g., Tipper & Cranston, 1985), which is a selective attention paradigm in which the distractor on one trial becomes the target on the next trial. In this situation, the participant typically performs worse (i.e., more slowly or inaccurately) if the target had previously served as a distractor on the immediately preceding trial than if it had not previously appeared. One popular interpretation of this result is that participants suppress or inhibit the activation of a distractor and that this inhibition dissipates slowly over time (see, e.g., May, Kane, & Hasher, 1995). However, older adults have impaired inhibitory mechanisms and hence do not suppress the distractor with the same efficiency as the younger adults: As a result, the older adults do paradoxically better than their younger counterparts when the initially ignored distractor subsequently reappears as the target (e.g., Hasher, Stoltzfus, Zacks, & Rypma, 1991). Therefore, studies that have included both negative priming and repetition priming conditions (e.g., McDowd & Filion, 1995) have revealed age declines in inhibition but age-invariance in facilitation. However, it should be emphasized that performance in the negative priming procedure does not reflect interference per se (D. C. Park, personal communication, January 1998). In other words, the interference paradigms described above have all involved overcoming the activation of previously relevant information. In contrast, the standard negative priming paradigm involves the lingering aftereffects of suppressing irrelevant information. In fact, several negative priming studies have included a further condition to assess interference: A no-distractor (i.e., target only) condition is compared with a control condition (i.e., target plus distractor) in order to assess the interfering effects of concurrent distraction. However, even if this type of interference is relevant to the present investigation, conflicting results have been obtained: Some negative priming studies that included this condition have reported an age-related increase in interference (e.g., McDowd & Oseas-Kreger, 1995), whereas other studies have reported no age difference (e.g., Kane, Hasher, Stoltzfus, Zacks, & Connelly, 1994).

In summary, the extant literature does not allow for any broad and sweeping conclusions regarding the differential effects of facilitation and interference on age differences for either direct or indirect memory tests. There have been few aging studies that have directly compared the facilitatory versus interfering effects of prior learning on either type of memory test. Nonetheless, two competing theoretical accounts each make clear predictions regarding the magnitude of age differences in memory under facilitation and interference conditions.

Impaired Inhibition or Impaired Recollection?

The disinhibition hypothesis. One relevant theoretical account of aging and memory is known as the *disinhibition hypothesis* (Hasher & Zacks, 1988). This account attributes age related memory decrements to declines in inhibition. The disinhibition hypothesis derives partly from selection-for-action theories (see, e.g., Tipper, 1992) that argue for the existence of both excitatory and inhibitory processes in selection. Hasher and Zacks hypothesized that aging impairs inhibitory attentional processes, while sparing excitatory processes. In their original framework, Hasher and Zacks assumed that inhibitory processes served two major functions, controlling *access* to and *deletion* from working memory. In other words, efficient inhibition prevents task-irrelevant information from entering working memory and allows for the removal of irrelevant or no-longer-relevant information from working memory. More recently, Hasher, Zacks and May (1999) suggested a third function of inhibitory mechanisms, *restraint* over strong responses such as highly practiced actions. These claims have received a fair amount of empirical support: For example, older adults are more distracted than younger adults by irrelevant text in an on-line reading task (e.g., Connelly, Hasher, & Zacks, 1991) and have greater difficulty suppressing irrelevant information on a negative priming task (e.g., Hasher et al., 1991). Moreover, older adults show larger "fan effects" or more cluttered working memories than do younger adults (e.g., Gerard, Zacks, Hasher, & Radvansky, 1991) and show more proactive interference than do their younger counterparts (e.g., Winocur & Moscovitch, 1983). Furthermore, older adults have greater difficulty than younger adults in abandoning disconfirmed ideas when reading sentences or prose passages (e.g., Hamm & Hasher, 1992; Hartman & Hasher, 1991) and in deliberately forgetting previously presented information (e.g., Zacks, Radvansky, & Hasher, 1996). Finally, older adults have greater difficulty than younger adults on an "antisaccade" task which requires participants to overcome reflexive responding by looking away from, rather than toward, a visual cue (e.g., Butler, Zacks, & Henderson, 1996; but see Munoz, Broughton, Goldring, and Armstrong, 1998, for a different result). However, despite the rather impressive evidence for the disinhibition hypothesis, this account has not been universally supported: For example, there are now several studies showing age-equivalence in both identity and location negative priming (see McDowd, 1997, for a review). Moreover, some of the findings of Hasher, Zacks and their colleagues are subject to alternative interpretations (see, e.g., Burke, 1997). However, on the whole, the inhibitory account remains a useful one that accounts for a broad range of findings in the cognitive aging literature (see also, e.g., Zacks & Hasher, 1997).

The disinhibition hypothesis is relevant to the present investigation in that it suggests the following prediction: *Age differences in memory should be larger under interference conditions than under facilitation conditions.* That is, if aging impairs inhibitory functioning, then older adults should be disproportionately harmed in the present interference conditions, as compared to the facilitation conditions. For example, in the interference conditions of the present studies, participants must be able to delete irrelevant information from working memory and overcome prepotent responses. Although the present facilitation conditions may provide some sources of distraction (e.g., infrequently presented alternative responses or internal "noise"), the amount of distraction should be less than that observed in the interference conditions. Finally, the finding of greater age decrements under interference conditions, as compared to facilitation conditions, should be obtained for both direct and indirect memory tests. In other words, regardless of whether retrieval is deliberate or not, the disinhibition hypothesis suggests that age declines in memory should largely be restricted to interference situations.

<u>The impaired recollection hypothesis.</u> An alternative account of aging and memory suggests that older adults' memory decrements are caused by failures of deliberate recollection

(see, e.g., Jacoby, 1991; and Light, 1991). The impaired-recollection hypothesis is based on the observation that age differences are small on indirect/implicit memory tests, which require no conscious awareness of a prior study episode, as compared to direct/explicit memory tests, which require deliberate referencing of the study episode (see, e.g., Light & LaVoie, 1993, for a review). However, age declines in implicit memory do exist, as was shown by Light and La Voie's (1993) meta-analysis of this literature. Some researchers (e.g., Jacoby, 1991) have commented that these small age decrements on implicit tests reflect the fact that tasks are not process pure: In other words, performance on a given memory test may reflect both consciously controlled and automatic uses of memory. Therefore, to the extent that younger adults show an advantage over older adults on indirect tests, the younger adults could be making greater use of conscious recollection than their older counterparts. As an alternative to task dissociations, Jacoby (1991) proposed a process dissociation procedure (PDP), which served to experimentally separate the contributions of automatic and controlled processes to memory performance. The process dissociation logic consists of pairing a facilitation condition, in which automatic and consciously controlled processes both promote successful responding, with an interference condition, in which consciously controlled but not automatic processes promote successful responding. These conditions then serve as a basis for estimating recollection and automaticity (see Jacoby, 1991, and Chapter 2 for a more thorough discussion of the process dissociation logic). Jacoby's procedure has yielded results that are largely congruent with the task dissociation literature: That is, age differences are typically large for recollection but quite small or even non-existent for automatic influences (see, e.g., Jacoby et al., 1997, for a review of process dissociation research). Interestingly enough, these results are congruent with an earlier suggestion by Hasher and Zacks (1979), that age differences should be greater for effortful than automatic processes, given the lesser cognitive capacity that was available for the former type of process with age (see also the "Processing resources" section, below). However, Hasher and Zacks (1988) subsequently shifted their views to focus on age declines in inhibition, rather than controlled processes per se.

The impaired-recollection account is of some relevance to the present investigation in that it makes the following prediction: *Age differences in direct memory tests should be reliable for both facilitation and interference conditions* (see, e.g., Hay & Jacoby, 1999, for a similar suggestion). This is the case because, in Jacoby's (1991) terms, facilitation reflects both consciously controlled and automatic memory influences operating in the same direction (the former being more age-sensitive than the latter), whereas interference reflects (unsuccessful) conscious attempts to overcome the effects of automatic memory influences. Moreover, if older adults are deficient in these consciously controlled uses of memory, then older people should show impairments whenever the retrieval task involves deliberate referencing of a prior study episode. In contrast, *smaller age differences should be observed in indirect memory tests*, which do not require deliberate recollection. Interestingly enough, the impaired-recollection hypothesis allows for the possibility that age differences might be smaller under the interference conditions of an indirect memory test, as compared to the facilitation conditions of the same test. This is the case because indirect facilitation conditions might be more prone to *explicit contamination*, or the deliberate referencing of a prior study episode, even when the experimenter fails to reveal the relation between study and test. This potential for explicit contamination is less likely in the interference conditions of the indirect test because there should be less conscious incentive to use the previously studied information. In other words, deliberate referencing of a study list would prove more costly in an interference paradigm (see also Reingold, 1995, for similar suggestions).

In summary, comparisons of older and younger adults in facilitation and interference conditions of the same experimental task provide a means of distinguishing between a deficit in inhibition and a deficit in conscious recollection. Namely, if age differences in memory arise solely as a result of impaired inhibitory processes, then older adults should differ from younger adults on the opposition trials of both direct and indirect tests (i.e., when prior learning interferes with correct responding) but not on in-concert trials (i.e., when prior learning facilitates correct responding). However, if age differences are caused by more general impairments of conscious recollection, then older adults should exhibit reliable declines in facilitation conditions relative to younger adults (see, e.g., Hay & Jacoby, 1999, for similar arguments).

In closing, however, there are two additional types of theoretical account that have some relevance to the present investigation: neuropsychological accounts (e.g., Moscovitch & Winocur, 1992), and processing-resource accounts (e.g., Craik & Byrd, 1982). These latter two accounts are at least partly consistent with the claims made by both Hasher and Zacks (1988) and Jacoby (1991). Each account is described in turn.

Other Relevant Theoretical Accounts

Neuropsychological accounts of memory aging. The pattern of age-related memory impairments has also been noted by neuropsychologists (e.g., Moscovitch & Winocur, 1992), who have linked these deficits to deterioration of the hippocampal system and the frontal lobes. In particular, hippocampal damage impairs memory with awareness, whereas frontal damage impairs self-initiated encoding and retrieval processes (e.g., Moscovitch, 1992, 1994). The hippocampus and adjacent cortical structures (e.g., medial temporal lobes) underlie the obligatory integrative and ecphoric processes performed upon all consciously apprehended information. These structures are more specifically involved in *associative* or cue-dependent uses of memory such as paired-associate learning and retrieval. In contrast, the frontal lobes subserve more *strategic* or deliberate uses of information, such as initiating memory search. These latter, frontally-mediated processes are termed "working with memory" because of their emphasis on active manipulation of encoded information (e.g., Moscovitch & Winocur, 1992). In summary, neuropsychological accounts of memory aging attribute age-related declines in memory control processes such as inhibition to frontal impairments (see also, e.g., Moscovitch & Winocur, 1995; and West, 1996), and age-related declines in deliberate recollection to both hippocampal (ecphoric) and frontal (strategic) impairments. In contrast, the age-invariance of nonstrategic memory processes involved in many indirect tests (e.g., perceptual fluency) is attributed to sparing of the posterior neocortex (e.g., Moscovitch & Winocur, 1992).

Processing resource accounts of memory aging. A further and somewhat related perspective attributes older adults' memory decrements to declines in the amount of psychological resources available for mental activity. The accounts in this area vary in the metaphor used to designate processing resources. Some researchers have claimed that older people's memory impairments reflect declines in mental energy or attentional resources (e.g., Craik & Byrd, 1982), whereas other researchers have pointed to impairments in the time or speed of processing (e.g., Salthouse, 1985), and still others have cited reduced mental capacity or working memory (e.g., Craik, Morris, & Gick, 1990). At present there appears to be ample evidence for age declines in all three of these areas. For example, a recent meta-analysis of the divided attention literature has revealed clear age-related declines in dual-task performance (Hartley, 1992). Moreover, psychomotor slowing is quite commonly reported in the cognitive aging literature (see, e.g., Salthouse, 1996, for a review), as is reduced working memory capacity (e.g., Park, Smith, Lautenschlager et al., 1996). Furthermore, structural equation modeling and path analysis have clearly identified working memory and, more fundamentally, speed of processing to be mediators of memory performance in older and younger adults (see, e.g., Park et al., 1996, and Salthouse, 1996). However, in Park et al.'s (1996) study, the working memory construct was directly related only to the two most effortful memory measures: free and cued recall. This latter finding provides further evidence that tasks differ with respect to their processing requirements and that age differences should be larger on more effortful memory tasks, as compared to more automatic tasks (see also, e.g., Craik, 1986, and Jacoby et al., 1997, for similar suggestions). It appears somewhat likely that age declines in attentional resource, processing speed, working memory, and even sensory function reflect the same underlying psychological resource (e.g., brain integrity) or at least share some common cause (see, e.g., Salthouse, 1985, and Lindenberger & Baltes, 1994, for similar suggestions). Finally, it is of some interest to the present investigation that some researchers consider attentional control and inhibition to be types of processing resource in their own right (see, e.g., Anderson, 1997). The relation of control processes to both working memory capacity and frontal lobe functioning has been noted elsewhere (e.g., Hasher & Zacks, 1988, Kimberg, D'Esposito, & Farah, 1997).

Research Overview

The present investigation was therefore conducted to assess the magnitude of age differences in the facilitation and interference conditions of a direct memory test (Experiments 1 and 2), as well as an indirect memory test (Experiments 3, 4, and 5). Facilitation conditions were those in which prior learning contributed to successful responding, whereas interference conditions were those in which learning opposed successful responding. The obtained patterns of age differences under these two types of condition were used to empirically distinguish between impaired inhibition (e.g., Hasher & Zacks, 1988) and impaired recollection (e.g., Jacoby, 1991) as an explanation of age differences in memory. The disinhibition hypothesis predicted greater age differences in the interference conditions of both memory tests, as compared to the facilitation conditions. In contrast, the impaired-recollection hypothesis predicted reliable age differences in the facilitation and interference conditions of the direct memory test, but little or no age difference for either type of condition in the indirect memory test.

Experiments 1 and 2 employed a version of the process dissociation procedure (see, e.g., Hay and Jacoby, 1996) in which older and younger adults studied sentence completions that had been either the dominant completion or the non-dominant completion for that sentence in a prior learning phase. These conditions served to examine the facilitatory versus interfering effects of repetition on age differences in recall. In other words, both age groups should find it easy to recall the study item when it had previously appeared frequently and with little competition from distractors, as compared to when the study item had appeared infrequently and with greater response competition. However, older adults should show greater interference than their younger counterparts. Supportive conditions in the first two experiments (i.e., repeated reinstatement of perceptual context, and of prior knowledge) were expected to reduce age differences in memory, in line with Craik and Jennings' (1992) suggestions regarding environmental support. More specifically, repeatedly presenting sentence completions in the same voice, rather than different voices (Experiment 1), and using plausible rather than implausible sentence completions (Experiment 2) were expected to differentially benefit older adults by reducing the demands placed on their limited processing resources. However, it was unclear whether the present facilitation conditions would actually eliminate age differences in recall.

Experiments 3, 4, and 5 employed a letter deletion task (Reingold, 1995) in which older and younger adults deleted one of two specified letters from a six-letter nonword in order to create a meaningful five-letter word. Reingold's task was used to examine age differences in the facilitation (Experiment 4) and interference (Experiments 3, 4, and 5) conditions of an indirect memory test. In the facilitation conditions, a previously presented nonword reappeared with the same candidate letters for deletion as in the nonword's previous appearance. However, in the interference conditions, a previously presented nonword reappeared, albeit with different letters specified as candidates for deletion than in the nonword's previous appearance. Facilitation and interference was assessed by changes (either decreases or increases) in the time to solve a word puzzle as a result of prior learning. Experiment 3 compared older and younger adults on the interference condition of Reingold's (1995) third experiment. Although it was expected that both age groups would show interference, it was unclear whether older adults would show as much interference, or more interference, than the younger adults. Experiments 4 and 5 subsequently employed manipulations (i.e., presentation frequency and working memory load) that were expected to differentially increase interference for older adults. More specifically, Experiment 4 varied whether a nonword was presented either once only, or four times, prior to a change in candidate letters for deletion. This manipulation resulted in the creation of facilitation conditions involving the precise reinstatement of stimuli. Finally, Experiment 5 varied whether the candidate letters for deletion were visibly displayed during each trial or were revealed only before each trial and subsequently retained in working memory. It was predicted that strengthening the initial response to a stimulus (Experiment 4) and increasing the contents of working memory (Experiment 5) would disproportionately harm the older adults in the interference trials of the latter two experiments. Experiment 4 was the only experiment in the series that included both facilitation and interference conditions: Although it was assumed that the older adults would show as much facilitation priming as the younger adults (e.g., La Voie & Light, 1994), it was unclear whether the older adults would show as much, or more, interference priming relative to the younger adults.

Taken as a whole, the five experiments serve to examine whether age differences in memory are caused by impaired inhibition or impaired recollection. It is possible that both types of deficit can occur with aging and that the nature and extent of the impairment varies across tasks and individuals. The present studies therefore examine correlations between memory performance and several individual-difference measures such as verbal ability, education, neurological functioning, and expertise (e.g., experience solving word puzzles). Finally, the studies serve to further identify which forms of environmental support are most useful to older, versus younger, adults.

CHAPTER TWO:
AGE DIFFERENCES IN FACILITATION AND INTERFERENCE IN A DIRECT MEMORY TEST

As described in Chapter 1, two predominant accounts of age differences in memory (i.e., impaired recollection vs. impaired inhibition) both predict sizable age differences in the interference conditions of direct memory tests. However, these accounts appear to generate somewhat different predictions regarding the magnitude of age differences under facilitation conditions. Namely, the impaired recollection hypothesis (e.g., Jacoby, 1991; Hay & Jacoby, 1999) predicts reliable age declines in the facilitation conditions of most direct tests, due to a general decline in explicit remembering. In contrast, the impaired inhibition hypothesis (e.g., Hasher & Zacks, 1988) appears to predict age-equivalence in the facilitation conditions of direct memory tests. In other words, conditions that minimize the potential for distraction (e.g., repetition of the same stimulus) should largely eliminate age differences in memory. However, the disinhibition hypothesis can account for any age difference in recall that is accompanied by an age-related increase in intrusions of either the experimental or extra-experimental variety.

Experiments 1 and 2 therefore examined age differences in the facilitatory and interfering effects of repetition on a direct memory test. Both studies were inspired by Jacoby's (1991) process-dissociation procedure (PDP) – specifically, the version used by Hay and Jacoby (1996). Hay and Jacoby's procedure involves creating a habit by frequently pairing a stimulus with two possible responses during a training phase. More specifically, an associative cue (e.g., "knee-b_n_") is paired with one response (e.g., "bone"), arbitrarily deemed typical, on the majority of training trials and an alternative response (e.g., "bend"), deemed atypical, on the remaining training trials. Then the subsequent study list contains the same associative cue with either the typical or atypical response from training. At test, the participant's task is always to retrieve the study item. Within this context, the reappearance of the typical training item at study serves as a facilitation condition in which successful recall reflects both recollection of the study item and habitual memory influences from the training phase. In contrast, the reappearance of the atypical training item at study serves as an interference condition, where intrusions of the typical response reflect automatic influences operating in the absence of recollection. Hay and Jacoby therefore use participants' performance in facilitation and interference conditions to derive estimates of recollection and habit.

The main measure of interest in Hay and Jacoby's (1996) version of the PDP is the number of typical responses given by participants, both when the study completion is the typical completion at training (a facilitation condition) and when the study completion is atypical (an interference condition). Performance in the facilitation and interference conditions is used to estimate consciously controlled and automatic memory influences, or recollection and habit, respectively. More specifically, the proportion of typical trials correctly completed with typical words equals the proportion of trials in which either recollection, $\underline{R}$, and/or habit, $\underline{H}$, exert an influence:

$$\underline{P}(\text{Typical}) = \underline{R} \text{ or } \underline{H} = \underline{R} + \underline{H} - \underline{R\&H} \qquad (1)$$

In contrast, the proportion of atypical trials incorrectly completed with typical words equals the proportion of trials in which habit operates in the absence of recollection:

$$\underline{P}(\text{Atypical}) = \underline{H} - \underline{R\&H} \qquad (2)$$

Recollection can therefore be estimated as:

$$\underline{R} = \underline{P}(\text{Typical}) - \underline{P}(\text{Atypical}) \qquad (3)$$

Finally, if independence is assumed between recollection and habit (e.g., Jacoby et al., 1997), then habit can be estimated as:

$$\underline{H} = \underline{P}(\text{Atypical}) / (1 - \underline{R}) \qquad (4)$$

Although the present experiments were not principally concerned with dissociating recollection and habit, these studies nonetheless used Hay and Jacoby's framework in order to create conditions that were likely to produce varying magnitudes of age differences in a direct memory test. Both of the first two experiments employed sentences as stimuli because previous research (e.g., Brown, 1994; Shaw & Craik, 1989) suggested that reinstatement of rich, distinctive contexts might be one means of reducing age differences in memory.

Experiment 1

Experiment 1 employed a variant of Hay and Jacoby's (1996) process dissociation procedure in which sentence frames (e.g., "The canoeists met at a _____.") appeared a total of four times during a training phase. More specifically, the sentence frame appeared with one completion (e.g., "competition") on three training trials and with another, alternative completion (e.g., "contest") on the remaining one training trial for that sentence. At training, participants were instructed to try to remember both alternative completions for a memory test. Then, at study, participants saw one of the two trained completions for each sentence (either the once-presented completion, or the thrice-presented completion). Participants were now told to remember this one study completion for the memory test and to disregard the alternative completion from training. Within this procedure, the reinstatement of the thrice-trained completion at study serves as a facilitation condition, whereas the reinstatement of the once-trained completion serves as an interference condition. This is the case because, when study items were thrice-trained, both consciously controlled influences (from the study phase) and automatic influences (from the training phase) should prompt successful retrieval. In contrast, when study items were once-trained, consciously controlled influences should promote correct responding, but automatic influences (e.g., inappropriate retrieval of the thrice-trained item) should not.

However, the study items varied not only in their frequency of presentation at an earlier training phase, but also in their perceptual similarity to the training phase. More specifically, the voice used to present a particular sentence at study was the same voice, or a different voice, from that used at training. (Each sentence was spoken in the same voice across the four training trials so that any changes in voice occurred at study.) In summary, the experiment examined age differences in the effects of reinstating both perceptual and conceptual information from a prior learning phase. Note that in this study, the conceptual manipulation involves the item itself (i.e., the particular completion that was predominant for a sentence at training), whereas the perceptual manipulation involves the context (i.e., the particular voice – male or female – that was employed for a sentence at training).

It was hypothesized that age differences would be least when both the conceptual and perceptual information repeated at training was reinstated at study: That is, older and younger adults should show equivalently high memory performance for the Thrice-trained/Same-voice condition. This view is in line with Craik and Jennings' (1992) suggestion that age differences should "disappear" when sufficiently high levels of support are provided (barring ceiling effects). This prediction is also consistent with the finding that older and younger adults sometimes show equivalently high memory performance when meaningful study cues are precisely reinstated at test (e.g., Brown, 1994; and Shaw & Craik, 1989). However, this prediction should be regarded somewhat cautiously, given the predominant finding in the literature that most experimental manipulations in the literature affect older and younger adults equivalently and therefore fail to reduce age differences (see, e.g., Light, 1991, and a relevant commentary by Hay & Jacoby, 1999).

In contrast to the above prediction of age-equivalence under very supportive conditions, it was hypothesized that age differences would become more sizable with lesser reinstatement of perceptual and/or conceptual information from the training phase to the study phase. For example, age differences should be larger in the Thrice-trained/Different-voice condition than in the Thrice-trained/Same-voice condition. Moreover, for the Once-trained conditions (i.e., reappearance of the once-trained item at study), age differences should be even larger still. This is because older adults would show more interference (i.e., intrusions of thrice-trained responses) than their younger counterparts. However, it was somewhat unclear whether age differences would be larger for the Once-trained/Same-voice condition or the Once-trained/Different-voice condition. The former situation might actually yield more source confusions between the two learning phases of the experiment than might the latter situation and hence disproportionately penalize older adults (see, e.g., Spencer & Raz, 1995).

Furthermore, it was predicted that older and younger adults would perform equally poorly under minimally supportive conditions. For example, in addition to the critical conditions (described above) in which sentence frames appeared at training, study, and test, there were baseline conditions in which items did not appear in all phases of the experiment. It was predicted that one baseline condition – e.g., the pre-experimental baseline, in which items appeared for the first time at test – would yield equivalently poor performance for the two age groups. There was also a pure episodic baseline, in which items appeared only at study and test; and a training baseline, in which items appeared only at training and test. It was expected that age differences would be reliable in the episodic baseline, in line with most studies on aging and memory (see, e.g., Craik & Jennings, 1992, for a review). However, no age differences were expected in the training baseline, given the largely automatic nature of memory for repeated items (see also Hay & Jacoby, 1999).

Methods

Design. Figure 2.1 outlines the critical and baseline conditions employed in Experiment 1. For the critical or main experimental conditions, sentences appeared in all three phases of the experiment (training, study, and test). In these critical conditions, sentences appeared a total of four times at training (three times with one completion and once with an alternative completion), and then once at study with either the once-trained or thrice-trained completion, before re-appearing at test without the completion. The completion presented at study therefore served as the "target" or the to-be-remembered completion for the memory test. In addition, there was a voice manipulation: Each sentence was consistently presented in one of two voices (male or female) for all four training presentations. However, following training, the sentence was presented in either the same voice as had been used previously, or the alternative voice not used previously. In summary, there were four critical conditions that represented the crossing of two within-subjects factors, with two levels each: presentation frequency of target completions (the use at study of either the once-trained or thrice-trained completion), and voice congruency (the use at study and test of either the same voice used at training, or a different voice). There was also one between-subjects factor, age, with two levels (younger vs. older adults). The analyses of the critical conditions therefore correspond to a 2x2x2 (Age x Frequency of presentation at training x Voice congruency) design.

In addition to the critical conditions in which items appeared in all three experimental phases, there were three types of baseline condition: The first of these appeared in Hay and Jacoby (1996) and the latter two were unique to this experiment. First, some sentences appeared at training and test only: These items, designated "guessing" items by Hay and Jacoby, never appeared at study and therefore could not be recollected in the sense of remembering a single episode from the study phase. Therefore, participants' production of repeatedly trained but

Figure 2.1. Experiment 1: Overview of design and procedure

Training phase | **Study phase** | **Test phase**

Critical conditions:

Training phase: At training, critical sentences were presented 4 times each: 3 times with one completion and 1 time with an alternative completion. The two completions were similar in plausibility.

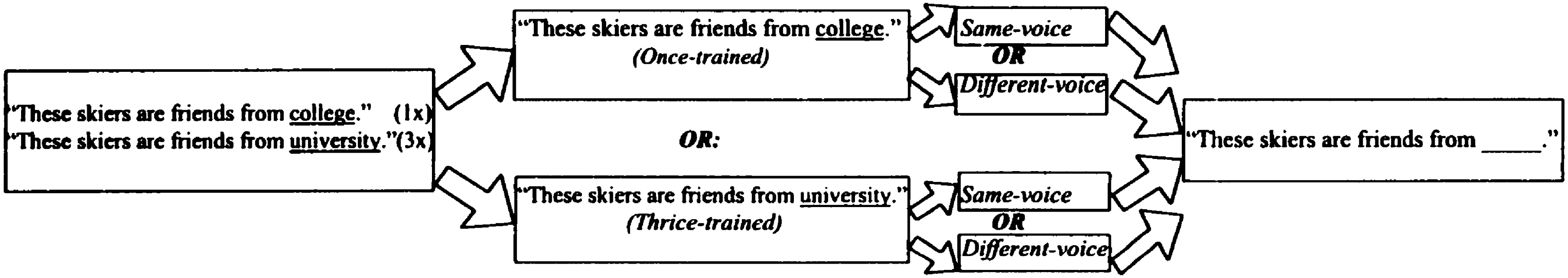
The same voice (either male or female) was used for a given sentence across training trials.

Study phase: At study, trained sentences were re-presented once only, with either the once-trained completion or the thrice-trained completion.

In addition, each sentence was presented in either the same voice used at training, or a different voice.

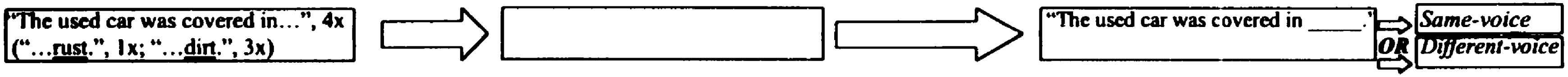
Test phase: At test, sentences were re-presented without any completions. These sentence frames served as retrieval cues for the studied completions.

Each sentence was presented in the same voice used at study.

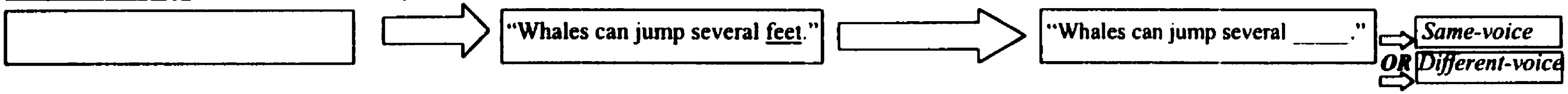


Baseline conditions:

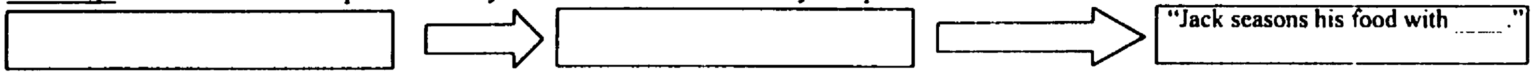
Training-and-test-only: Each sentence was presented 4 times at training (3 times with one completion, 1 time with another completion), and once again at test, in either the same voice used at training, or a different voice.



Study-and-test-only: Each sentence was presented once at study -- with one completion -- and then again at test, in either the same voice or a different voice.



Test-only: Each sentence was presented only once -- at test -- without any completions



unstudied items is thought to provide a baseline measure of training influences. This "guessing" base rate typically converges with other estimates of automatic memory influences – namely, the PDP estimates of habit and the probability at which typical completions appear at training (e.g., Hay & Jacoby, 1996). Second, some additional sentences appeared at study and test only: These items, which did not appear at training, provided a retrospective baseline for once-presented items, in the absence of prior learning. Third, a last group of sentences appeared for the first time at test, without completions, and therefore served as a baseline for pre-experimental influences (i.e., the probability of producing a target completion for a sentence without prior experimental exposure). Each participant received all three types of baseline condition.

The first two types of baseline item (training-and-test-only, and study-and-test-only) could both be further subdivided into two voice congruency conditions: either the same voice across experimental phases (i.e., from either training to test, or study to test), or a different voice from one phase to another. The third type of baseline item appeared only once – at test – and therefore could not be further subdivided by voice condition because there was no prior experimental context for these items. The analyses of the test-only condition therefore concern the variable Age, whereas the other two baseline conditions reflect a 2x2 (Age x Voice congruency) analysis. The conceptualization of the target or "correct" response also varied across the different baseline conditions: For the training-and-test-only items, the term "hits" referred to the proportion of typical (thrice-presented) items produced from the training phase; for the study-and-test-only baseline, it was the proportion of study items produced; and for the test-only condition, it was the proportion of "target" items produced (the target being one of the two experimenter-generated completions for each sentence that was arbitrarily designated as the correct answer in each experimental format). In summary, there were nine experimental conditions of interest: four critical or main conditions, in which items appeared in all three experimental phases (training, study, and test); and five baseline conditions, in which items appeared in two or fewer experimental phases.

A final note regarding the experimental design is that the target and alternative completions were presented varying numbers of times across the experimental conditions. Table 2.1 lists the total number of presentations for the two types of completion for each condition. Of particular interest was the fact the total number of target presentations (across the three experimental phases) ranged from 0 in the test-only condition to 4 in the thrice-trained (and studied) conditions. Moreover, all target frequencies between these two extremes were represented across the experimental conditions. Therefore, the experimental design lent itself to further presentation-frequency analyses that bridged across the critical and baseline conditions. These analyses corresponded to a 2x5 (Age x Total number of target presentations) design, with age varying between subjects and presentation condition within subjects.

Participants. Demographic information on the 24 younger and 24 older adults is presented in Table 2.2. Older adults were healthy volunteers who were offered reimbursement for their travel expenses, whereas younger adults were psychology undergraduates who received either course credit or $15 for their participation in a 2-hour session. On a 20-item version of the Mill Hill Vocabulary Scale (Raven, 1965), the older adults scored higher than the younger adults, $t(46) = 3.46$, $p < .01$. In contrast, on a perceptual speed measure – the WAIS Digit Symbol Substitution Task (Wechsler, 1981) – the younger adults outperformed the older adults, $t(46) = 8.28$, $p < .001$. The finding of an age-related increase in vocabulary, despite a decrease in perceptual speed, has been reported in other studies of aging and memory (e.g., Anderson, Craik, & Naveh-Benjamin, 1998). In years of formal education, the older adults were superior, $t(46) = 2.73$, $p < .01$, to the younger adults. However, the high educational attainment of the older adults in this sample was not considered problematic because the majority of younger adults were

Table 2.1

Experiments 1 and 2: Presentation Frequency of Target and Alternative Completions, By Experimental Condition

Experimental Condition	Frequency of target completion	Frequency of alternative completion
Test-only	0	0
Study-and-test-only	1	0
Once-trained	2	3
Training-and-test-only	3	1
Thrice-trained	4	1

Note. Frequency refers to the total number of presentations of either the target completion or the alternative experimental completion across all three experimental phases (training, study, and test). The once-trained and thrice-trained conditions are critical conditions (targets presented in all three experimental phases); the remaining conditions are baselines.

expected to complete their degrees and therefore come to resemble the older adults in educational status. The two age groups were equivalent in self-reported health, $t(46) < 1$, as assessed for two time frames: the last two to three months, and today. Both measures were on a 5-point scale, ranging from 1 ("excellent") to 5 ("poor").

Materials. A total of 147 sentence frames were generated to serve as experimental stimuli; each sentence frame was one word short of being a complete sentence (e.g., "These skiers are friends from _____."). These sentence frames were selected so as to be as thematically distinct from one another as possible. For each of these sentence frames, two alternative completions were generated by the experimenter that were plausible but somewhat unlikely to be generated without prior experimental exposure (e.g., "college" and "university" for the previous example). Both completions for a given sentence came from the same taxonomic category (e.g., educational institutions; see, e.g., Battig & Montague, 1969) and care was taken to employ a different category for each sentence. Of the total 147 sentence frames, 144 consistently served as critical items and 3 as practice items. The 147 sentence frames and their completions are listed in Appendix 2.1.

The 144 critical sentence frames and their alternative completions were divided into nine sets of 16 so that each participant was exposed to all nine conditions: The four critical conditions in which items appeared at training, study, and test; and the five baseline conditions in which items appeared in two or fewer of the experimental phases. With respect to the voice conditions, the nine sets were further subdivided in two so that all possible combinations of the two voices occurred equally often for each participant. For example, sentences in the Thrice-trained/Different-voice condition could be further subdivided into one of two presentation conditions: either a female voice at training and a male voice at study and test, or a male voice at training and a female voice at study and test. Similarly, sentences in the test-only condition

Table 2.2
Experiment 1: Mean Participant Characteristics

Variable	Age Group				t
	Young		Old		
	M (SD)	n	M (SD)	n	
Age	21.0 (1.3)	24	71.5 (5.2)	24	46.18***
Age range	19-25		62-81		-------
Gender:					-------
Female		21		19	
Male		3		5	
Handedness:					-------
Right-handed		21		23	
Left-handed		3		0	
Ambidextrous		0		1	
Education (years)	14.5 (1.0)		16.0 (2.3)		2.73**
Self-reported health:					
Last 2 to 3 months	1.9 (0.8)		1.9 (0.6)		0.20
Today	1.9 (0.8)		1.9 (0.6)		0.10
Mill Hill Vocabulary Scale	14.4 (1.6)		16.3 (2.3)		3.46**
Digit Symbol Substitution Task	69.5 (10.4)		44.3 (10.7)		8.28***

Note. For gender, $\chi^2_{(1, N=48)} = .60$, $p > .30$. For handedness, $\chi^2_{(2, N=48)} = 4.09$, $p > .10$. Self-reported health was assessed for two time frames and was rated on a 5-point scale, ranging from 1 (excellent) to 5 (poor). Mill Hill Vocabulary Scale maximum = 20. Digit symbol score is the number of correct completions in 90 sec (maximum = 93). t values are for younger versus older adults (df = 46); * = $p < .05$, ** = $p < .01$; *** = $p < .001$.

could be subdivided into conditions of either a female voice or a male voice at test. However, the large number of sentence frames employed (a total of 624 presented to each participant) made it impractical to produce the eighteen experimental formats required for exhaustive rotation of stimuli through all conditions. Therefore, four completely randomized experimental formats were generated so that in each format, all stimuli were equally likely to appear in any given condition. For each sentence in a given format, one of the two alternative completions was randomly selected to be the target completion and the other alternative to be the non-target completion.

Sentence frames for the training and study phases appeared on laminated sheets of paper kept in a three-ring binder. These sentence frames were quintuple-spaced, approximately six sentence frames per page, and appeared in bold print with a character size of approximately 3 x 4 mm, in an 18-point, Times New Roman font. A blank five characters long (i.e., "_____") was used to indicate that a word was missing from each sentence frame and that this word could change across stimulus presentations. This missing word was revealed in an auditory version of the stimulus that was presented at the same time as the visual version. No visual formats were produced for the test sentence frames, which appeared only auditorily so that their voice contexts might be emphasized. A paper mask (i.e., two pieces of white cardboard glued together) was fashioned with a rectangular hole the size of the largest sentence frame (approximately 2 x 18 cm) in order to display sentences one at a time. For all three experimental phases, audio recordings were prepared in which stimuli were spoken in either a male voice or a female voice. For the training and study phases, entire sentences were recorded, including the missing words from the printed version; for the test phase, sentence frames were recorded, with the word "blank" indicating the missing word in each case. Sentences and sentence frames were preceded by beeps which signaled that the experimenter should advance the paper mask from one stimulus to the next.

Procedure. Participants were tested individually and were seated approximately 30 cm in front of a desk with a three-ring binder elevated at a 30-degree angle. Participants were informed that they would be hearing a series of sentences on audio tape and that the same sentences would simultaneously be presented visually, except that one word would be missing in the visual version. That is, a blank or underline would appear on the page to indicate the missing word that would be presented auditorily. Participants were instructed that they should be able to provide this missing word when the sentence cue subsequently re-appeared at testing. In addition, they were informed that each sentence would have two possible combinations across the experimental trials. Participants were asked to remember both possible completions for the memory test a few minutes later. A short (12-item) practice list was then presented in which three different sentence cues were presented four times each: three times with one completion, and one time with another completion. Sentence frames appeared in a random order for 5 sec each, during which time the entire sentence including completion was presented auditorily. Then participants received a practice test in which they saw and heard the sentence cues and were instructed to produce both completions for each sentence cue. Participants had a total of 10 sec (until the next beep on the audio recording) to produce a response.

Following the practice phase, participants underwent a training phase in which they were exposed to 96 sentence cues, presented four times each: three times with one completion, and one time with another, alternative completion. Participants were once again informed that they should remember both completions for each sentence. As before, participants saw sentence frames (the entire sentence except for one missing word) while simultaneously hearing entire sentences, at the rate of one sentence every 5 sec. Sixty-four sentence cues corresponded to critical items (i.e., those appearing at training, study, and test), whereas the remaining 32 cues corresponded to "guessing" items (see Hay & Jacoby, 1996) appearing only at training and test. In summary, there were 384 training trials in which the item ordering was random, with the restriction that no fewer than three intervening trials appeared between repetitions of the same sentence.

Participants completed a demographic questionnaire before proceeding to the study phase. Participants were then informed that, instead of receiving a memory test for the previous sentences, they would now be receiving a new list of sentences. Participants were asked to disregard the previous list and instead remember the list that would follow. Moreover, they were told that each sentence would appear with only one completion and that this was the "real" completion to be remembered for the memory test. The study phase resembled the training phase in that sentences were presented visually while the entire sentence was presented auditorily, at a rate of one sentence every 4 sec. The study phase consisted of 96 sentence cues, presented once each, for a total of 96 study trials. Sixty-four of these sentence cues were critical items that appeared in all three experimental phases; the remaining 32 items appeared at study and test only and provided a baseline for once-presented items. Once again, the experimental conditions varied randomly across experimental trials.

The study phase was followed by four short interpolated tasks: the Digit Symbol Substitution task (Wechsler, 1981); a 20-item version of the Mill Hill Vocabulary Test (Raven, 1965); a version of the Stroop (1935) Color and Word Task; and the verbal fluency test (e.g., Borkowski, Benton, & Spreen, 1967). These interpolated tasks were used to collect additional measures from the participants while simultaneously guarding against ceiling effects. The interpolated tasks took no more than 12 minutes to complete.

Stroop task. The Stroop task was a paper version of a computerized blocked Stroop task previously employed by West and Baylis (1998). The task consisted of three parts or blocks, each corresponding to a different condition: a color-naming baseline, a word-reading baseline, and the critical interference condition. In each part, participants were instructed to respond as quickly and as accurately as possible. First, participants completed the color-naming baseline task, in which they received a page with two columns of 8 colored rectangles. These rectangles were 10x43 mm in size and were spaced 10 mm apart within each column, with a minimum of 22-mm separation between columns. The rectangles were filled with red, green, blue, or brown ink (four rectangles of each color) and were presented in a fixed random ordering. Participants were asked to name aloud the colors of all 16 rectangles. Secondly, participants completed the word-reading baseline task, in which they received a page containing two columns of 8 words in black ink. All words were printed in bold capital letters of size 6x6mm and were spaced 13 mm apart within each column, with a maximum of 42 mm separation between columns. The words were the color words red, green, blue, and brown, which occurred equally often and in a fixed random ordering within the two columns. Participants were instructed to read these words aloud. Finally, participants completed the interference task. They received a page on which the four color words were presented -- four times each -- in red, green, blue, or brown ink. The color words were each presented in an ink color incompatible with their meaning and were presented in a fixed random ordering. Participants were now asked to name the ink colors aloud, irrespective of the words' meaning.

For each part of the Stroop task, the experimenter recorded the participants' oral responses (i.e., the specific color names spoken) and their reaction times in completing each condition. The third, interference condition (i.e., naming ink colors incompatible with color words) was the main condition of interest, whereas the first and second conditions (i.e., naming colors of rectangles, and word-reading, respectively) provided baseline measures. As per West and Baylis (1998), the word-reading baseline was included because it was of some interest whether age differences would be smaller on this highly-practiced task than on the other tasks employed. This expectation was based on the assumption that word-reading is more automatic than color-naming (e.g., Posner & Snyder, 1975) and that automatic processes are relatively unaffected by the aging process (e.g., Hasher & Zacks, 1979). However, for the calculation of interference effects, the most meaningful comparison was between the incongruent color-naming (interference) condition and the color-naming baseline, as described by West and Baylis (1998).

Verbal fluency task. The verbal fluency task consisted of two parts: phonemic fluency (e.g., Borkowski et al., 1967), and semantic fluency (e.g., Newcombe, 1969). For phonemic fluency, participants were instructed to generate words that began with a particular letter (i.e., f, a, or s), excluding proper nouns and variants of the same word (e.g., the same root word but different suffixes). For semantic fluency, participants were instructed to generate the names of animals. Sixty seconds were allotted for each of the three phonemic trials and one semantic trial. The main measure of interest was the number of words generated, excluding errors and repetitions. For the phonemic fluency trials, the total number of words generated were summated across the three phonemic fluency trials for each participant. In addition, however, the number of perseverative errors (repetitions of the same word or root word) were scored. Furthermore, the verbal fluency data were subsequently analyzed in terms of clustering and switching performance, or the number of successive responses in a particular phonemic or semantic subcategory and the number of switches between subcategories (Troyer, Moscovitch, & Winocur, 1997). Troyer et al. (1997) argue that the clustering subscale principally reflects temporal-lobe processes, whereas the switching subscale provides a relatively pure assessment of frontal-lobe functioning (see also Neuropsychological analyses, below, for more information on scoring).

The interpolated tasks were followed by a test phase in which participants heard 144 sentence frames, presented once each, with the word "blank" indicating the missing word in each case. Participants were instructed to produce the word for each sentence that was presented during the study list (i.e., the last list of sentences that they saw and heard). They were further instructed that, if they could not remember any completion from the previous study list, then they should guess. The sentence cues were presented at a 10-sec rate, so the participant had until the next beep on the recording to make a response. The experimenter recorded all the participants' oral responses. Of the 144 sentence cues, 64 were critical items appearing in all three experimental phases and the remaining 80 were baseline items appearing in two or fewer experimental phases. Of these 80 baseline items, 32 were "guessing" items appearing only at training and test, an additional 32 were once-presented items appearing only at study and test, and a final 16 were new items appearing for the first time at test. The ordering of these conditions varied randomly across test trials. The experiment took approximately 100 minutes per individual session. The experimenter's instructions to participants are shown in Appendix 2.2.

Results

Cued recall in the critical conditions. Table 2.3 indicates participants' mean recall performance in the critical conditions, as a function of age group and memory measure. In these critical conditions, sentences appeared in all three phases of the experiment: training, study, and test. Four separate three-way (Age x Frequency of presentation at training x Voice congruency) analyses of variance (ANOVAs) were conducted for the various dependent measures of interest,

Table 2.3
Experiment 1: Cued Recall Performance for the Critical Conditions

Measure and age group	Condition (Presentation frequency/Voice congruency)							
	Once trained				Thrice trained			
	Different voice		Same voice		Different voice		Same voice	
	M	(SD)	M	(SD)	M	(SD)	M	(SD)
Hits								
Young	.75	(.19)	.75	(.21)	.89	(.13)	.89	(.10)
Old	.66	(.17)	.66	(.20)	.82	(.12)	.88	(.14)
FAs, experimental								
Young	.19	(.09)	.19	(.15)	.06	(.05)	.08	(.08)
Old	.22	(.11)	.24	(.14)	.06	(.06)	.07	(.07)
FAs, extra-experimental								
Young	.05	(.14)	.06	(.13)	.05	(.13)	.03	(.07)
Old	.10	(.12)	.08	(.09)	.08	(.10)	.04	(.09)
Non-responses								
Young	.01	(.03)	.01	(.02)	.003	(.01)	.003	(.01)
Old	.02	(.04)	.02	(.04)	.04	(.07)	.01	(.03)

Note. Critical conditions are those in which items appeared in all three experimental phases (training, study, and test). Voice congruency refers to the use, at study and test, of either the same voice used for a sentence at training, or the alternative voice not used at training. False alarms (FAs) were one of two types: experimental, or the alternative completion provided for a sentence at training; and extra-experimental, or any other response that did not correspond to an experimental completion. Non-responses refer to trials in which participants failed to make a response within the 10-sec deadline.

with each factor having two levels. Each ANOVA was conducted with age (younger vs. older adults) as a between-subjects variable and with frequency of presentation at training (once- vs. thrice-trained) and voice congruency (same voice vs. different voices at training and in subsequent phases) as within-subject variables. The first ANOVA was for hits, or the proportion of target completions correctly recalled. The next two ANOVAs corresponded to two different types of false alarms (FAs): experimental, and extra-experimental. Experimental FAs were trials in which the participant incorrectly produced the alternative completion provided for a sentence at training, whereas extra-experimental FAs were all other incorrect responses that did not correspond to an experimental completion. Finally, an ANOVA was conducted for non-responses, or the proportion of trials in which participants failed to respond.

The ANOVA for hits revealed a main effect of presentation frequency at training, $F(1,46) = 75.69$, $p < .001$, $MSE = 0.017$; and a marginally reliable effect of age, $F(1,46) = 3.30$, $p < .08$, $MSE = .062$. These main effects revealed that recall was higher when completions were thrice-trained, as compared to once-trained, and when participants were younger, as compared to older, adults. No other effects reached statistical significance.

Although the three-way interaction for hits was not reliable, $F(1,46) < 1$, $p > .24$, it was nonetheless of some theoretical interest to assess the reliability of age differences in each critical condition. Pairwise comparisons revealed that the age-related decline in recall was marginally reliable in the Once-trained/Different-voice condition, $t(46) = 1.82$, $p < .08$, was non-reliable in the Once-trained/Same-voice condition, $t(46) = 1.41$, $p > .16$, was marginally reliable in the Thrice-trained/Different-voice condition, $t(46) = 1.75$, $p < .09$, and was non-reliable in the Thrice-trained/Same-voice condition, $t(46) < 1$. In other words, there was a non-reliable trend toward greater age differences when different voices were employed across experimental phases than when the same voice was employed in the critical conditions.[2.1]

Moreover, although the Age x Frequency interaction was not reliable, $F(1,46) = 1.55$, $p = .22$, $MSE = 0.017$, it was also of theoretical interest to compare the reliability of the age difference in memory for once-trained completions, as compared to thrice-trained completions. As discussed previously, the once-trained target is a case of interference because it is opposed by an alternative completion on three training trials, whereas the thrice-trained target is a case of facilitation because it is opposed by a competing completion on only one training trial. Furthermore, a central hypothesis of the dissertation is that age differences in memory should be larger for interference conditions than for facilitation conditions. Consistent with this view, between-groups t-tests revealed that the older adults showed marginally worse performance than the younger adults in the once-trained condition, $t(46) = 1.82$, $p < .08$, but were equivalent to the younger adults in the thrice-trained condition, $t(46) = 1.39$, $p > .17$.

The ANOVA for experimental FAs revealed as its only effect a main effect of presentation frequency, $F(1,46) = 73.56$, $p < .001$, $MSE = 0.013$. This effect indicated that there were more intrusions of the alternative experimental completion in the once-trained condition than in the thrice-trained condition.

The ANOVA for extra-experimental FAs revealed a main effect of presentation frequency, $F(1,46) = 5.87$, $p < .05$, $MSE = 0.004$; and a marginally reliable effect of voice congruency, $F(1,46) = 3.86$, $p < .06$, $MSE = 0.004$. These main effects revealed that there were more intrusions of non-experimental completions when completions were once-trained, rather than thrice-trained, and when different voices were used at training and at subsequent

[2.1] The assessment of age differences through t-tests, rather than a priori contrasts, should be regarded with some caution (K. Dion, personal communication, October 1999). Nonetheless, the t-statistic is used both here and elsewhere in the dissertation as a relatively simple means of making between-group comparisons.

experimental phases, rather than the same voice across all experimental phases. No other effects reached statistical significance.

The ANOVA for the probability of not responding revealed only a main effect of age, $F(1,46) = 7.01$, $p < .05$, $MSE = 0.002$. This effect indicated that the older adults failed to respond on a greater proportion of trials than did the younger adults.

Cued recall in the baseline conditions. Table 2.4 indicates participants' mean recall performance in the baseline conditions, as a function of age group and memory measure. Separate analyses were conducted for each of the three main baseline conditions: test-only, study-and-test-only, and training-and-test-only. For each baseline condition, four ANOVAs were conducted, or one for each of the four dependent measures described above: hits, experimental FAs, extra-experimental FAs, and non-responses. Age was a factor for all three baseline conditions (younger vs. older adults). Voice congruency was a factor for the two baseline conditions in which sentences appeared in two experimental phases: In these two conditions, sentences were presented in either the same voice across experimental phases (i.e., from study to test, or from training to test), or in different voices from one phase to another.

For the test-only condition, a one-way ANOVA (Age) was conducted for each of the four dependent measures. Voice congruency was not included as a factor because, in this condition, there was no prior voice context with which to compare the voice context at testing. Moreover, in this "pre-experimental" baseline, neither of the two experimenter-generated completions appeared during the experiment: Therefore, one completion was arbitrarily designated the target completion and the other was designated the alternative completion. In summary, a response was scored as a hit if it corresponded to one of the experimenter-generated completions (arbitrarily deemed the target) and an experimental FA if it corresponded to the alternative experimenter-generated completion. In the test-only condition, the younger adults produced

Table 2.4
Experiment 1: Cued Recall Performance for the Baseline Conditions

Measure and age group	**Condition *(Presentation frequency/Voice congruency)***									
	Test only		**Study and test only**				**Training and test only**			
			Different voice		Same voice		Different voice		Same voice	
	M	(SD)	M	(SD)	M	(SD)	M	(SD)	M	(SD)
Hits										
Young	.10	(.06)	.68	(.22)	.65	(.23)	.71	(.17)	.73	(.19)
Old	.07	(.05)	.53	(.22)	.46	(.15)	.69	(.13)	.67	(.13)
FAs, experimental										
Young	.10	(.10)	.04	(.05)	.01	(.02)	.25	(.13)	.21	(.15)
Old	.06	(.06)	.04	(.05)	.04	(.06)	.16	(.09)	.19	(.11)
FAs, extra-experimental										
Young	.66	(.16)	.25	(.20)	.29	(.23)	.04	(.11)	.06	(.14)
Old	.57	(.25)	.29	(.15)	.34	(.17)	.11	(.09)	.09	(.09)
Non-responses										
Young	.14	(.15)	.03	(.05)	.06	(.07)	.00	(.00)	.003	(.01)
Old	.30	(.27)	.14	(.16)	.16	(.17)	.03	(.06)	.05	(.07)

Note. Baseline conditions are those in which a sentence appeared in fewer than three experimental phases. Voice congruency refers to the use at test of either the same voice used previously for a sentence, or the alternative voice not used previously. False alarms (FAs) were one of two types: experimental, or the alternative experimental completion used for a sentence (although not necessarily for that participant); and extra-experimental, or any other response that did not correspond to an experimental completion. For the study-and-test-only condition, the alternative experimental completion did not appear during the experiment. For the test-only condition, neither of the two experimenter-generated completions appeared during the experiment: Therefore, one completion was arbitrarily designated the target completion and the other was designated the alternative completion.

more target completions than the older adults, $F(1,46) = 4.10$, $p < .05$, $MSE = 0.003$, and also produced more alternative experimenter-generated completions, than the older adults, $F(1,46) = 3.73$, $p = .06$, $MSE = 0.006$. However, the two age groups produced equivalent numbers of non-experimental completions in the test-only condition, $F(1,46) = 2.17$, $p > .14$, $MSE = 0.043$. On the other hand, the older adults were more likely than the younger adults to fail to respond to new items at test, $F(1,46) = 7.03$, $p < .05$, $MSE = 0.046$.

For the study-and-test-only condition, a two-way ANOVA (Age x Voice congruency) was conducted for each of the four memory measures. In this episodic baseline for once-presented items, voice congruency refers to the use of either the same voice or different voices for a sentence from study to test. Moreover, in this baseline condition, the alternative experimental completion for a sentence never appeared during the experiment. The ANOVA for hits revealed main effects of age, $F(1,46) = 9.45$, $p < .01$, $MSE = 0.074$; and of voice congruency, $F(1,46) = 5.87$, $p < .05$, $MSE = 0.013$. These effects indicated that recall of once-presented items was greater for younger than for older adults and, surprisingly, for the different-voice condition than for the same-voice condition. The ANOVA for experimental FAs revealed two marginally reliable effects: a main effect of age, $F(1,46) = 3.70$, $p < .07$, $MSE = 0.002$; and an Age x Voice congruency interaction, $F(1,46) = 2.90$, $p < .10$, $MSE = 0.002$. The main effect suggested that intrusions of the never-presented alternative experimental completion were more likely in older than in younger adults. However, tests of simple effects for the interaction revealed that this age-related increase in experimental FAs was reliable for the same-voice condition, $t(46) = 2.72$, $p < .01$, but not for the different-voice condition, $t(46) < 1$. Further tests of simple effects revealed that the younger adults produced more experimental FAs in the different-voice condition than in the same-voice condition, $t(46) = 3.11$, $p < .01$, but that the older adults produced equivalent numbers of experimental FAs in these two voice congruency conditions, $\underline{t}(46) = 0.00$, $\underline{p} = 1.00$. The ANOVA for extra-experimental FAs revealed only a main effect of voice congruency, $\underline{F}(1,46) = 6.60$, $\underline{p} < .05$, $\underline{MSE} = 0.009$, indicating that intrusions of non-experimental completions were more likely in the same-voice condition of the episodic baseline than in the different-voice condition. The ANOVA for non-responses revealed only a main effect of age, $\underline{F}(1,46) = 9.97$, $\underline{p} < .01$, $\underline{MSE} = 0.027$, indicating that older adults were less likely than younger adults to respond to once-presented (studied but untrained) items.

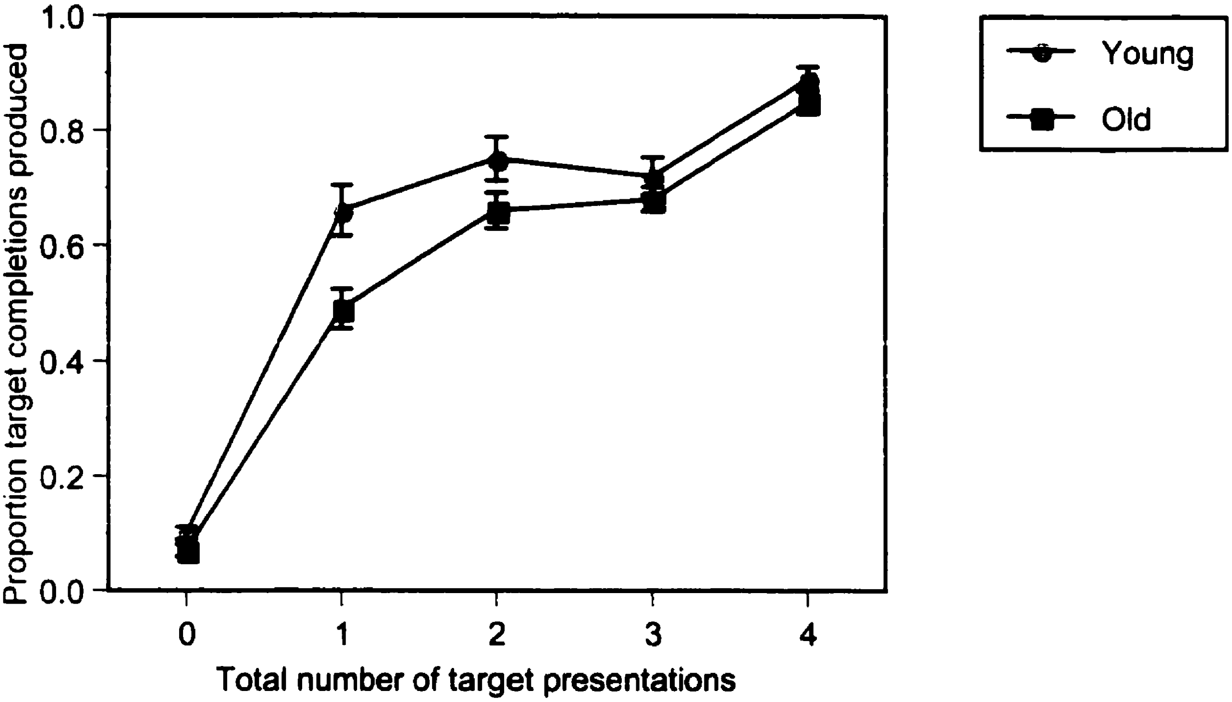
For the training-and-test-only condition, a two-way ANOVA (Age x Voice congruency) was once again conducted for each of the four memory measures. In this "guessing" baseline assessing training influences, voice congruency refers to the use of either the same voice or different voices for a sentence from training to test. A response was scored as a hit if it corresponded to the repeatedly trained or "target" completion for a sentence (presented three times at training) but was scored as an experimental FA if it corresponded to the alternative experimental completion (presented only once at training). The ANOVA for hits failed to reveal any reliable effects, either for age, $\underline{F}(1,46) = 1.07$, $\underline{p} > .30$, $\underline{MSE} = 0.037$; or for voice congruency or the interaction of these two variables ($\underline{F}$s < 1). That is, both age groups showed equivalent recall of repeatedly trained but unstudied items, and this measure was unaffected by voice context. The ANOVA for experimental FAs revealed two marginally reliable effects: a main effect of age, $\underline{F}(1,46) = 3.39$, $\underline{p} < .08$, $\underline{MSE} = 0.019$; and an Age x Voice congruency interaction, $\underline{F}(1,46) = 3.00$, $\underline{p} < .10$, $\underline{MSE} = 0.009$. The main effect suggested that the older adults were less likely than the younger adults to produce the alternative, infrequently trained completion in the training baseline. However, tests of simple effects for the interaction revealed that the age decline in the production of infrequently trained completions was reliable for the different-voice condition, $\underline{t}(46) = 2.74$, $\underline{p} < .01$, but not for the same-voice condition, $\underline{t}(46) < 1$. With further regard to this interaction, the older adults showed a non-reliable trend toward greater experimental FAs in the different-voice condition than in the same-voice condition, whereas younger adults showed a non-reliable trend in the opposite direction, $ts < 1.48$, $ps > .15$. The ANOVA for extra-experimental FAs also revealed a marginally reliable effect of age, $F(1,46) = 3.24$, $p < .08$, $MSE = 0.021$; and a two-way interaction, $F(1,46) = 3.58$, $p < .07$, $MSE = 0.003$. The main effect suggested that intrusions of non-experimental completions in the training baseline were more likely in older adults than in younger adults. However, tests of simple effects revealed that this age-related increase in extra-experimental responses was reliable for the different-voice condition, $t(46) = 2.55$, $p < .05$, but not for the same-voice condition, $t(46) = 1.00$, $p > .32$. Further tests of simple effects revealed that the likelihood of producing extra-experimental FAs was equivalent across voice congruency conditions for the younger adults, $t(46) = 1.70$, $p > .10$, as well as the older adults, $t(46) = 1.10$, $p > .28$. Finally, the ANOVA for non-responses revealed only a main effect of age, $F(1,46) = 11.67$, $p < .01$, $MSE = 0.003$, indicating that the older adults were less likely to respond to trained, but unstudied items, than were the younger adults.

Cued recall as a function of total presentation frequency. An alternative way of analyzing the cued recall data involves examining memory performance as a function of the total number of target presentations. In other words, across the critical and baseline conditions, the target completion was presented a varying number of times during the experiment, ranging from 0 presentations across the experimental phases (i.e., the test-only condition) to 4 presentations (i.e., the thrice-trained condition; see Table 2.1). This type of analysis therefore provides a way of bridging across the critical and baseline conditions. It was expected that the age difference in recall of targets would be reduced with greater numbers of stimulus repetitions, in line with Craik and Jennings' (1992) suggestion that highly supportive conditions should reduce age differences in memory. In contrast, it was expected that there would be an age-related increase in experimental FAs, or greater intrusions of the alternative experimental completion for older than younger adults, particularly when the target was presented for 2 presentations (i.e., the once-trained condition). In this once-trained condition, the alternative completion was presented a total of 3 times at training and was therefore thought to produce the greatest interference of the experimental conditions. Figure 2.2 indicates the mean probabilities of producing the target completion and the alternative experimental completion, as a function of the total number of target presentations. The mean probabilities of producing extra-experimental responses and non-responses are not depicted, as these dependent measures were of lesser theoretical interest.

The analyses for presentation frequency consisted of four two-way ANOVAs, one for each of the memory measures described above: hits, experimental FAs, extra-experimental FAs, and non-responses. In each ANOVA, age served as a between-subjects factor with two levels (younger vs. older adults) and frequency of target presentation as a within-subjects factor with five levels (0 vs. 1 vs. 2 vs. 3 vs. 4 target presentations). Voice congruency was not included as a factor in these analyses because this factor did not apply to the test-only condition (0 presentations), in which there is no prior experimental exposure to either the sentence or its voice context. Of principal interest in these 2x5 (Age x Presentation frequency) ANOVAs was whether the magnitude of age differences varied with the total number of target presentations.

The ANOVA for hits revealed main effects of age, $F(1,46) = 6.62$, $p < .05$, $MSE = 0.050$; and frequency of target presentation, $F(4,184) = 365.20$, $p < .001$, $MSE = 0.012$; as well as an Age x Presentation frequency interaction, $F(4,184) = 3.44$, $p < .05$, $MSE = 0.012$. The effect of age occurred because the younger adults had more hits, overall, than did the older adults. Post-hoc Tukey HSD tests were conducted to assess the nature of the main effect of frequency: These analyses revealed that the number of hits increased from 0 to 2 prior presentations, remained constant from 2 to 3 presentations, and then increased again from 3 to 4 presentations. Tests of

A. **Hits**



B. **Experimental False Alarms**

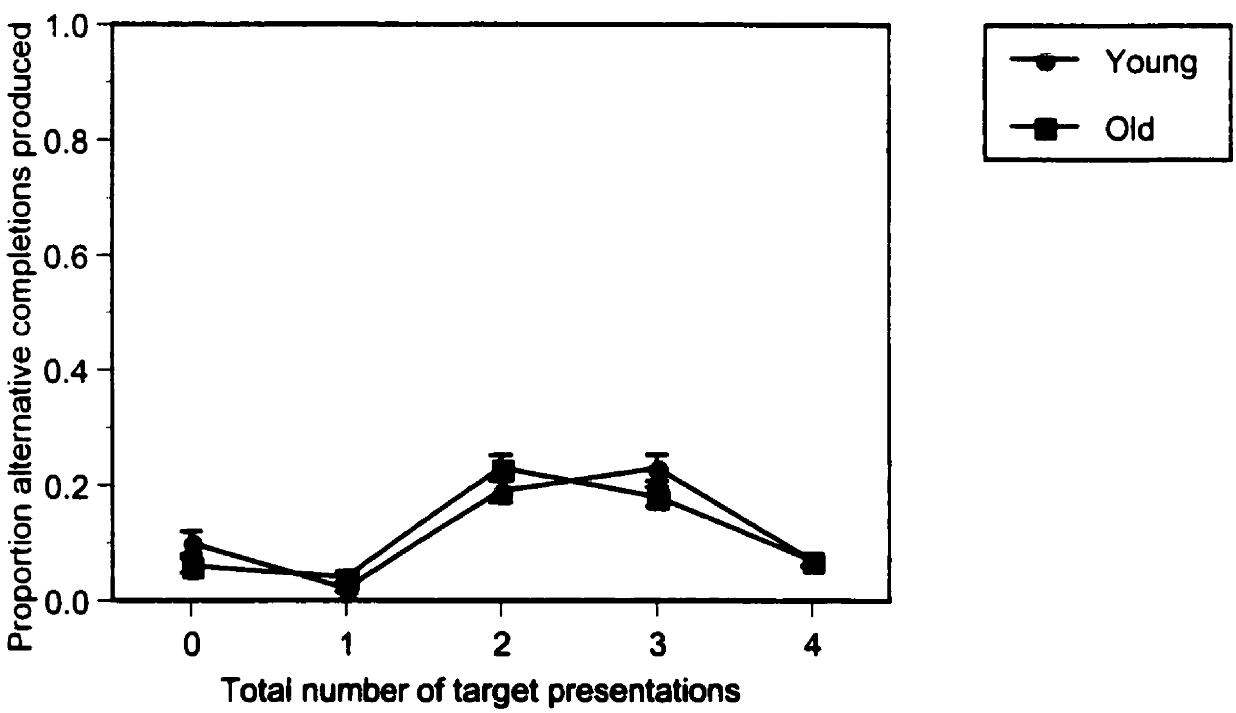


Figure 2.2. Experiment 1: Mean recall performance, as a function of memory measure, age group, and total number of target presentations. Error bars of one standard error are plotted for each point but are only visible when greater in magnitude than the size of the symbol.

simple effects were performed to assess the nature of the two-way interaction. First, t-tests were conducted to assess the reliability of age differences for each frequency level: These analyses revealed that the age-related decline in hits was reliable at 0 target presentations, $t(46) = 2.03$, $p < .05$, and at 1 presentation, $t(46) = 3.07$, $p < .01$; and was marginally reliable at 2 presentations, $t(46) = 1.82$, $p < .08$; but was not reliable for either 3 presentations, $t(46) = 1.03$, $p > .30$, or 4 presentations, $t(46) = 1.39$, $p > .17$. Secondly, two separate one-way ANOVAs assessed the effect of frequency for each group: A main effect was obtained for the younger adults, $F(4,92) = 151.93$, $p < .001$, $MSE = 0.015$, as well as for the older adults, $F(4,92) = 237.66$, $p < .001$, $MSE = 0.009$. Tukey tests were conducted to assess the nature of these simple main effects. For the younger adults, the number of hits increased from 0 to 1 presentations, remained constant from 1 to 3 presentations, and then increased again from 3 to 4 presentations. For the older adults, the number of hits increased from 0 to 2 presentations, remained constant from 2 to 3 presentations, and then increased again from 3 to 4 presentations. In summary, these analyses for hits suggest that both age groups benefited from stimulus repetition, but that the older adults required more repetitions before they showed the same degree of benefit. Moreover, the older adults were equivalent to the younger adults at 3 and 4 stimulus repetitions.

The ANOVA for experimental FAs revealed a main effect of frequency of target presentation, $F(4,184) = 54.23$, $p < .001$, $MSE = 0.006$; and an Age x Presentation frequency interaction, $F(4,184) = 2.98$, $p < .05$, $MSE = 0.006$. Post-hoc Tukey HSD tests revealed that the main effect occurred because there were more experimental FAs when the target was presented 2 or 3 times than when the target was presented 0, 1, or 4 times; moreover, the number of experimental FAs remained constant from 2 to 3 target presentations. Furthermore, there were significantly more experimental FAs when the target was presented 0 times, as compared to 1 time. (The number of experimental FAs in the 4-presentations condition did not vary reliably from that in either the 0- or 1-presentations conditions.) Tests of simple effects assessed the nature of the Age x Presentation frequency interaction. First, pairwise comparisons assessing the reliability of age differences revealed that the older adults showed more experimental FAs than the younger adults at 1 target presentation, $t(46) = 1.92$, $p < .07$, but showed fewer experimental FAs than the younger adults in two other frequency conditions: 0 presentations, $t(46) = 1.93$, $p < .07$, and 3 presentations, $t(46) = 1.84$, $p < .08$. There were no age differences in experimental FAs either at 2 presentations, $t(46) = 1.19$, $p > .23$, or at 4 presentations, $t(46) < 1$. Secondly, two separate one-way ANOVAs assessed the effect of frequency for each age group: A main effect was obtained for the younger adults, $F(4,92) = 26.54$, $p < .001$, $MSE = 0.007$, as well as the older adults, $F(4,92) = 31.23$, $p < .001$, $MSE = 0.005$. Tukey tests were conducted to assess the nature of these simple main effects. These analyses revealed that, in most respects, each age group showed the same pattern of pairwise differences that was obtained for the effect of frequency in the overall two-way (Age x Presentation Frequency) ANOVA. Most importantly, both age groups showed more experimental FAs when the target was presented 2 or 3 times than when the target was presented 0, 1, or 4 times. However, the two age groups differed in that the younger adults produced reliably more experimental FAs in the 0-presentations condition than in the 1-presentation condition, whereas the older adults produced equivalent numbers of experimental FAs in these two frequency conditions. In summary, there was little evidence for an age-related increase in experimental FAs across most presentation conditions.

The ANOVA for extra-experimental FAs revealed a main effect of frequency of target presentation, $F(4,184) = 223.43$, $p < .001$, $MSE = 0.013$; and an Age x Presentation frequency interaction, $F(4,184) = 3.25$, $p < .05$, $MSE = 0.013$. Tukey tests revealed that the main effect occurred because the proportion of extra-experimental FAs decreased from 0 target presentations ($M_{0x} = .62$) to 2 target presentations ($M_{2x} = .07$) and then remained constant from 2 presentations to 4 presentations ($\underline{M}_{4x}$ = .05). Tests of simple effects assessed the nature of the interaction. First, pairwise comparisons revealed that the two-way interaction occurred because there were no reliable age differences in extra-experimental FAs in any condition (all $\underline{t}$s < 1.48, $\underline{p}$s > .14) except in the 3-presentations condition, where the older adults were marginally more likely to produce extra-experimental FAs ($\underline{M}_{old}$ = .10) than were the younger adults ($\underline{M}_{young}$ = .05), $\underline{t}(46)$ = 1.80, p < .08. Secondly, two separate one-way ANOVAs assessed the effect of presentation frequency for each age group: A main effect was obtained for the younger adults, $\underline{F}(4,92)$ = 140.58, $\underline{p}$ < .001, $\underline{MSE}$ = 0.012, as well as the older adults, $\underline{F}(4,92)$ = 87.51, $\underline{p}$ < .001, $\underline{MSE}$ = 0.013. Tukey tests revealed that each age group showed the same pattern of pairwise differences that was obtained for the effect of frequency in the two-way ANOVA: That is, the number of extra-experimental FAs decreased from 0 to 2 target presentations and then remained constant from 2 to 4 presentations. In short, there was little evidence for an age-related increase in extra-experimental FAs.

The ANOVA for non-responses revealed main effects of age, $\underline{F}(1,46)$ = 10.74, $\underline{p}$ < .01, MSE = 0.027, and of presentation frequency, $\underline{F}(4,184)$ = 41.88, $\underline{p}$ < .001, $\underline{MSE}$ = .009; as well as an interaction of these two variables, $\underline{F}(4,184)$ = 5.50, $\underline{p}$ < .001, $\underline{MSE}$ = .009. The effect of age occurred because the older adults failed to respond on a greater proportion of trials ($\underline{M}_{old}$ = .11) than did the younger adults ($\underline{M}_{young}$ = .04). Tukey tests revealed that the effect of presentation frequency occurred because the likelihood of not responding decreased monotonically from 0 target presentations ($\underline{M}_{0x}$ = .22) to 2 target presentations ($\underline{M}_{2x}$ = .01) and then remained constant from 2 presentations to 4 presentations ($\underline{M}_{4x}$ = .01). Tests of simple effects assessed the nature of the Age x Presentation frequency interaction. First, pairwise comparisons revealed that the two-way interaction occurred because the age-related increase in non-responses was statistically reliable in all presentation conditions (all $\underline{t}$s > 2.10, $\underline{p}$s < .05), but was numerically the largest for

0 presentations (M_{old-0x} = .30 vs. $M_{young-0x}$ = .14) and 1 presentation (M_{old-1x} = .15 vs. $M_{young-1x}$ = .05). Secondly, two separate one-way ANOVAs assessed the effect of frequency for each group: A main effect was obtained for the younger adults, $F(4,92) = 18.50$, $p < .001$, $MSE = 0.004$, as well as for the older adults, $F(4,92) = 25.31$, $p < .001$, $MSE = 0.014$. Tukey tests were conducted to assess the nature of these simple main effects. For the younger adults, the number of non-responses decreased from 0 target presentations ($M_{young-0x}$ = .14) to 1 target presentation ($M_{young-1x}$ = .05) and then remained constant from 1 presentation to 4 presentations ($M_{young-4x}$ = .003). For the older adults, the number of non-responses decreased monotonically from 0 target presentations (M_{old-0x} = .30) to 2 target presentations (M_{old-2x} = .02) and then remained constant from 2 presentations to 4 presentations (M_{old-4x} = .03). In summary, although both age groups showed fewer non-responses with increasing target presentations, there was an age-related increase in non-responses which was reliable across all presentation conditions.

Estimates of recollection and automatic influences. Table 2.5 reports mean estimates of recollection and automatic influences by age group and voice congruency condition. Two separate two-way (Age x Voice congruency) ANOVAs were conducted, one for recollection and one for habit. Each ANOVA was conducted with age as a between-groups variable and with voice congruency as a within-subjects variable.

The ANOVA for recollection failed to reveal any reliable effects, either for age, $F(1,46) = 2.69$, $p < .11$, $MSE = 0.053$; or for voice congruency or the interaction of these two factors (Fs < 1). The ANOVA for habit also failed to reveal any reliable main effects (Fs < 1) or interaction, $F(1,46) = 2.34$, $p > .13$, $MSE = 0.069$. In summary, both age groups showed equivalent recollection and habitual influences for repeated sentences, and this pattern was not affected by voice context. Although age-invariance in habit has been found in several other PDP studies (see

Table 2.5
Experiment 1: Estimates of Recollection and Automatic Influences

Estimate and age group	**Condition *(Voice Congruency)***			
	Different Voice		Same Voice	
	M	(SD)	M	(SD)
Recollection				
Young	.70	(.18)	.70	(.16)
Old	.60	(.19)	.64	(.20)
Habit				
Young	.69	(.25)	.63	(.32)
Old	.56	(.25)	.67	(.23)
"Guessing"				
Young	.73	(.19)	.71	(.17)
Old	.67	(.13)	.69	(.13)

Note. Recollection and habit were derived based on equations provided by Hay and Jacoby (1996). "Guessing," in Hay and Jacoby's terminology, refers to performance in the training-and-test-only condition (i.e., recall of items repeated at training but not at study) and is expected to correspond with derived estimates of habit.

Jacoby et al., 1997, for a review), the present finding of age-invariance in recollection is relatively unique among PDP studies. Nonetheless, there is a non-reliable age decrement (approximately .08) in recollection in the present study, which is numerically larger than the non-reliable age decrements observed for either the habit estimates or the "guessing" scores (approximately .04).

With further regard to automatic influences, an additional analysis sought converging evidence for the derived estimates of habit. As described above, both the "guessing" scores (recall in the training-and-test-only conditions) and the derived habit estimates were assumed to measure the same automatic influences induced in the training phase (Hay & Jacoby, 1996). Therefore, a three-way (Measure x Age x Voice congruency) MANOVA was conducted, with each factor having two levels each. Age served as a between-subjects factor, whereas measure (habit vs. "guessing") and voice congruency served as within-subjects factors. This analysis revealed a main effect of measure, $F(1,46) = 6.68$, $p < .05$, $MSE = 0.028$, indicating the "guessing" scores were higher than the derived estimates of habit; however, no other effects attained statistical significance. Interestingly enough, a similar difference between habit and "guessing" was also found in Hay and Jacoby's (1999) Experiment 1, but not in their subsequent experiments. However, despite this difference in scores, it is noteworthy that both estimates of automatic influences yielded similar results – that is, neither measure showed an effect of age, voice context, or the interaction of these two factors.

Neuropsychological task performance. Table 2.6 indicates participants' performance on the two neuropsychological tasks (the verbal fluency task and the Stroop task), as a function of age group and performance measure.

In the Stroop task, there were two baseline conditions (reading color words presented in black ink, and naming the colors of colored rectangles) and one interference condition (naming

Table 2.6
Experiment 1: Neuropsychological Task Performance

Measure	**Age Group**				t
	Young		Old		
	M	(SD)	M	(SD)	
Stroop Task:					
Word-Reading Baseline					
RT	5.49	(0.92)	5.90	(1.07)	1.42
Accuracy	1.00	(0.00)	1.00	(0.00)	-----
Color-Naming Baseline					
RT	7.28	(1.65)	8.58	(1.84)	2.58*
Accuracy	0.995	(0.02)	1.00	(0.00)	-----
Interference *(Incongruent)*					
RT	11.86	(3.04)	19.99	(4.47)	7.38***
Accuracy	0.99	(0.03)	0.93	(0.10)	2.53*
Verbal Fluency Task:					
Phonemic					
Total Generated	42.29	(9.84)	43.63	(12.35)	0.41
Perseverative Errors	0.58	(0.78)	0.92	(1.28)	1.09
Clustering	0.30	(0.13)	0.42	(0.21)	2.41*
Switching	31.21	(8.37)	29.54	(10.15)	0.62
Semantic					
Total Generated	20.67	(5.35)	17.96	(4.97)	1.82⁺
Perseverative Errors	0.42	(0.78)	0.38	(0.58)	0.21
Clustering	1.28	(0.94)	1.61	(0.72)	1.38
Switching	9.13	(3.26)	6.38	(2.16)	3.44**

Note. Reaction times (RTs) for the Stroop task are the number of seconds taken to either name 16 colors or read 16 words. The clustering and switching measures for the verbal fluency task refer to the mean cluster size and number of shifts between subcategories, respectively (see Troyer et al., 1997). t values are for younger versus older adults ($df = 46$); $^{+} = p < .10$, $* = p < .05$, $** = p < .01$, $*** = p < .001$.

the ink colors in which color words were presented). For the baseline conditions, the older adults were as fast as the younger adults when reading words in black ink, $t(46) = 1.42$, $p > .15$, but were slower than the younger adults when naming colors of rectangles, $t(46) = 2.58$, $p < .05$. Both age groups showed perfect accuracy in these two conditions, except for one error made by each of two younger adults in the color-naming baseline. For the interference condition, or the main condition of interest, the older adults were slower, $t(46) = 7.37$, $p < .001$, and less accurate, $t(46) = 2.53$, $p < .05$, in naming color words presented in incompatible ink colors than were the younger adults. With further regard to Stroop performance, an interference effect was calculated by computing the absolute difference in RT between the incongruent color-naming (interference) condition and the color-naming baseline: This analysis revealed that the older adults showed a greater increase in RT ($M = 11.42$ sec, $SD = 3.76$) than did the younger adults ($M = 4.58$ sec, $SD = 2.73$), $t(46) = 7.21$, $p < .001$. An additional analysis assessed Stroop interference in relative, rather than absolute, terms, given some researchers' suggestions (e.g., Salthouse & Meinz, 1995) that an age-related increase in interference may result from general slowing with increasing age. Therefore, a proportionate interference ratio (i.e., [incongruent color-naming – color-naming baseline] / color-naming baseline) was constructed for each participant, as described by West and Baylis (1998): An analysis of these ratios revealed that interference continued to be greater for the older adults ($M = 1.37$, $SD = 0.47$) than for the younger adults ($M = 0.66$, $SD = 0.39$), $t(46) = 5.73$, $p < .001$.[2.2] This finding of greater Stroop interference in older than younger adults has been reported elsewhere in the literature, where it has been taken as evidence for age-related

[2.2] The means of the Stroop interference ratios for individual participants deviate slightly from the interference ratios based on group means. This deviation of group means from the means of individual participants' means has been reported by other researchers. For example, Jennings and Jacoby (1993) reported that their PDP estimates varied slightly depending on whether these estimates were calculated based on a group's mean inclusion and exclusion performance, or the inclusion and exclusion performance of individual participants.

deterioration of the frontal lobes (e.g., Moscovitch & Winocur, 1992; but see, e.g., Verhaeghen & De Meersman, 1998, for an alternative interpretation).

In the verbal fluency task, participants generated words either beginning with a particular letter (i.e., phonemic fluency) or belonging to a specific category (i.e., semantic fluency). For phonemic fluency, the older adults generated as many words, $t(46) < 1$, as the younger adults. For semantic fluency, the older adults generated marginally fewer words, $t(46) = 1.82$, $p < .08$, than the younger adults. The finding of smaller age differences in phonemic fluency than in semantic fluency has been reported elsewhere in the literature (e.g., Troyer et al., 1997). The two age groups showed equivalent numbers of perseverative errors for both phonemic and semantic fluency ($ts < 1$): This age-equivalence in perseveration is likely due to the low numbers of errors committed on this task.

With further regard to the verbal fluency task, additional measures of clustering and switching were derived for both phonemic and semantic fluency, based on the work of Troyer et al. (1997). The specific subscales of clustering and switching are thought to correlate with temporal lobe and frontal-lobe functioning, respectively (Troyer et al., 1997). Cluster size refers to the mean number of successive responses belonging to the same phonemic or semantic subcategory (e.g., words sharing the same initial two letters, or animals from the same geographic area; see Troyer et al. for a complete list of categories). Switches refer to the number of shifts between phonemic or semantic clusters within a participant's verbal fluency protocol. For phonemic fluency, cluster size was larger for older than younger adults, $t(46) = 2.41$, $p < .05$, but number of switches was equivalent for the two age groups, $t(46) < 1$. For semantic fluency, cluster size was equivalent for the two age groups, $t(46) = 1.38$, $p > .17$, but number of switches was greater for younger than older adults, $t(46) = 3.44$, $p < .01$. These results are entirely consistent with Troyer et al.'s findings. Namely, for phonemic fluency, there were age differences in clustering favoring the old, but no age differences in switching. Moreover, for semantic fluency, there were no age differences in clustering but age differences favoring the young for switching.

Correlational analyses. Table 2.7 shows selected correlations among variables for younger and older adults. In these correlational analyses, four memory measures were of principal theoretical interest. First, recollection was the PDP-derived estimate of consciously controlled memory influences in the critical experimental conditions. Secondly, automaticity was assessed by taking the mean of two z-score-transformed estimates of automatic memory influence (habit and "guessing"): These z-score transformations were carried out due to possible scaling differences for habit and "guessing" and were conducted separately for each age group. Thirdly, episodic influences for once-presented items was assessed by calculating the proportion hits in the study-only condition, which served as an episodic baseline. Fourth, intrusions are the number of thrice-presented completions erroneously recalled when the alternative, once-presented completion was the target. It was hypothesized that demographic variables such as age, education, vocabulary, and perceptual speed might correlate with these memory measures. For instance, one might reasonably expect chronological age to correlate negatively with recollection and episodic influences but show no correlation with automatic influences (see, e.g., Jacoby et al., 1997). Moreover, age should positively correlate with intrusions, at least in the older adults, given Hasher and Zacks' (1988) hypothesis of age declines in inhibitory functioning. Furthermore, it was speculated that individuals with high levels of education, verbal ability, and perceptual speed would show better memory performance than would their same-aged peers who are less educated, less verbal, or slower-performing. More specifically, for each age group, these three individual difference measures (years of education, Mill Hill performance, and Digit Symbol performance) should correlate positively with recollection and episodic

Table 2.7

Experiment 1: Selected Correlations for Younger and Older Adults

Younger Adults

Memory measures	***Demographic and Neuropsychological Measures***						
	Age	Education	Vocabulary	Digit Symbol	Stroop Interference	Verbal Fluency – Frontal	Verbal Fluency – Temporal
Recollection	.02	.10	.09	.08	.07	.16	**-.50***
Automatic	-.03	-.05	-.03	.06	.17	-.11	**-.46***
Episodic	.25	.32	.22	**.47***	.04	.19	-.33
Intrusions	-.13	-.21	-.12	.10	.02	-.16	.17

Older Adults

Memory measures	***Demographic and Neuropsychological Measures***						
	Age	Education	Vocabulary	Digit Symbol	Stroop Interference	Verbal Fluency – Frontal	Verbal Fluency – Temporal
Recollection	-.13	-.23	.17	.16	-.07	.08	-.22
Automatic	.14	.27	.30	-.24	.23	-.20	.10
Episodic	-.20	-.23	**.58****	.17	**-.37**$^{+}$	**.41**$^{+}$	-.08
Intrusions	.12	.29	-.12	-.34	.19	-.29	.13

Note. Recollection is the PDP-derived estimate of consciously controlled memory influences. "Automatic" is a composite measure combining two z-score-transformed estimates of automatic memory influence: habit and "guessing". "Episodic" refers to hits in the episodic baseline, where items appeared only once, at study. Intrusions refers to the number of thrice-presented items erroneously recalled in the once-presented critical conditions. Stroop interference is the difference in mean RT between the incongruent (interference) condition and color-naming baseline of the blocked Stroop test. Verbal fluency-Frontal is a composite measure of frontal-lobe functioning that combines two z-score-transformed measures (generation and switching) from both the phonemic and semantic blocks of the verbal fluency test. Verbal fluency-Temporal is a composite measure of medial-temporal functioning that combines two z-score-transformed clustering measures (phonemic and semantic) from the verbal fluency test. N = 24 for each age group; $^{+}$ = $p < .10$, * = $p < .05$, ** = $p < .01$, *** = $p < .001$.

influences but should correlate negatively with intrusions and show no correlation with automatic influences.

Some neuropsychological measures were also included in these correlational analyses. Stroop interference was operationalized as the absolute mean difference in mean RT between the incongruent (interference) condition and the color-naming baseline of the blocked Stroop test. The Stroop task has been used in previous research as a rough index of frontal lobe functioning (e.g., Golden, 1978). Verbal fluency test performance was subdivided into a frontal component and a medial-temporal component based on work by Troyer et al. (1997). Each of these two components was a composite measure, based on the arithmetic means of two or more z-score-transformed subscales. A verbal fluency frontal scale combined two subscales (generation and switching) from both the phonemic and semantic blocks of the verbal fluency test. A verbal fluency medial-temporal scale combined two clustering subscales (phonemic and semantic) from the test. Based on current theorizing (e.g., Moscovitch, 1992, 1994), measures of frontal lobe functioning were expected to correlate with strategic or consciously controlled uses of memory. Therefore, the verbal fluency frontal scale should correlate positively with recollection and negatively with intrusions, whereas Stroop interference should show the opposite relation to these two memory measures. In contrast, measures of medial-temporal or hippocampal functioning should correlate with automatic or associative uses of memory (Moscovitch, 1992, 1994). Therefore, the verbal fluency medial-temporal scale should correlate positively with automatic memory influences and with cued recall of words paired once with a particular sentence context.

For the younger adults, there were three reliable correlations: a positive correlation between the episodic base rate and perceptual speed, $r(22) = .47$, $p < .05$; a negative correlation between the recollection and the verbal fluency medial-temporal scale, $r(22) = -.50$, $p < .05$; and a negative correlation between automaticity and the verbal fluency medial-temporal scale, $r(22) = -.46$, $p < .05$. The first correlation, between episodic influences and speed, suggested that recall of once-presented completions increased with greater perceptual speed. The latter two correlations involving medial-temporal functioning were not predicted by theory: The suggestion of decreasing medial temporal functioning with increasing recollection or automaticity is clearly counterintuitive. However, there was one visible statistical outlier: One younger adult with high clustering scores showed poor recall in the episodic baseline. When this outlier was removed, the verbal fluency medial-temporal scale was no longer correlated with either recollection, $r(21) = -.22$, $p = .31$, or automatic influences, $r(21) = -.12$, $p = .59$.

For the older adults, there were two reliable correlations: a positive correlation between the episodic base rate and vocabulary, $r(22) = .58$, $p < .01$; and a positive correlation between episodic influences and the verbal fluency frontal scale, $r(22) = .41$, $p < .05$. There was also a marginally reliable negative correlation between episodic influences and Stroop interference, $r(22) = -.37$, $p < .10$. The first correlation, between episodic influences and vocabulary, suggested that recall of once-presented items increased with higher levels of verbal ability. The second correlation, between episodic influences and the frontal scale of the verbal fluency test, suggested that cued recall of once-presented items increased with higher levels of frontal-lobe functioning. This finding was not readily predicted from theory, as it was thought that the episodic base rate in this paradigm was more related to medial-temporal functioning than frontal functioning, due to the lack of experimental distractors and the presence of an associative retrieval cue. The third correlation, between episodic influences and Stroop interference, suggested that cued recall of once-presented items increased with decreased susceptibility to interference. However, the reliability of this latter finding was partly due to a statistical outlier, or one older adult with high Stroop interference and a low episodic base rate: When this outlier was removed, the correlation between these two variables was no longer reliable, $r(21) = -.27$, $p = .21$. In summary, there was some evidence that frontal lobe functioning and verbal ability were correlated with recall of once-presented items in older adults. It is possible that recall of once-presented items was of a strategic, rather than associative, nature for this age group.

Discussion

To summarize the main findings for Experiment 1, recall in the critical conditions was higher when the study completion was thrice-trained, rather than once-trained, and when participants were younger, rather than older, adults. However, voice context had no reliable effect on recall of repeated sentences. Moreover, no reliable interactions were obtained for these critical conditions. Nonetheless, there were non-reliable trends in the direction of lesser age differences with increasing environmental support: Namely, age differences in recall were numerically smaller when the same voice was employed for a sentence across experimental phases, as compared to different voices. Moreover, age differences were numerically smaller when sentences were thrice-trained, as compared to once-trained. Furthermore, the magnitude of the age difference was negligible in the Thrice-trained/Same-voice condition (M difference = .01), as compared to the other three critical conditions (M difference = .08). In other words, there were non-reliable trends in each of the predicted directions, as far as age differences in hits were concerned.

Further findings of interest in the critical conditions include a failure to find reliable age differences for either the experimental or extra-experimental false alarms. However, there were numeric trends toward an age-related increase for both types of false alarm. With regard to the PDP-derived measures (see Hay & Jacoby, 1996), the two age groups were statistically equivalent in both recollection and automatic influences of memory (i.e., both habit and guessing scores), with voice context having no effect on either type of memory influence. However, there was a slight numeric advantage for younger adults in recollection (i.e., M age difference = .08), as well as habit (i.e., M age difference = .04). Moreover, for both recollection and habit, there was a numeric trend toward a larger age difference in the different-voice condition than in the same-voice condition. In fact, the older adults actually showed higher habitual influences in the same-voice condition than did the younger adults. This finding of a small age-related increase in habit, together with the apparent age-equivalence in recall for the Thrice-trained/Same-voice condition, suggests that older and younger adults may achieve equivalently high performance as a consequence of age-related increases in automatic influences (see, e.g., Brown, 1994, and Rybash & Hoyer, 1996, for similar findings).

With regard to the baseline conditions, there was an age-related decrement in the test-only condition and an even larger decrement in the study-and-test-only condition. That is, older adults performed worse than younger adults when producing never-presented words (i.e., in the pre-experimental baseline) and words that were presented once only (i.e., in the episodic baseline). However, there was no age difference in the training-and-test-only condition, suggesting that stimulus repetition can reduce age differences in a direct memory test, as compared to one stimulus presentation. Voice congruency had little bearing on performance in these conditions, except for the surprising finding in the study-and-test-only condition of higher recall when different voices were used for a sentence at study and test, rather than the same voice. This latter finding was likely a distinctiveness or contrast effect, given the number of other sentences that appeared repeatedly in the same voice. That is, a change in voice for once-presented sentences may have been especially surprising or memorable for participants. However, this finding was not of particular theoretical interest (but see Brown, Neblett, Jones, & Mitchell, 1991, for other transfer inappropriate findings). Finally, there was little consistent evidence for an age-related increase in false alarms in the baseline conditions.

Additional analyses examined memory performance as a function of the total number of target presentations. These analyses, which bridged across the critical and baseline conditions, revealed that recall was higher when participants were younger, rather than older, adults and when targets were presented frequently, rather than infrequently, across experimental phases. Moreover, there was an Age x Presentation frequency interaction suggesting that the magnitude of the age difference decreased with increasing presentations of the target: More specifically, age differences were reliable at 2 or fewer presentations of a stimulus, but not at 3 or more presentations. However, these presentation frequency analyses provided little evidence for an age-related increase in false alarms, of either the experimental or extra-experimental variety.

The present results provided little evidence for an age-related decline in either deliberate recollection or inhibition. However, in both cases, there were non-reliable age decrements. That is, relative to the younger adults, the older adults showed a slight numerical decrease in the PDP-derived estimate of recollection and a slight numerical increase in false alarms. On the other hand, the rather sizeable age difference observed in the episodic baseline (i.e., a mean age difference of .17 in the study-and-test-only condition) suggests an age-related decline in explicit memory for once-presented items. Moreover, the finding of a greater age difference in recall for the interference (i.e., once-trained) conditions, as compared to the facilitation (i.e., thrice-trained) conditions, is compatible with the impaired-inhibition account of aging and memory. However, this finding is also somewhat consistent with the impaired-recollection account. This is the case because the once-trained condition has fewer repetitions of the target, and more repetitions of the alternative experimental completion, than does the thrice-trained condition. One might therefore expect the former condition to be less supportive of both recollective and inhibitory processes than the latter condition, and hence, more likely to produce age differences. Moreover, age decrements in recall were reliable in both the once-trained and thrice-trained conditions, which is consistent with an age decline in recollection.

The present findings are also in line with Craik and Jennings' (1992) suggestion that age differences in memory may be reduced when conditions are highly supportive at encoding and retrieval (e.g., identical semantic cues at study and test). In particular, the present experiment showed that repeated pairing of words and sentence frames benefited older adults more than younger adults, based on the finding that age differences "disappeared" when there were 3 or more presentations of the same stimulus. Moreover, secondary analyses of the critical conditions suggested that the two age groups showed equivalently high recall when both conceptual and perceptual stimulus information was repeatedly reinstated. However, in a departure from Craik and Jennings' (1992) notions of environmental support, there were no conditions in the present experiment where the two age groups performed equivalently poorly. Namely, the younger adults outperformed the older adults in the pre-experimental baseline, in which the target completion appeared at no point in the experiment. This result is somewhat surprising, given previous findings of age-invariance in semantic associations (e.g., Howard, 1980). However, this age difference likely reflects older adults' greater reluctance to produce a response when the target completion was infrequently presented.

Experiment 2

Experiment 2 was largely an extension of Experiment 1. However, instead of simply varying "experimental expertise" (i.e., the frequency with which associations appeared in a prior experimental phase), the present study varied the extent to which participants could rely on their pre-experimental expertise or schematic knowledge to recall associations. For example, participants seeing a sentence frame such as "He had to fill the truck with ____" would be more likely to complete the sentence with the word "gas" than with the word "cement." This is the case because, in real life, people are more likely to fill a truck with gas than with cement. Within the context of the present experiment, if the ability to learn new associations declines with age, then older adults should have more difficulty learning a low-probable sentence completion than a higher-probable sentence completion. In other words, older adults should have reasonably little difficulty learning a high-probable response such as "gas" in the above example because such responses come automatically to mind. In fact, older adults might show more activation of such schematic knowledge than younger adults, given recent findings that semantic knowledge (e.g., vocabulary) appears to remain constant or even increase over the lifecourse (see, e.g., Hess, 1990, and Light, 1992). In contrast, older adults should have more difficulty learning a lower-probable response such as "cement" because the word "gas" comes more readily to mind. According to researchers like Hasher and Zacks (1988), inappropriate high-probable responses should be especially difficult for older people to suppress because this age group has deficient inhibitory mechanisms (see also, e.g., Hartman & Hasher, 1991, for a relevant finding). In summary, age differences should be more pronounced when participants are learning a low-probable sentence completion, as compared to a high-probable completion.

Therefore, in the present experiment, participants saw and heard sentence frames, paired with either low-cloze or medium-cloze endings. (*Cloze* is a linguistic term referring to the pre-experimental baseline of someone supplying a particular word for a given sentence frame.) The mean pre-experimental baseline completion rates for low- and medium-cloze endings were .10 and .30, respectively. High-cloze endings were avoided in order to reduce the potential for ceiling effects.

However, the present manipulation did not end with varying the congruency of sentence completions with prior schematic knowledge. An additional variable was the frequency at which sentence completions appeared in a prior training phase. That is, for the critical items appearing in all experimental phases, half of the sentence completions were predominant at training (i.e., appearing on three of four training trials), whereas the remaining half were non-dominant at training (i.e., appearing on only one of the four training trials). In other words, sentence frames were either once-trained or thrice-trained, as in Experiment 1. Moreover, the frequency of completions' occurrence at training was orthogonal to, or crossed with, the completions' schema-congruency or cloze value. In summary, the critical conditions reflect a three-way factorial design, with each factor having two levels: Age (young versus old) x Frequency at training (once- vs. thrice-trained) x Cloze (low- vs. medium-cloze). In addition, the same baseline conditions were employed as were used in Experiment 1 (i.e., there were baselines for pre-experimental influences, pure episodic influences, and training influences).

As in the prior experiment, it was expected that age differences in recall would decrease with increased frequency of presentation at training. However, in addition to this interaction of age and frequency at training, a three-way interaction was expected. Namely, age differences were expected to be small when the study item was high in both experimental and pre-experimental typicality (i.e., the Thrice-trained/Medium-cloze condition) but were expected to be sizable when the study item was low on both dimensions (i.e., the Once-trained/Low-cloze condition). Age differences were expected to be moderate under conditions of intermediate typicality (i.e., the Thrice-trained/Low-cloze and Once-trained/High-cloze conditions).

Said another way, older adults were expected to benefit from two types of expertise, one experimenter-created and the other previously-acquired, when study items were both schema-congruent and frequently presented at training. If retrieval of such items is highly automatic, then age differences might be reduced. In contrast, older adults were predicted to be especially error-prone (more so than younger adults) when study items were both schema-incongruent and infrequently presented. This is the case because, under such circumstances, the individual must consciously overcome two influences that would automatically predispose the individual to retrieve the incorrect sentence completion.

Methods

Design. Figure 2.3 outlines the critical and baseline conditions employed in Experiment 2. As in the previous experiment, there were critical sentences that appeared in all three experimental phases (training, study, and test). In these critical conditions, sentences appeared a total of four times at training (three times with one completion and once with an alternative completion), and then once at study with either the once-trained or thrice-trained completion, before re-appearing at test without the completion. The completion presented at study therefore served as the "target" or the to-be-remembered completion for the memory test. In addition, there was a manipulation of the predictability or cloze value of the sentence completions: For each sentence frame, one of the two experimental completions was low in predictability (low cloze), whereas the other completion was moderately predictable (medium cloze). Moreover, the cloze value of the completions was manipulated independently of their presentation frequency. Therefore, for half of the sentences, the low-cloze completion was presented thrice and the medium-cloze completion was presented once, whereas for the remaining half of the sentences, the opposite was true. Furthermore, the low-cloze completion appeared as the "target" or study item for half of the sentences, and the medium-cloze completion appeared as the target for the remaining half of the sentences. In summary, there were four critical conditions that represented the crossing of two within-subjects factors, with two levels each: presentation frequency of target completions (the use at study of either the once-trained or thrice-trained completion), and cloze value (the use at study of either the low-cloze or medium-cloze completion for a sentence). There was also one between-subjects factor, age, with two levels (younger vs. older adults). The

Figure 2.3. Experiment 2: Overview of design and procedure

Training phase

Critical conditions:
At training, critical sentences were presented 4 times each: 3 times with one completion and 1 time with an alternative completion. The two completions differed in plausibility, or cloze value (low- vs. medium-cloze).

For half of the sentences, the low-cloze completion was presented 1 time and the medium-cloze completion 3 times; for the remaining sentences, the opposite was true.

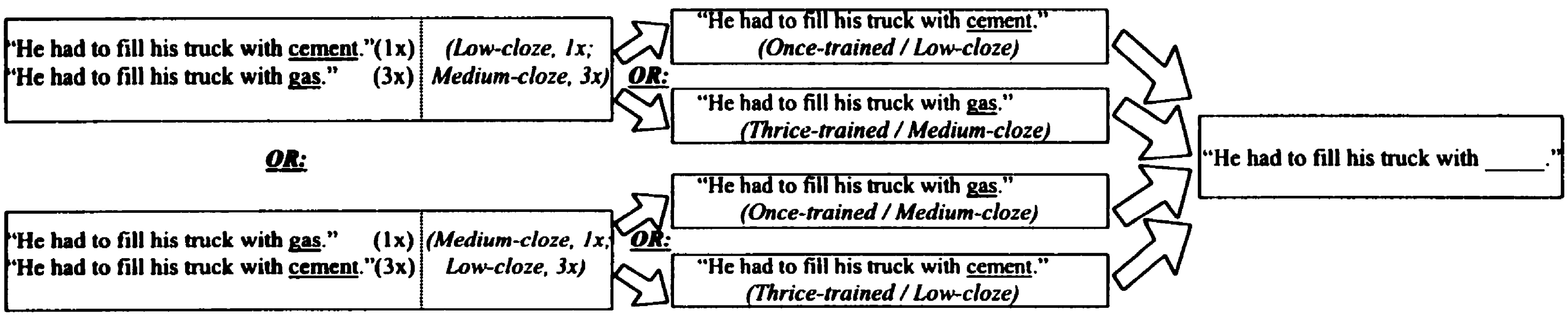
Study phase

At study, trained sentences were re-presented once only, with either the once-trained completion or the thrice-trained completion.

In addition, the studied completion was the low-cloze completion for half of the sentences and the medium-cloze completion for the remaining sentences.

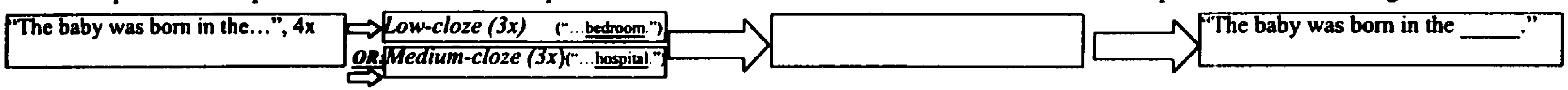
Test phase

At test, sentences were re-presented without any completions. These sentence frames served as retrieval cues for the studied completions.

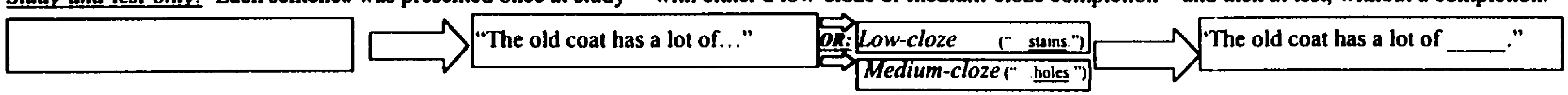


Baseline conditions:

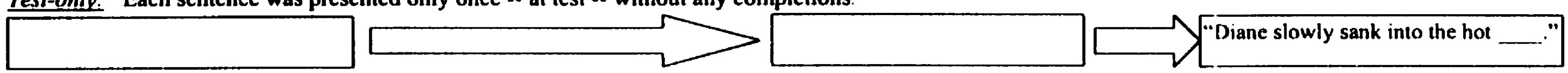
Training-and-test-only: Each sentence was presented 4 times at training (3 times with one completion, 1 time with another completion), and once again at test. The thrice-presented completion was the low-cloze completion for half of the sentences and was the medium-cloze completion for the remaining sentences.



Study-and-test-only: Each sentence was presented once at study -- with either a low-cloze or medium-cloze completion – and then at test, without a completion.



Test-only: Each sentence was presented only once -- at test -- without any completions.



analyses of the critical conditions therefore correspond to a 2x2x2 (Age x Frequency of presentation at training x Cloze value) design.

As in Experiment 1, there were three types of baseline condition: a training or "guessing" baseline (cf. Hay & Jacoby, 1996) in which items appeared only at training and test but not at study; an episodic baseline in which items appeared only at study and test but not at training; and a pre-experimental baseline in which sentences appeared for the first time at test. Each participant received all three types of baseline condition. The conceptualization of the target or "correct" response varied across the different baseline conditions: For the training-and-test-only items, the term "hits" referred to the proportion of thrice-presented items produced from the training phase; for the study-and-test-only baseline, it was the proportion of study items produced; and for the test-only condition, it was the proportion of "target" items produced (the target being one of the two experimenter-generated completions for each sentence that was arbitrarily designated as the correct answer in each experimental format). The first two baselines could be readily subdivided on the basis of cloze value. More specifically, for the training baseline, half of the thrice-presented (target) completions at training were low-cloze and the remaining half of the typical completions were medium-cloze. Moreover, for the episodic baseline, half of the study completions were low-cloze and the remaining half were medium-cloze. Although the third, pre-experimental baseline could also be subdivided into low-cloze and medium-cloze responses, no completion of either type was presented for these items at any point during the experiment. Moreover, for these test-only items, it simplified the analyses to arbitrarily designate the medium-cloze response as a "hit" and the low-cloze response as an "experimental FA." This arbitrary classification system facilitated comparisons between Experiments 1 and 2 by employing similar analyses in each case. In summary, there were nine experimental conditions of interest: four critical or main conditions, in which items appeared in all three experimental phases (training, study, and test); and five baseline conditions, in which items appeared in two or fewer experimental phases.

Finally, as was also true of Experiment 1, the total number of target presentations varied across the baseline and critical conditions. Of particular interest was the fact the total number of target presentations across the three experimental phases ranged from 0 in the test-only condition to 4 in the thrice-trained conditions; moreover, all target frequencies between these two extremes were represented across the experimental conditions. Therefore, the experimental design lent itself to further presentation-frequency analyses that bridged across the critical and baseline conditions. These analyses corresponded to a 2x5 (Age x Total number of target presentations) design, with age varying between subjects and presentation condition within subjects.

Participants. Demographic information on the 24 younger and 24 older adults is presented in Table 2.8. None of the participants had taken part in Experiment 1. Older adults were healthy volunteers who were offered reimbursement for their travel expenses, whereas younger adults were psychology undergraduates who received either course credit or $15 for their participation in a 2-hour session. On a 20-item version of the Mill Hill Vocabulary Scale (Raven, 1965), the older adults scored higher than the younger adults, $t(46) = 3.80$, $p < .001$. In contrast, on a perceptual speed measure – the WAIS Digit Symbol Substitution Task (Wechsler, 1981) – the younger adults outperformed the older adults, $t(46) = 7.84$, $p < .001$. The finding of an age-related increase in vocabulary, despite a decrease in perceptual speed, was reported in Experiment 1 and in previous published studies of aging and memory (e.g., Anderson et al., 1998). The two age groups were equivalent in years of formal education and in self-reported health, $ts < 1$, as assessed for two time frames: the last two to three months, and today. Both health measures were on a 5-point scale, ranging from 1 ("excellent") to 5 ("poor").

Table 2.8
Experiment 2: Mean Participant Characteristics

Variable	**Age Group**				**t**
	Young		Old		
	M (SD)	n	M (SD)	n	
Age	23.0 (3.2)	24	72.5 (4.9)	24	41.48***
Age range	19-32		60-80		-------
Gender:					-------
Female		22		21	
Male		2		3	
Handedness:					-------
Right-handed		23		20	
Left-handed		1		1	
Ambidextrous		0		3	
Education (years)	16.0 (1.5)		15.2 (2.7)		1.28
Self-reported health:					
Last 2 to 3 months	1.8 (0.8)		1.9 (0.8)		0.54
Today	1.8 (0.9)		2.0 (0.7)		0.73
Mill Hill Vocabulary Scale	13.6 (2.6)		16.4 (2.5)		3.80***
Digit Symbol Substitution Task	70.1 (10.2)		48.6 (8.7)		7.84***

Note. For gender, $\chi^2_{(1, N=48)} = .22$, $p > .50$. For handedness, $\chi^2_{(2, N=48)} = 3.21$, $p > .20$. Self-reported health was assessed for two time frames and was rated on a 5-point scale, ranging from 1 (excellent) to 5 (poor). Mill Hill Vocabulary Scale maximum = 20. Digit symbol score is the number of correct completions in 90 sec (maximum = 93). t values are for younger versus older adults (df = 46); * = $p < .05$, ** = $p < .01$; *** = $p < .001$.

Materials. A total of 147 sentence frames were generated to serve as experimental stimuli; each sentence frame was one word short of being a complete sentence. Three of the sentence frames were employed as practice items and were identical to the practice items used in Experiment 1. The remaining 144 sentence frames served as critical items and were selected so that, among the possible completions for that sentence frame, one completion had a low probability of being produced without prior experimental exposure (low cloze) and a second completion had a moderate probability of being produced (medium cloze). Low-cloze completions had baseline completion rates of approximately .10, whereas medium-cloze completions had base rates of approximately .30. For example, in the sentence frame, "He had to fill his truck with _____," the completion "cement" was considered low-cloze and the completion "gas" was medium-cloze. Of the 144 critical sentence frames, 85 were taken from Bloom and Fischler's (1980) sentence completion norms and the remaining 59 were taken from unpublished sentence completion norms collected by the author. For this latter set of norms, 60 graduate students at the University of Toronto (M age = 25.7 years, SD = 3.5; M education = 18.8 years, SD = 2.3) responded to 128 sentence frames by writing the first word that came to mind. Thirty-five of the respondents were psychology graduate students; the remaining 25 respondents were graduate students in other disciplines who resided at a graduate residence of the university. Each participant was given sheets of paper with 64 sentence frames, presented in one of two fixed random orderings, and was asked to respond to the survey individually and without trying to be "creative" in completing the sentences. Although no older adults participated in either normative study, this was not considered problematic for the present research, given previous findings of similar semantic associations for older and younger adults (e.g., Howard, 1980). Based on the two sets of norms, sentences were selected so that the mean baseline completion rate for low-cloze completions was exactly .10 (SD = .04; range = .02 to

.19), whereas the mean baseline completion rate for medium-cloze completions was exactly .30 (SD = .06; range = .21 to .44). The sentence frames and their two alternative experimental completions are listed in Appendix 2.3.

The 144 critical sentence frames and their alternative completions were divided into nine sets of 16 so that each participant was exposed to all nine conditions: The four critical conditions in which items appeared at training, study, and test; and the five baseline conditions in which items appeared in two or fewer of the experimental phases. However, the large number of sentence frames employed (a total of 624 presented to each participant) made it impractical to produce the nine experimental formats required for exhaustive rotation of stimuli through all conditions. Therefore, three completely randomized experimental formats were generated so that in each format, all stimuli were equally likely to appear in any given condition. For each sentence in a given format, one of the two alternative completions was randomly selected to be the target completion and the other alternative to be the non-target completion.

Sentence frames for all three experimental phases appeared on laminated sheets of paper kept in a three-ring binder. As in the previous experiment, the sentence frames were quintuple-spaced and appeared in bold print in an 18-point, Times New Roman font. A blank five characters long (i.e., "_____") was used to indicate that a word was missing from each sentence frame and that this word could change across stimulus presentations. This missing word was revealed in an auditory version of the stimulus that was presented at the same time as the visual version. Unlike the previous study, visual formats were produced for the test sentence frames because it was no longer necessary to emphasize the particular voice contexts employed. A paper mask (i.e., two pieces of white cardboard glued together) was fashioned with a rectangular hole (approximately 2 x 18 cm) in order to display sentences one at a time. For all three experimental phases, audio recordings were prepared in which stimuli were spoken in a male voice. For the training and study phases, entire sentences were recorded, including the missing words from the printed version; for the test phase, sentence frames were recorded, with the word "blank" indicating the missing word in each case. Sentences and sentence frames were preceded by two beeps in rapid succession that signaled that the experimenter should advance the paper mask from one stimulus to the next. These double beeps appeared at 10-sec intervals. Moreover, two seconds before the double-beep (or 8 sec after the presentation of a stimulus), a further, single beep was sounded. This additional, warning beep alerted the participant that their time limit was almost up and that they should make sure to come up with an answer before the next beep, even if they had to guess. The warning signal was provided in an effort to reduce the number of non-responses in the present study, as compared to the previous study.

Procedure. In most important respects, the procedure was the same as that employed in Experiment 1. The ordering and types of experimental task were virtually identical to those employed in the previous experiment, as were the list lengths and presentation rates for each experimental phase. There were only two noticeable changes in procedure: a change in instructions, and the addition of a further interpolated task. First, there was a small change in instructions prior to the test phase. Participants were now informed that they would have 10 seconds to respond to each sentence frame, but that there would be a single warning beep after 8 seconds. This beep warned them that their time limit was almost up and that they should immediately provide their best guess as to how the sentence would be completed. Participants were advised that they should always come up with a response, no matter what. Participants were further advised that, after 10 seconds, they would hear a double beep which advised them that the trial was over. The experimenter's instructions to participants are shown in Appendix 2.4. A second change from the first experiment was the Trail Making Task was added as a fifth interpolated task in between the study and test phases. This task provided an additional measure of frontal lobe functioning and was a further attempt to avert ceiling effects in the memory test.

Trail making task. The Trail Making test (Reitan, 1955, 1958) is a measure of mental agility that has been correlated with frontal lobe functioning. This is a timed pencil-and-paper test consisting of two parts: Trails A and Trails B. In both parts, participants were instructed to connect dots on a page in a pre-determined order, as quickly and as accurately as possible. In Trails A, a total of 25 dots appeared on a page, which were numbered from "1" to "25" in a fixed random ordering. Participants were instructed to connect the dots in ascending numerical order. In Trails B, the same number of dots were shown, but the numbers "1" through "13" and the letters "A" through "L" appeared next to the dots. Participants were instructed to begin at "1" and alternate between numbers and letters in ascending numerical and alphabetical order as they proceeded (e.g., the participants should connect the dots labeled "1," "A," "2," "B," "3," "C," and so on). Each test page was preceded by a practice page with eight dots. The experimenter recorded each participant's speed in completing each page and the number of errors committed. The two parts of the test took no more than 5 min. to complete.

Results

Cued recall in the critical conditions. Table 2.9 indicates participants' mean recall performance in the critical conditions, as a function of age group and memory measure. In these critical conditions, sentences appeared in all three phases of the experiment: training, study, and test. Four three-way (Age x Frequency of presentation at training x Cloze value) ANOVAs were conducted for the critical conditions, with each factor having two levels. Each ANOVA was conducted with age (younger vs. older adults) as a between-subjects variable and with frequency of presentation at training (once- vs. thrice-trained completions at study) and cloze value (low- versus medium-cloze sentence completions at study) as within-subject variables. As in

Table 2.9

Experiment 2: Cued Recall Performance for the Critical Conditions

Measure and age group	**Condition *(Presentation frequency/Cloze value)***							
	Once trained				**Thrice trained**			
	Low cloze		Medium cloze		Low cloze		Medium cloze	
	M	(SD)	M	(SD)	M	(SD)	M	(SD)
Hits								
Young	.77	(.15)	.80	(.13)	.89	(.14)	.91	(.12)
Old	.63	(.16)	.71	(.15)	.83	(.21)	.83	(.13)
FAs, experimental								
Young	.19	(.10)	.17	(.12)	.08	(.09)	.06	(.06)
Old	.29	(.10)	.22	(.11)	.10	(.13)	.09	(.07)
FAs, extra-experimental								
Young	.04	(.08)	.03	(.05)	.03	(.07)	.03	(.10)
Old	.08	(.12)	.07	(.11)	.06	(.08)	.09	(.11)
Non-responses								
Young	.01	(.02)	.00	(.00)	.003	(.01)	.00	(.00)
Old	.003	(.01)	.00	(.00)	.01	(.03)	.00	(.00)

Note. Critical conditions are those in which items appeared in all three experimental phases (training, study, and test). Cloze value refers to the probability of producing a particular completion for a sentence frame without any prior experimental exposure. False alarms (FAs) were one of two types: experimental, or the alternative completion provided for a sentence at training; and extra-experimental, or any other response that did not correspond to an experimental completion. Non-responses refer to trials in which participants failed to make a response within the 10-sec deadline.

Experiment 1, an ANOVA was conducted for each of the following dependent measures: hits, experimental FAs, extra-experimental FAs, and non-responses.

The ANOVA for hits revealed main effects of age, $F(1,46) = 6.91$, $p < .05$, $MSE = 0.065$, presentation frequency at training, $F(1,46) = 91.11$, $p < .001$, $MSE = 0.010$, and cloze value, $F(1,46) = 6.24$, $p < .05$, $MSE = 0.009$. These main effects indicated that the younger adults showed higher recall than did the older adults. Moreover, recall was higher when the studied sentence completion was thrice-trained, as compared to once-trained, and when the target completion was medium-cloze, as compared to low-cloze. There were also two marginally reliable interactions: Presentation frequency x Cloze, $F(1,46) = 3.47$, $p < .07$, $MSE = 0.006$, and Age x Presentation frequency x Cloze, $F(1,46) = 3.04$, $p < .09$, $MSE = 0.006$. Tests of simple effects revealed that the two-way interaction occurred because the superiority of medium-cloze over low-cloze completions was reliable when the completions were once-trained, $t(47) = 2.73$, $p < .01$, but not when they were thrice-trained, $t(47) < 1$. On the other hand, the superiority of thrice-trained over once-trained completions was reliable for both low-cloze completions, $t(47) = 8.27$, $p < .001$, and medium-cloze completions, $t(47) = 6.60$, $p < .001$.

Further tests of simple effects were conducted to assess the nature of the three-way interaction for hits. First, two separate two-way ANOVAs assessed the reliability of the Presentation frequency x Cloze interaction for each age group: These analyses revealed that the two-way interaction was reliable for the older adults, $F(1,23) = 4.83$, $p < .05$, $MSE = 0.008$, but not for the younger adults, $F(1,23) < 1$, $p > .91$, $MSE = 0.004$. Tests of these simple main effects revealed that, for the older adults, the superiority of medium-cloze over low-cloze completions was reliable when the completions were once-trained, $t(23) = 2.61$, $p < .05$, but not when they were thrice-trained, $t(23) = 0.00$, $p = 1.00$. In contrast, for the younger adults, the superiority of medium-cloze over low-cloze completions was not reliable when completions were once-trained,

$F(1,23) = 1.19$, $p > .24$, and was only marginally reliable when completions were thrice-trained, $F(1,23) = 1.80$, $p < .09$. On the other hand, both age groups showed a superiority of thrice-trained over once-trained completions at each cloze value (all ts > 4.45, ps < .001). A second set of analyses assessed the nature of the three-way interaction by examining the reliability of age differences for each of the four critical conditions: Pairwise comparisons revealed that the superiority of younger over older adults was not statistically reliable in one critical condition -- Thrice-trained/Low-cloze, $t(46) = 1.17$, $p = .25$ -- but was reliable in the three other critical conditions: Once-trained/Low-cloze, $t(46) = 3.17$, $p < .01$, Once-trained/Medium-cloze, $t(46) = 2.42$, $p < .05$, and Thrice-trained/Medium-cloze, $t(46) = 2.53$, $p < .05$.

With further regard to hits, although the Age x Frequency interaction was non-reliable, $F(1,46) = 2.56$, $p > .11$, $MSE = 0.010$, pairwise comparisons examined the reliability of the age difference in memory for once-trained completions, as compared to thrice-trained completions. As discussed previously, the once-trained target is a case of interference because it is opposed by an alternative completion on three training trials, whereas the thrice-trained target is a case of facilitation because it is opposed by a competing completion on only one trial. Moreover, a central hypothesis of the dissertation is that age differences in memory should be larger for interference conditions than for facilitation conditions. Consistent with this view, the older adults showed worse recall than the younger adults in the once-trained condition, $t(46) = 2.98$, $p < .01$, but were equivalent to the younger adults in the thrice-trained condition, $t(46) = 1.16$, $p > .25$.

The ANOVA for experimental FAs revealed main effects of age, $F(1,46) = 7.29$, $p = .01$, $MSE = 0.016$, presentation frequency at training, $F(1,46) = 96.57$, $p < .001$, $MSE = 0.009$, and cloze value, $F(1,46) = 4.80$, $p < .05$, $MSE = 0.008$. These main effects revealed that the older adults produced more experimental FAs than did the younger adults. Moreover, the likelihood of producing experimental FAs was higher when the target completion was once-trained, as compared to thrice-trained, and when the target completion was low-cloze, as compared to medium-cloze. In other words, when expressed in terms of alternative completions rather than target completions, the likelihood of experimental FAs was greater when the alternative experimental completion was thrice-trained, rather than once-trained, and when the alternative completion was medium-cloze, rather than low-cloze. There was also a marginally reliable Age x Presentation frequency interaction, $F(1,46) = 3.57$, $p < .07$, $MSE = 0.009$. Tests of simple effects revealed that this interaction occurred because the age-related increase in experimental FAs was reliable for once-trained completions, $t(46) = 2.98$, $p < .01$, but not for thrice-trained completions, $t(46) = 1.16$, $p > .24$. On the other hand, the increase in intrusions of alternative completions from the once-trained condition to the thrice-trained condition was reliable for the younger adults, $t(46) = 6.58$, $p < .001$, as well as for the older adults, $t(46) = 7.34$, $p < .001$.

The ANOVA for extra-experimental FAs revealed a marginally reliable effect of age, $F(1,46) = 3.72$, $p = .06$, $MSE = 0.027$, and a marginally reliable Presentation frequency x Cloze interaction, $F(1,46) = 3.20$, $p = .08$, $MSE = 0.002$. The main effect of age suggested that there were more intrusions of extra-experimental completions when participants were older, as compared to younger, adults. The two-way interaction occurred because in the low-cloze condition, the once-trained completions yielded marginally more extra-experimental intrusions than did the thrice-trained completions, $t(47) = 1.63$, $p < .11$, whereas in the medium-cloze condition, the once-trained and thrice trained completions yielded equivalent numbers of extra-experimental intrusions, $t(47) < 1$.

The analyses for non-responses were conducted for only the low-cloze conditions because there was no variance (i.e., means of exactly zero) in the two medium-cloze conditions: In other words, both age groups responded on all of the trials when medium-cloze completions appeared at study. Therefore, a two-way ANOVA (Age x Presentation frequency) was conducted for the low-cloze completions. This analysis revealed only a marginally reliable two-way interaction, $F(1,46) = 3.81$, $p < .06$, MSE = 0.0003. The interaction occurred because the older adults failed to respond to more thrice-trained completions than once-trained completions, $t(23) = 1.70$, $p < .11$, whereas the younger adults were equally likely to not respond to the two types of completion, $t(23) = 1.00$, $p > .32$. In any case, the very low numbers of non-responses reflect a floor effect in the critical conditions.

Cued recall in the baseline conditions. Table 2.10 indicates participants' mean recall performance in the baseline conditions, as a function of age group and memory measure. Separate analyses were conducted for each of the three main baseline conditions: test-only, study-and-test-only, and training-and-test-only. For each baseline condition, four ANOVAs were conducted, or one for each of the four dependent measures described above: hits, experimental FAs, extra-experimental FAs, and non-responses. Age (younger vs. older adults) was a factor for all three baseline conditions. The cloze value (low- vs. medium-cloze) of the presented completion was a factor for the two baseline conditions in which sentences appeared in two experimental phases: study-and-test-only, and training-and-test-only. For the remaining baseline – the test-only condition -- sentences appeared only once, without any completions. Therefore, in line with the Experiment 1 analyses of test-only sentences, one of the two experimenter-generated completions was arbitrarily designated the target completion and the other experimenter-generated completion was designated the alternative completion. More specifically, the medium-cloze completion always served as the target completion for the test-only sentences, whereas the low-cloze completion always served as the alternative completion.

For the test-only condition, a one-way ANOVA (Age) was conducted for each of the four dependent measures. In this "pre-experimental" baseline, neither of the two experimenter-

Table 2.10
Experiment 2: Cued Recall Performance for the Baseline Conditions

Measure and age group	**Condition *(Presentation frequency/Cloze value)***									
	Test only		**Study and test only**				**Training and test only**			
			Low cloze		Medium cloze		Low cloze		Medium cloze	
	M	(SD)	M	(SD)	M	(SD)	M	(SD)	M	(SD)
Hits										
Young	.31	(.14)	.63	(.23)	.72	(.13)	.69	(.16)	.72	(.17)
Old	.30	(.12)	.52	(.15)	.65	(.15)	.62	(.18)	.63	(.15)
FAs, experimental										
Young	.08	(.06)	.11	(.10)	.03	(.04)	.26	(.12)	.22	(.14)
Old	.07	(.06)	.12	(.08)	.05	(.06)	.29	(.12)	.24	(.10)
FAs, extra-experimental										
Young	.60	(.15)	.26	(.16)	.24	(.14)	.05	(.10)	.06	(.10)
Old	.60	(.12)	.35	(.13)	.28	(.13)	.09	(.11)	.13	(.13)
Non-responses										
Young	.01	(.02)	.01	(.03)	.01	(.02)	.00	(.00)	.00	(.00)
Old	.03	(.07)	.01	(.03)	.02	(.05)	.01	(.04)	.003	(.01)

Note. Baseline conditions are those in which a sentence appeared in fewer than three experimental phases. Cloze value refers to the probability of producing a particular completion for a sentence frame without any prior experimental exposure. False alarms (FAs) were one of two types: experimental, or the alternative experimental completion used for a sentence (although not necessarily for that participant); and extra-experimental, or any other response that did not correspond to an experimental completion. For the study-and-test-only condition, the alternative experimental completion did not appear during the experiment. For the test-only condition, neither of the two experimenter-generated completions appeared during the experiment: However, for the purposes of description, the medium-cloze completion has been arbitrarily designated the target completion and the low-cloze completion is the alternative completion.

generated completions appeared during the experiment: Therefore, a response was arbitrarily designated a hit if it corresponded to the medium-cloze experimenter-generated completion and an experimental FA if it corresponded to the alternative, low-cloze completion. In these analyses, there were no age differences in hits, experimental FAs, or extra-experimental FAs (all $\underline{F}$s < 1). However, the older adults failed to respond to a larger number of the completely novel sentences than did the younger adults, $\underline{F}(1,46) = 3.65$, $\underline{p} < .07$, $\underline{MSE} = 0.06$. Nonetheless, the proportion of non-responses was negligible for both age groups.

For the study-and-test-only condition, a two-way ANOVA (Age x Cloze value) was conducted for each of the four memory measures. The ANOVA for hits revealed main effects of age, $\underline{F}(1,46) = 4.08$, $\underline{p} < .05$, $\underline{MSE} = 0.047$, and cloze value, $\underline{F}(1,46) = 31.91$, $\underline{p} < .001$, $\underline{MSE} = 0.010$. These effects indicated that recall of once-presented (studied but untrained) items was greater for younger than older adults, and for medium-cloze than low-cloze completions. In the ANOVA for experimental FAs, there was only a main effect of cloze value, $\underline{F}(1,46) = 26.21$, $\underline{p} < .001$, $\underline{MSE} = 0.005$, indicating that the likelihood of intruding the alternative experimental completion was greater for low-cloze completions than for medium-cloze completions. The ANOVA for extra-experimental FAs revealed a main effect of cloze value, $\underline{F}(1,46) = 6.90$, $\underline{p} < .05$, $\underline{MSE} = 0.006$, and a marginally reliable effect of age, $\underline{F}(1,46) = 3.27$, $\underline{p} < .08$, $\underline{MSE} = 0.034$. These effects indicated that the likelihood of intruding extra-experimental responses was greater for older than younger adults and for low-cloze than medium-cloze completions. Finally, the ANOVA for non-responses failed to reveal any reliable effects, either for age, $\underline{F}(1,46) = 1.57$, $\underline{p} > .21$, $\underline{MSE} = 0.001$, or for cloze value or the interaction of these two variables, $\underline{F}$s < 1.

For the training-and-test-only condition, a two-way ANOVA (Age x Cloze value) was once again conducted for each of the four memory measures. The ANOVA for hits revealed only a marginally reliable main effect of age, $\underline{F}(1,46) = 3.74$, $\underline{p} < .06$, $\underline{MSE} = 0.042$, suggesting that the younger adults had superior recall of trained, but unstudied items, than did the older adults. The ANOVA for experimental FAs revealed only a main effect of cloze value, $F(1,46) = 4.26$, $p < .05$, $MSE = 0.012$, indicating that there were more intrusions of the alternative experimental completion when the low-cloze completion was repeatedly presented at training than when the medium-cloze completion was repeated at training. The ANOVA for extra-experimental FAs revealed a main effect of cloze value, $F(1,46) = 4.19$, $p < .05$, $MSE = 0.002$, as well as a marginally reliable main effect of age, $F(1,46) = 2.90$, $p < .10$, $MSE = 0.022$, and a marginally reliable two-way interaction, $F(1,46) = 3.21$, $p = .08$, $MSE = 0.002$. The main effects indicated that the likelihood of intruding non-experimental responses was higher when participants were older, rather than younger, adults and when the medium-cloze completion was repeatedly presented at training, rather than the low-cloze completion. Tests of simple effects were conducted to assess the nature of the Age x Cloze interaction: Pairwise comparisons revealed that the age-related increase in non-experimental completions was reliable when medium-cloze completions were repeatedly trained, $t(46) = 2.11$, $p < .05$, but not when low-cloze completions were repeatedly trained, $t(46) = 1.09$, $p > .28$. Moreover, the increase in extra-experimental FAs for the medium-cloze condition, as compared to the low-cloze condition, was reliable for the older adults, $t(23) = 2.33$, $p < .05$, but not for the younger adults, $t(23) < 1$. Finally, with regard to non-responses in the training baseline, the younger adults showed no variance, or a mean of exactly zero; therefore, a one-way ANOVA (Cloze value) was conducted for only the older adults: This analysis revealed no effect of cloze value, $F < 1$.

Cued recall as a function of total presentation frequency. As in Experiment 1, the cued recall data were further analyzed as a function of the total number of target presentations. In other words, across the critical and baseline conditions, the target completion was presented a varying number of times during the experiment, ranging from 0 presentations across the experimental phases (i.e., the test-only condition) to 4 presentations (i.e., the thrice-trained condition; see Table 2.1). The presentation frequency analyses therefore bridged across the critical and baseline conditions. It was expected that age differences in recall should decrease with increasing stimulus presentations, in line with Craik and Jennings' (1992) suggestion that highly supportive conditions should minimize age differences in memory. Moreover, the older adults were expected to produce more experimental FAs than the younger adults, particularly in the 2 target-presentations condition (i.e., the once-trained condition), where there were three competing presentations of the alternative experimental completion. Figure 2.4 indicates the mean probabilities of producing the target completion and the alternative experimental completion, as a function of the total number of target presentations. The mean probabilities of producing extra-experimental responses and non-responses are not depicted, as these dependent measures were of lesser theoretical interest.

The analyses for presentation frequency consisted of four two-way ANOVAs, one for each of the memory measures described above: hits, experimental FAs, extra-experimental FAs, and non-responses. In each ANOVA, age served as a between-subjects factor with two levels (younger vs. older adults) and frequency of target presentation as a within-subjects factor with five levels (0 vs. 1 vs. 2 vs. 3 vs. 4 target presentations). Note that for the sentences in the 0-presentations (test-only) condition, neither experimental completion was presented at any point during the experiment. Therefore, for these completely novel sentences, one of the two experimenter-generated completions (i.e., the medium-cloze completion) was arbitrarily designated the target completion, whereas the other experimenter-generated completion (i.e., the low-cloze completion) was designated the alternative completion (refer to the baseline analyses, above). For the remaining four presentation frequencies, performance was collapsed across cloze conditions, in order to simplify the analyses. Of principal interest in these 2x5 (Age x

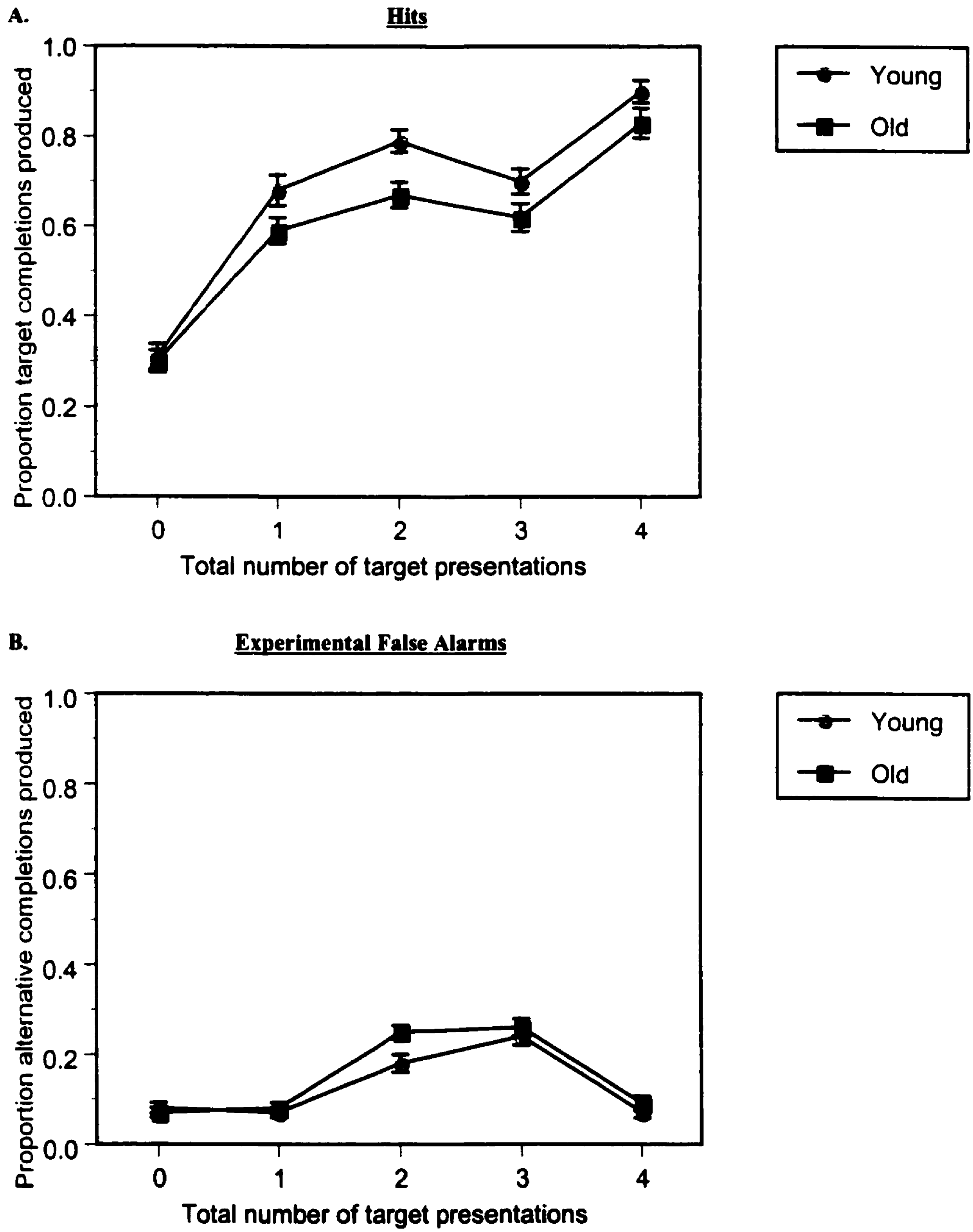


Figure 2.4. Experiment 2: Mean recall performance as a function of memory measure, age group, and total number of target presentations. Error bars of one standard error are plotted for each point but are only visible when greater in magnitude than the size of the symbol.

Frequency) ANOVAs was whether the magnitude of age differences varied with the total number of target presentations.

The ANOVA for hits revealed main effects of age, $F(1,46) = 6.94$, $p < .05$, $MSE = 0.050$, and presentation frequency, $F(4,184) = 169.28$, $p < .001$, $MSE = 0.012$, but no Age x Presentation frequency interaction, $F(4,184) = 1.44$, $p > .22$, $MSE = 0.012$. The main effects of age indicated that recall was higher for younger than older adults. Post-hoc Tukey HSD tests were conducted to assess the nature of the main effect of presentation frequency: These pairwise comparisons revealed that the number of hits increased from 0 to 2 target presentations, then decreased from 2 to 3 presentations, and then increased again from 3 to 4 presentations. Moreover, there were equivalent numbers of hits in the 1- and 3-presentations conditions. Although the two-way interaction was not reliable, it was nonetheless of some theoretical interest to examine the reliability of age differences at each frequency level. T-tests revealed that the age-related decline in hits was not reliable at 0 target presentations, $t(46) < 1$, but was reliable at 1 presentation, $t(46) = 2.02$, $p < .05$, and at 2 presentations, $t(46) = 3.18$, $p < .01$, and was only marginally reliable for the two remaining frequency conditions: 3 presentations, $t(46) = 1.93$, $p < .06$, and 4 presentations, $t(46) = 1.79$, $p = .08$. Secondly, two separate one-way ANOVAs assessed the effect of frequency for each group: A main effect was obtained for the younger adults, $F(4,92) = 102.24$, $p < .001$, $MSE = 0.011$, as well as for the older adults, $F(4,92) = 70.09$, $p < .001$, $MSE = 0.013$. Tukey tests were conducted to assess the nature of these simple main effects. For the younger adults, the number of hits increased from 0 to target 2 presentations, then remained constant from 2 to 3 presentations, and then increased again from 3 to 4 presentations. However, for this age group, the number of hits at 3 presentations did not vary reliably from that at 1 presentation. For the older adults, the number of hits increased from 0 to 1 target presentations, then remained constant from 1 to 3 presentations, and increased again from

3 to 4 presentations. In summary, the two age groups performed equally poorly at 0 presentations, then showed an age difference favoring the younger adults at 1 and 2 presentations, and finally showed a marginally reliable age difference at 3 and 4 presentations. In other words, both age groups appeared to benefit from stimulus repetition, but the older adults benefited less than their younger counterparts.

The ANOVA for experimental FAs revealed main effects of age, $F(1,46) = 4.58$, $p < .05$, $MSE = 0.007$, and presentation frequency, $F(4,184) = 70.54$, $p < .001$, $MSE = 0.005$, as well as an Age x Presentation frequency interaction, $F(4,184) = 2.53$, $p < .05$, $MSE = 0.005$. The main effect of age indicated that the older adults produced more experimental FAs than did the younger adults. Post-hoc Tukey HSD tests assessed the nature of the main effect of presentation frequency: These analyses revealed that there were more experimental FAs when the target was presented 2 or 3 times than when the target was presented 0, 1, or 4 times. Moreover, the numbers of experimental FAs did not vary from 2 to 3 target presentations, nor did they vary across the 0-, 1-, and 4-presentations conditions. Tests of simple effects were conducted to assess the nature of the Age x Presentation frequency interaction. First, pairwise comparisons revealed that the age-related increase in experimental FAs was reliable in the 2-presentations condition, $t(46) = 2.98$, $p < .01$, but not in the other four frequency conditions (all $ts < 1.17$, $ps > .25$). Secondly, two separate one-way ANOVAs assessed the effect of frequency for each age group: A main effect was obtained for the younger adults, $F(4,92) = 30.08$, $p < .001$, $MSE = 0.005$, as well as the older adults, $F(4,92) = 42.30$, $p < .001$, $MSE = 0.005$. Tukey tests were conducted to assess the nature of these simple effects. For each age group, there were more experimental FAs when the target was presented 2 or 3 times than when it was presented 0, 1, or 4 times; moreover, for each age group, the numbers of experimental FAs did not vary across the 0-, 1-, and 4-presentations conditions. However, the two age groups differed in that the younger adults showed an increase in experimental FAs from 2 to 3 target presentations, whereas the older adults showed equivalent numbers of experimental FAs across these two frequency conditions. In summary, the predicted age-related increase in experimental FAs was obtained and was largest in the 2-target-presentations condition, when the competing experimental completion appeared three times.

The ANOVA for extra-experimental FAs revealed a marginally reliable main effect of age, $F(1,46) = 3.99$, $p < .06$, $MSE = 0.027$, and a reliable main effect of presentation frequency, $F(4,184) = 319.55$, $p < .001$, $MSE = 0.008$. The effect of age suggested that there were more intrusions of non-experimental responses for the older adults ($M_{old} = .24$) than for the younger adults ($M_{young} = .19$). Tukey tests revealed that the effect of presentation frequency occurred because the number of extra-experimental FAs decreased monotonically from 0 target presentations ($M_{0x} = .60$) to 2 target presentations ($M_{2x} = .06$) and then remained constant from 2 presentations to 4 presentations ($M_{4x} = .05$).

The ANOVA for non-responses revealed a marginally reliable main effect of age, $F(1,46) = 3.87$, $p < .06$, $MSE = 0.001$, a main effect of presentation frequency, $F(4,184) = 4.53$, $p < .01$, $MSE = 0.001$, and an Age x Presentation frequency interaction, $F(4,184) = 2.65$, $p < .05$, $MSE = 0.001$. The effect of age suggested that the older adults had a higher rate of non-responses ($M_{old} = .01$) than did the younger adults ($M_{young} = .003$). Tukey tests revealed that the effect of presentation frequency occurred because the likelihood of not responding was greater for the 0-presentations condition ($M_{0x} = .02$) than for the 2-presentations condition ($M_{2x} = .002$); no other pairwise differences were statistically reliable. Tests of simple effects assessed the nature of the two-way interaction. First, pairwise comparisons revealed that the age-related increase in experimental FAs was marginally reliable in the 0-presentations condition ($M_{old\text{-}0x} = .03$ vs. $M_{young\text{-}0x} = .005$), $t(46) = 1.91$, $p < .07$, but was not reliable in the other frequency conditions (all

$\underline{t}$s < 1.49, $\underline{p}$s > .14). (No $\underline{t}$-test could be performed to assess age differences for the 3-presentations condition, because the younger adults had a mean of exactly zero in this condition.) Secondly, two separate one-way ANOVAs assessed the effect of frequency for each age group. For the younger adults, this analysis omitted the 3-presentations condition because there was no variance in this cell. A main effect was obtained for the older adults, $\underline{F}(4,92) = 3.79$, $\underline{p} < .01$, $\underline{MSE} = 0.001$, but not for the younger adults, $\underline{F}(3,69) = 1.12$, $\underline{p} > .34$, $\underline{MSE} = 0.0001$. Tukey tests revealed that the simple main effect for the older adults occurred because the likelihood of not responding was greater in the 0-presentations condition than in the other four frequency conditions. However, regardless of age or experimental condition, the number of non-responses was at floor in this experiment.

Estimates of recollection and automatic influences. Table 2.11 reports mean estimates of recollection and automatic influences by age group. It was not possible to derive unambiguous estimates of recollection and habit for each cloze value because this variable was confounded with item type (i.e., the target versus the alternative experimental completion for a sentence) in this experiment. This was problematic for Hay and Jacoby's (1996) PDP logic because their equations assume that a given instance of remembering reflects recollection and habitual influences from the same response category, rather than different response categories for each type of influence.[2.3] However, it was possible to compute overall estimates of the two types of

---

[2.3] It was not possible to arrive at separate estimates of recollection and habit for each cloze value of Experiment 2. This is the case because there was a complete confounding of cloze value with response type (i.e., designation of a sentence completion as either the target completion, or the alternative experimental completion). In other words, for each sentence, the completion at one cloze value (either low- or medium-cloze) was the target completion, whereas the completion at the other cloze value was the alternative completion.

Therefore, correct recall of a thrice-trained target item reflects recollection and/or habitual influences for an item of a particular cloze value. The probability of recall in each thrice-trained condition is expressed as follows, where $\underline{R}_L$ and $\underline{R}_M$ refer to recollection of the low-cloze and medium-cloze completions, respectively, and $\underline{H}_L$ and $\underline{H}_M$ refer to habitual influences of the low-cloze and medium-cloze completions, respectively:

$$P(\text{Recall of Thrice-trained/Low-cloze completion}) = \underline{R}_L + \underline{H}_L*(1-\underline{R}_L) = \underline{R}_L + \underline{H}_L - \underline{H}_L*\underline{R}_L$$

$$P(\text{Recall of Thrice-trained/Medium-cloze completion}) = \underline{R}_M + \underline{H}_M*(1-\underline{R}_M) = \underline{R}_M + \underline{H}_M - \underline{H}_M*\underline{R}_M$$

memory process, collapsing across cloze values. Therefore, two separate one-way (Age) ANOVAs were conducted, one for recollection and one for habit. These analyses revealed that recollection was higher for younger than older adults, $\underline{F}(1,46) = 7.64$, $\underline{p} < .01$, $\underline{MSE} = 0.035$, but that habit was equivalent for the two age groups, $\underline{F}(1,46) < 1$. The finding of an age-related decline in recollection, albeit with age-invariance in habitual influences, has been reported in several other studies using the PDP (see, e.g., Jacoby et al., 1997, for a review).

With further regard to automatic influences, an additional analysis sought converging evidence for the derived estimates of habit. As described above, both the "guessing" scores (recall in the training-and-test-only condition) and the derived habit estimates were assumed to measure the same automatic influences (Hay & Jacoby, 1996). Therefore, a two-way (Measure x Age) MANOVA was conducted, with each factor having two levels each. Age served as a between-subjects factor, whereas measure (habit vs. "guessing") served as a within-subjects factor. This analysis failed to reveal any reliable effects, either for age, $\underline{F}(1,46) = 1.97$, $\underline{p} > .16$,

---

In contrast, when the target completion is once-trained, intrusions of the alternative (thrice-trained) completion reflect habitual influences of an item at one cloze value, when there is no recollection of the (target) item at the other cloze value. In other words, recollection now corresponds to one cloze value, whereas habit corresponds to the other cloze value:

P(Intrusions of Thrice-trained/Low-cloze completion) = $\underline{H}_L*(1-\underline{R}_M)$ = $\underline{H}_L - \underline{H}_L*\underline{R}_M$

P(Intrusions of Thrice-trained/Medium-cloze completion) = $\underline{H}_M*(1-\underline{R}_L)$ = $\underline{H}_M - \underline{H}_M*\underline{R}_L$

It is therefore not possible to estimate recollection for a particular cloze value by subtracting the probability of intrusions from the probability of correct recall. For example, subtracting intrusions of low-cloze completions from recall of low-cloze completions yields the following disappointing result:

P(Recall minus Intrusions of Low-cloze completion) = $\underline{R}_L + \underline{H}_L*(\underline{R}_M - \underline{R}_L)$

The above result indicates that the subtractive logic of Hay and Jacoby's (1996) PDP equations fails when the two alternative responses for a stimulus come from different response categories. In other words, the simple subtraction of intrusions from correct recall for a particular response category fails to eliminate the memory influences (i.e., $\underline{R}$ and $\underline{H}$) for the other response category.

Nevertheless, it is possible in this experiment to obtain overall estimates of recollection and habit for each participant by collapsing across cloze values. In other words:

P(Correct recall of thrice-trained completion) = $\underline{R} + \underline{H}*(1-\underline{R})$

P(Intrusions of thrice-trained completion, in once-presented condition) = $\underline{H}*(1-\underline{R})$

Recollection and habit can therefore be estimated through simple algebra:

$\underline{R}$ = P(Recall of thrice-trained completion) – P(Intrusions of thrice-trained completion)

$\underline{H}$ = P(Intrusions of thrice-trained completion) / $(1-\underline{R})$

Table 2.11
Experiment 2: Estimates of Recollection and Automatic Influences

Estimate and age group	**Overall**
	M (SD)
Recollection	
Young	.72 (.19)
Old	.57 (.19)
Habit	
Young	.69 (.19)
Old	.65 (.18)
"Guessing"	
Young	.70 (.14)
Old	.62 (.15)

Note. Recollection and habit were derived based on equations provided by Hay and Jacoby (1996). "Guessing," in Hay and Jacoby's terminology, refers to performance in the training-and-test-only condition (i.e., recall of items repeated at training but not at study) and is expected to correspond with derived estimates of habit. It was not possible to derive separate estimates of recollection and habit for each cloze value because this variable was completely confounded with item type (i.e., target vs. alternative experimental completion) in this experiment (see Footnote 2.3).

MSE = 0.042, or for measure or the Age x Measure interaction, Fs < 1. Note, however, that in the above analyses for the training-and-test-only condition, there was a marginally reliable age difference in hits favoring the young, $F(1,46) = 3.74$, $p < .06$, MSE = 0.021. This analysis of "guessing" scores therefore yields a different result than either the analyses for habit or the combined analyses for habit and "guessing." It therefore appears that there is a non-reliable trend in these data toward an age-related decline in automatic influences, but that this age difference is less than that observed for recollection.

Neuropsychological task performance. Table 2.12 indicates participants' performance on the three neuropsychological tasks (the Stroop, Trails, and verbal fluency tasks), as a function of age group and performance measure.

In the Stroop task, there were two baseline conditions (reading color words presented in black ink, and naming the colors of colored rectangles) and one interference condition (naming the ink colors in which color words were presented). For the baseline conditions, the older adults were slower than the younger adults, both when reading words in black ink, $t(46) = 3.04$, $p < .01$, and when naming the colors of rectangles, $t(46) = 2.70$, $p = .01$. Both age groups showed perfect accuracy in the reading baseline and were equally accurate in the color-naming baseline, $t(46) < 1$, with near-perfect accuracy in this condition. For the interference condition, the older adults were slower than the younger adults, $t(46) = 6.12$, $p < .001$, but were as accurate as the younger adults, $t(46) = 1.17$, $p > .24$, in naming color words presented in incompatible ink colors. With further regard to the Stroop interference effect, the older adults showed a larger increase in mean RT from the color-naming baseline to the incongruent color-naming condition ($M = 12.17$, $SD = 5.70$) than did the younger adults ($M = 4.75$, $SD = 2.24$), $t(46) = 5.94$, $p < .001$. When the interference effect was analyzed in relative rather than absolute terms, the older adults showed a greater proportion increase in mean RTs across these two conditions ($M = 1.41$, $SD = 0.65$) than

Table 2.12
Experiment 2: Neuropsychological Task Performance

Measure	Age Group				t
	Young		Old		
	M	(SD)	M	(SD)	
Stroop Task:					
Word-Reading Baseline					
RT	5.54	(0.81)	6.52	(1.37)	3.04**
Accuracy	1.00	(0.00)	1.00	(0.00)	-----
Color-Naming Baseline					
RT	7.49	(1.57)	8.77	(1.72)	2.70*
Accuracy	0.995	(0.02)	0.99	(0.02)	0.46
Interference *(Incongruent)*					
RT	12.24	(2.93)	20.94	(6.33)	6.12***
Accuracy	0.97	(0.05)	0.95	(0.06)	1.17
Trails Task:					
Trails A					
RT	20.96	(6.76)	36.86	(12.98)	5.32***
Errors	0.13	(0.34)	0.13	(0.34)	0.00
Trails B					
RT	53.49	(18.73)	90.68	(32.16)	4.90***
Errors	0.54	(1.47)	0.50	(0.89)	0.12
Verbal Fluency Task:					
Phonemic					
Total Generated	44.29	(11.34)	47.96	(11.83)	1.10
Perseverative Errors	1.29	(2.24)	1.08	(1.47)	0.38
Clustering	0.28	(0.11)	0.61	(0.60)	2.64*
Switching	34.25	(9.40)	29.92	(8.62)	1.66
Semantic					
Total Generated	21.58	(4.62)	18.88	(5.34)	1.88⁺
Perseverative Errors	0.17	(0.38)	0.38	(0.82)	1.12
Clustering	1.45	(0.73)	2.07	(1.76)	1.62
Switching	8.79	(2.96)	6.75	(3.04)	2.36*

Note. Reaction times (RTs) for the Stroop task are the number of seconds taken to either name 16 colors or read 16 words. RTs for the Trails task are the number of seconds taken to connect 25 dots in sequence. The clustering and switching measures for the verbal fluency task refer to the mean cluster size and number of shifts between subcategories, respectively (see Troyer et al., 1997). t values are for younger versus older adults ($df = 46$); ⁺ = $p < .10$, * = $p < .05$, ** = $p < .01$, *** = $p < .001$.

did the younger adults ($M = 0.65$, $SD = 0.32$), $t(46) = 5.12$, $p < .001$. This finding of an age-related increase in Stroop interference RTs was reported in Experiment 1 and in published studies of aging and the Stroop effect (e.g., West & Baylis, 1998).

In the verbal fluency task, participants generated words either beginning with a particular letter (i.e., phonemic fluency) or belonging to a specific category (i.e., semantic fluency). For phonemic fluency, the older adults generated as many words, $t(46) = 1.10$, $p > .27$, as the younger adults. For semantic fluency, the older adults generated marginally fewer words, $t(46) = 1.88$, $p < .07$, than did the younger adults. The finding of smaller age differences in phonemic fluency than in semantic fluency has been reported elsewhere in the literature (e.g., Troyer et al., 1997). The two age groups showed equivalent numbers of perseverative errors for both phonemic fluency, $t(46) < 1$, and semantic fluency, $t(46) = 1.12$, $p > .26$: This age-equivalence in perseveration is likely due to the low numbers of errors committed on this task.

With further regard to the verbal fluency task, additional measures of clustering and switching were derived for both phonemic and semantic fluency, based on the work of Troyer et al. (1997). The specific subscales of clustering and switching are thought to correlate with temporal lobe and frontal-lobe functioning, respectively (Troyer et al., 1997). Cluster size refers to the mean number of successive responses belonging to the same phonemic or semantic subcategory, whereas switches refer to the number of shifts between phonemic or semantic clusters within a participant's verbal fluency protocol. For phonemic fluency, cluster size was larger for older than younger adults, $t(46) = 2.64$, $p < .05$, but number of switches was equivalent for the two age groups, $t(46) = 1.66$, $p > .10$. For semantic fluency, cluster size was equivalent for the two age groups, $t(46) = 1.62$, $p > .11$, but number of switches was greater for younger than older adults, $t(46) = 2.36$, $p < .05$. These results are entirely consistent with the verbal fluency results of Experiment 1 and of Troyer et al. (1997). Namely, for phonemic fluency, there were age differences in clustering favoring the old but no age differences in switching, whereas for semantic fluency, there were no age differences in clustering but age differences favoring the young for switching.

For the Trail Making test, there was one baseline condition (Trails A) that involved connecting consecutively numbered dots, and one task-switching condition (Trails B) that involved alternating between numbers and letters. There was no age difference in errors for either Trails A or Trails B, $\underline{t}$s < 1. However, the older adults were significantly slower than the younger adults in completing both Trails A, $\underline{t}(46) = 5.32$, $\underline{p} < .001$, and Trails B, $\underline{t}(46) = 4.90$, $\underline{p} < .001$. With further regard to Trail Making performance, the older adults showed a larger increase in mean RT from Trails A to Trails B ($\underline{M} = 53.81$ sec, $\underline{SD} = 25.22$) than did the younger adults ($\underline{M} = 32.52$ sec, $\underline{SD} = 17.78$), $\underline{t}(46) = 3.38$, $\underline{p} < .01$. However, when this increase in RT was assessed in relative terms (i.e., [Trails B – Trails A] / Trails A), the older adults showed the same proportion increase ($\underline{M} = 1.56$, $\underline{SD} = 0.73$) as did the younger adults ($\underline{M} = 1.70$, $\underline{SD} = 1.13$), $\underline{t} < 1$.

Correlational analyses. Table 2.13 shows selected correlations among variables for younger and older adults. The correlations of interest concerned the relations of four memory measures described for Experiment 1 (recollection, automatic influences, episodic influences of once-presented items, and intrusions of thrice-presented items) with the various demographic and neuropsychological measures collected. Experiment 2 employed the same individual-difference variables as Experiment 1, except for the addition of the Trails Making task. This task produced slowing in the Trails B condition, relative to the Trails A condition, due to the necessity of switching between numbers and letters in Trails B. For the purposes of these analyses, Trail Making interference was operationalized as the absolute difference in mean RT between the Trails B and Trails A conditions. The present analyses were expected to yield the same

Table 2.13

Experiment 2: Selected Correlations for Younger and Older Adults

Younger Adults

Memory measures	Age	Education	Vocabulary	Digit Symbol	Stroop Interference	Trail Making Interference	Verbal Fluency – Frontal	Verbal Fluency – Temporal
	Demographic and Neuropsychological Measures							
Recollection	-.20	-.09	.15	.34	-.17	**-.44***	.29	.06
Automatic	**-.41***	**-.36**$^{+}$	.04	.13	-.17	**-.41***	.19	-.30
Episodic	-.08	-.02	.22	**.46***	-.10	-.33	**.48***	-.07
Intrusions	-.20	-.21	-.18	-.30	.26	.17	-.34	-.11

Older Adults

Memory measures	Age	Education	Vocabulary	Digit Symbol	Stroop Interference	Trail Making Interference	Verbal Fluency – Frontal	Verbal Fluency – Temporal
	Demographic and Neuropsychological Measures							
Recollection	-.08	.21	.18	**.58****	**-.38**$^{+}$	-.30	.29	.29
Automatic	-.23	.17	.08	.28	**-.37**$^{+}$	-.24	.28	.09
Episodic	-.10	.06	.06	**.50***	**-.36**$^{+}$	-.29	.15	.34
Intrusions	-.11	-.26	-.30	-.35	**.39**$^{+}$	-.05	-.15	**-.45***

Note. Recollection is the PDP-derived estimate of consciously controlled memory influences. "Automatic" is a composite measure combining two z-score-transformed estimates of automatic memory influence: habit and "guessing". "Episodic" refers to hits in the episodic baseline, where items appeared only once, at study. Intrusions refers to the number of thrice-presented items erroneously recalled in the once-presented critical conditions. Stroop interference is the difference in mean RT between the incongruent (interference) condition and color-naming baseline of the blocked Stroop test. Trail making interference is the difference in mean RT between the Trails A and Trails B tasks. Verbal fluency-Frontal is a composite measure of frontal-lobe functioning that combines two z-score-transformed measures (generation and switching) from both the phonemic and semantic blocks of the verbal fluency test. Verbal fluency-Temporal is a composite measure of medial-temporal functioning that combines two z-score-transformed clustering measures (phonemic and semantic) from the verbal fluency test. $N = 24$ for each age group; $^{+} = p < .10$, $^{*} = p < .05$, $^{**} = p < .01$, $^{***} = p < .001$.

correlations expected in Experiment 1. Moreover, Trail Making interference was expected to show the same relations to the memory measures as was Stroop interference. Namely, this measure was expected to be negatively correlated with recollection and positively correlated with intrusions in recall.

For the younger adults, there were six correlations that were either reliable or marginally reliable. First, the episodic base rate was positively correlated with both perceptual speed, $r(22) = .46$, $p < .05$; and the verbal fluency frontal scale, $r(22) = .48$, $p < .05$. Next, there was a negative correlation between recollection and Trail Making interference, $r(22) = -.44$, $p < .05$. Finally, automaticity was negatively correlated with three measures: age, $r(22) = -.41$, $p < .05$; education, $r(22) = -.36$, $p < .09$; and Trail Making interference, $r(22) = -.41$, $p < .05$. Some of these correlations for younger adults are consistent with those obtained in Experiment 1 for older adults only: Namely, frontal lobe functioning appears related to recall of once-presented items, contrary to earlier predictions that this episodic baseline would largely reflect associative, medial-temporal processes. In contrast, other correlations are not predicted by either existing theory or data (e.g., changes in automatic influences with age and frontal-lobe functioning). However, all of the correlations for younger adults may be spurious, as each correlation ceases to be reliable when one or two outlying observations are removed from the sample ($|r|$s $< .28$, ps $>$ .23).

For the older adults, there were seven correlations that were either reliable or marginally reliable. First, perceptual speed was positively correlated with two memory measures: recollection, $r(22) = .58$, $p < .01$; and the episodic base rate, $r(22) = .50$, $p < .05$. The role of speed as a mediator of age differences in memory performance has been suggested by other researchers (e.g., Salthouse, 1996). Secondly, the intrusion of thrice-presented completions was negatively correlated with the verbal fluency medial-temporal scale, $r(22) = -.45$, $p < .05$. This correlation was somewhat contrary to prediction in the sense that intrusions were expected to correlate with frontal, rather than medial-temporal, functioning. Finally, there were four marginally reliable correlations involving Stroop interference: This measure was negatively correlated with three memory measures – recollection, $r(22) = -.38$, $p < .10$; automaticity, $r(22) = -.37$, $p < .10$; and the episodic base rate, $r(22) = -.36$, $p < .10$ – and was positively correlated with a fourth measure, intrusions, $r(22) = .39$, $p < .10$. It is somewhat surprising that Stroop interference bore some linear relation to all four memory measures; however, it is consistent with current theorizing that a frontal lobe measure should be correlated with both recollection and intrusions, which purportedly reflect strategic uses of memory. Moreover, the negative correlation between Stroop interference and episodic influences was previously obtained for the older adults in Experiment 1, so it appears that cued recall of once-presented completions may be more strategic than was originally thought. However, the relation between Stroop interference and automaticity is somewhat peculiar, given the view that recall of repeated stimuli should be largely independent of the frontal lobes.

Of the seven correlations described for the older adults, only three appear to be produced by outlying observations. First, the correlation between speed and episodic influences was unduly influenced by one older adult who performed quite poorly on both of these measures. When this individual was removed from the analysis, the correlation no longer remained reliable, $r(21) = .28$, $p = .19$. Secondly, the correlation between intrusions and the verbal fluency medial-temporal scale was unduly influenced by two older adults who showed high clustering scores but few experimental FAs. When these two individuals were removed from the analysis, the correlation was no longer reliable, $r(20) = -.33$, $p = .13$. Thirdly, the correlation between Stroop interference and intrusions was unduly influenced by one older adult who was quite slow in the interference condition of the Stroop task and who produced many experimental FAs. When this individual was removed from consideration, the correlation was no longer reliable, $r(21) = .17$, $p = .44$.

Discussion

To summarize the main findings for Experiment 2, recall in the critical conditions was higher for younger than older adults. In addition, recall was higher when studied sentence completions were thrice-trained, rather than once-trained, and when target completions were medium-cloze, rather than low-cloze. These main effects were qualified by a Presentation Frequency x Cloze interaction, indicating that the benefit of medium-cloze over low-cloze completions was reliable when target completions were once-trained, but not when they were thrice-trained. Moreover, there was a three-way interaction (Age x Presentation frequency x Cloze value), indicating that age differences in recall were reliable in all critical conditions, except in the Thrice-trained/Low-cloze condition (M age difference = .06). Post-hoc analyses further revealed that the age difference in the Once-trained/Low-cloze condition (M difference = .14) was larger than that in the other three critical conditions (M difference = .08). This result was somewhat surprising: Although a three-way interaction was expected, it was predicted that the smallest age difference would appear in the Thrice-trained/Medium-cloze condition (M difference = .08), which had the highest levels of experimental and extra-experimental expertise. However, the older adults did not show any added benefit from schema-congruent sentence completions, as compared to schema-incongruent completions, when these target completions appeared thrice at training. This result is somewhat inconsistent with Craik and Jennings' (1992) formulation of the environmental support hypothesis, which predicts that age differences should be minimal when conditions are highly supportive (e.g., identical semantic cues at study and test). Nonetheless, the results are in line with Craik and Jennings' suggestion that older adults should be especially penalized by open-ended conditions – in this case, the use of sentence completions that were both infrequently presented and inconsistent with prior knowledge. Finally, although there were no other reliable interactions in the analyses for hits, it was interesting that the age difference in recall was numerically smaller for the thrice-trained completions than for the once-trained completions. In contrast, the size of the age difference was equivalent for the two cloze conditions. In other words, the older adults appeared to benefit somewhat more from repetition than did the younger adults; however, both age groups benefited equally from target completions that were schema-congruent, as compared to schema-incongruent completions.

With further regard to the critical conditions, the likelihood of experimental false alarms (FAs) was greater for older than younger adults. Moreover, there were more intrusions of the alternative experimental completion when the target completion was low-cloze, rather than medium-cloze, and when the target was once-trained, rather than thrice-trained. Moreover, there was an Age x Presentation frequency interaction, indicating that the age-related increase in intrusions was reliable for once-trained completions but was not reliable for thrice-trained completions. Furthermore, although there was no three-way interaction for experimental FAs, the older adults were especially likely to intrude alternative experimental completions in the Once-trained/Low-cloze condition ($\underline{M}$ = .29), as compared to the other three critical conditions ($\underline{M}$ = .14). Finally, there was an age-related increase in extra-experimental FAs that did not vary across experimental conditions. These results are consistent with Hasher and Zacks' (1988) claim that older adults have an inhibitory deficit, which makes them less able to suppress irrelevant information.

The critical conditions also allowed for the calculation of two types of memory influence – recollection and habit – based on equations described by Jacoby and Hay (1996). However, it was not possible to calculate these estimates separately for each cloze value (see Footnote 2.2).

Nonetheless, it was possible to derive overall estimates of recollection and habit for each participant. These estimates revealed a reliable age decrement for recollection (M age difference = .15), but not for habit (M difference = .04). Although another estimate of automatic influences – the "guessing" or training baseline – revealed a marginally reliable age decrement (M difference = .08), this age difference is still smaller than that obtained for recollection. This finding of a reliable age difference in recollection with little or no age difference in automatic influences is consistent with most other studies of aging and the PDP (see, e.g., Jacoby et al., 1997, for a review).

With regard to the baseline conditions, the two age groups performed equivalently poorly in the test-only condition, in which the target completions did not appear during the experiment. In contrast, an age difference favoring the younger adults emerged in the study-and-test-only condition, as well as in the training-and-test-only condition. Although an age-related decline in recall was expected when the target completion appeared only once (i.e., at study only), it was somewhat surprising to have found an age difference when the target was repeated (i.e., at training only). This result suggests that the older adults in the present experiment had more difficulty acquiring associations between sentences and their completions than did the younger adults, regardless of the frequency with which these sentences were presented. On the other hand, there was no evidence for an age-related increase in experimental FAs in these baseline conditions. However, there was an age-related increase in extra-experimental FAs in the episodic and training baselines.

Additional analyses examined memory performance as a function of the total number of target presentations. These analyses, which bridged across the critical and baseline conditions, revealed that recall increased under conditions of young adulthood, as compared to old age, and of frequent target presentation, as compared to infrequent presentation. Although the Age x

Presentation frequency interaction failed to attain statistical significance, the magnitude of age differences in each frequency condition was nonetheless of some theoretical interest. Secondary analyses revealed that age differences were non-reliable at 0 target presentations and became reliable at 1 and 2 target presentations but were only marginally reliable at 3 and 4 presentations. In other words, both age groups benefited from stimulus repetition, but the older adults benefited less than the younger adults did. In addition, the presentation frequency analyses revealed an age-related increase in experimental FAs that was largest in the 2-target-presentations condition, where there were three competing presentations of the alternative experimental completion. Moreover, these analyses also revealed an age-related increase in extra-experimental FAs.

The present experiment supported both the impaired-recollection and impaired-inhibition accounts of aging and memory. There was evidence for reliable age declines in the PDP-derived estimate of recollection, as well as an age-related increase in false alarms, particularly when the target item was trained once only and the competing experimental completion was trained thrice. The finding that the older adults were less able than the younger adults to recollect repeated target items and to inhibit repeated distractor items is consistent with Hay and Jacoby's (1999) claim that age-related memory difficulties stem from general failures of cognitive control.

The present data are also somewhat consistent with Craik and Jennings' (1992) formulation of the environmental support hypothesis. Namely, the older and younger adults performed equally poorly when conditions were minimally supportive (i.e., when producing sentence completions that were never presented in the experiment). However, when somewhat greater support was provided (i.e., one or two presentations of a target item), the younger adults benefited more than the older adults did. Moreover, at the highest levels of support (i.e., reinstatement of both experimental and pre-experimental expertise, as in the Thrice-trained/Medium-cloze condition) the magnitude of the age difference was smaller but was still reliable. This latter result was somewhat surprising, given the finding in Experiment 1 of equivalently high recall in younger and older adults when relatively unpredictable sentence completions were repeated. It was expected that the addition of moderately predictable sentence completions in the present experiment would produce an even more robust demonstration of compensation, or greater benefits to older than younger adults. However, the present experiment failed even to replicate the earlier finding of age-equivalence in recall of repeated, low-probable sentence completions. It is possible that the lesser education of the older adults in the present experiment ($\underline{M}$ = 15.2 years) as compared to those previous experiment ($\underline{M}$ = 16.0 years) made the present sample of older participants less capable of benefiting from environmental support. This possibility is in line with Craik, Byrd, and Swanson's (1987) suggestion that the degree to which memory performance is "repaired" in older adults depends on both task and subject parameters. Alternatively, it is possible that the older adults were more penalized than the younger adults by the longer retention interval produced by the addition of a fifth interpolated task (i.e., Trail Making). A further possibility is that the use of high-probable completions, rather than moderately probable completions, may have been required to eliminate age differences in the present experiment.

General Discussion

Two experiments employed a variant of Hay and Jacoby's (1996) process dissociation procedure to investigate age differences in the facilitation and interference conditions of a direct memory test. In both experiments, younger and older adults underwent a training phase in which sentences were presented four times each: three times with one completion, and one time with an alternative completion. The trained sentences then reappeared at study, with either the thrice-trained completion or the once-trained completion. It was predicted that age differences in recall would be smaller for the thrice-trained completions (i.e., a facilitation condition) than for the once-trained completions (i.e., an interference condition). There was a trend in this direction for both experiments; however, in both cases, the Age x Presentation frequency interaction failed to attain statistical significance.

In addition to varying presentation frequency, the two experiments each varied a particular stimulus attribute that was expected to influence the magnitude of age differences in recall. In Experiment 1, the perceptual characteristics of stimuli were varied by presenting sentences in either the same voice or different voices across experimental phases. It was expected that repeatedly reinstating both the perceptual and conceptual attributes of a stimulus would yield a compensation pattern in which the older adults benefited more than the younger adults (see, e.g., Brown, 1994, and Shaw & Craik, 1989, for relevant findings). In Experiment 2, the extent to which participants could use prior knowledge to recall information was varied: Specifically, sentence completions were either medium-cloze (moderately predictable) or low-cloze (relatively unpredictable). It was expected that reinstating both experimental and pre-experimental expertise would benefit the older adults more than the younger adults (see, e.g., Hess, 1990, for similar ideas). However, these predictions were only partly borne out: In Experiment 1, age differences in recall were smaller when both the conceptual and perceptual characteristics of a stimulus were repeatedly reinstated, as compared to conditions that reinstated these characteristics to a lesser degree. However, the three-way interaction (Age x Presentation frequency x Voice context) failed to attain statistical significance. Similarly, in Experiment 2, there was some evidence that age differences in recall were reduced when information either was moderately consistent with prior knowledge *or* was frequently presented in the experiment, as compared to when information was low in both schema-congruency and presentation frequency. However, no further reduction in age differences was observed when the information was both schema-congruent and infrequently presented. In summary, the present results partly corroborated Craik and Jennings' (1992) notions of environmental support: Namely, the precise reinstatement of experimental contexts and utilization of existing knowledge reduced age differences, but did not eliminate them altogether.

It was somewhat surprising that age differences in recall were almost eliminated under highly supportive conditions in Experiment 1, but not in Experiment 2. For example, older and younger adults were almost equivalent in their recall of thrice-trained, voice-congruent sentence completions (M age difference = .01) in Experiment 1, but not in Experiment 2 (M age difference = .07). In addition, it was surprising that the addition of schema-congruent sentence completions in Experiment 2 was of no further benefit to the older adults. It is possible that the participants sampled in these two experiments differed along some individual-difference variable that affected their ability to utilize environmental support. Namely, in Experiment 1 the older adults were more highly educated than the younger adults, whereas Experiment 2 showed a trend in the opposite direction. The possibility that subject parameters such as education can influence the nature of Age x Condition interactions has been suggested by other researchers (e.g., Craik et al., 1987; Craik & Jennings, 1992).

With regard to the impaired-recollection hypothesis, both experiments provided some evidence for an age decrement in the PDP-derived estimate of recollection. However, this age difference was only reliable in Experiment 2. Moreover, both experiments revealed reliable age-related declines in recall for repeated sentences, as well as once-presented sentences (i.e., in the episodic baseline). This age difference in recall was eliminated only in the higher presentation-frequency conditions of Experiment 1 (i.e., 3 or 4 total target presentations). In contrast, the PDP-derived estimate of habit was relatively age-invariant in the two experiments. With regard to the impaired-inhibition hypothesis, both experiments showed a trend toward an age-related increase in experimental and extra-experimental false alarms. However, these age differences were reliable only in the second experiment. The finding of an age-related decrease in recall of repeated targets, together with an age-related increase in intrusions of repeated distractors, suggests that the older adults suffered from a general failure of cognitive control, rather than a more specific inhibitory failure (see, e.g., Hay & Jacoby, 1999, for a similar argument).

CHAPTER THREE:
AGE DIFFERENCES IN FACILITATION AND INTERFERENCE IN AN INDIRECT MEMORY TEST

Three experiments were conducted to assess age differences in an indirect memory test called the letter deletion task (Reingold, 1995). All three experiments examined age differences in interference priming, or the extent to which older and younger adults showed impairments in indirect test performance (e.g., speed or accuracy) as a result of prior learning. The use of an interference paradigm in an indirect test situation is somewhat novel, as most indirect tests in use are facilitation paradigms that assess improvements in speed and accuracy as a result of prior learning, also known as priming (see, e.g., Roediger & McDermott, 1993; and Schacter, 1987, for reviews). The one major exception to this reliance on facilitation paradigms is the negative priming paradigm (see, e.g., Tipper & Cranston, 1985), in which the response time to a target item increases if that target previously served as a distractor item that was to be ignored. However, negative priming is thought to primarily reflect inhibition, or the lingering effects of suppressing a response, because that suppression dissipates slowly (May et al., 1995; but see also, e.g., Neill & Valdes, 1992, for an alternative explanation). In contrast, the present investigation concerns an interference paradigm, in which the (automatic) effects of prior memory activations must be overcome at retrieval. Therefore, although inhibition may be involved in the present paradigm, it is thought to be initiated during the second appearance of a stimulus, rather than its first appearance (see May et al., 1995, and below).

The almost exclusive use of facilitation paradigms in indirect memory tests is problematic because these paradigms may not be "process pure" (see, e.g., Jacoby, 1991; Reingold, 1995). In other words, recent empirical work has cast doubt on the traditional claim that indirect tests reflect exclusively automatic or implicit memory processes. Instead, researchers like Jacoby and Reingold argue that most traditional indirect tests allow for the possibility of *explicit contamination* or the strategic use of information from a prior study episode. This strategic use of memory may differ as a function of age: Namely, younger adults may be more likely than older adults to exhibit "test awareness," or the realization that the task somehow relates to prior learning. The greater test awareness of younger adults may stem from this age group's greater familiarity with laboratory situations and may be responsible for the small but non-reliable age decrements reported in several studies of implicit memory (see, e.g., Light & La Voie, 1993, for a review).

One possible exemption from this possibility of explicit contamination in indirect tests is an interference-priming paradigm, such as Reingold's (1995) letter deletion task, which resembles the opposition conditions of Jacoby's (1991) process dissociation procedure. In Reingold's task, participants are presented with a nonword (e.g., "ANGILE") and must delete one of two specified letters (e.g., "ANGILE") in order to produce a word from the remaining letters. Participants are asked to respond as quickly and as accurately as possible. In the interference condition of this task, a previously presented nonword reappears, albeit with different letters specified as candidates for deletion than in the nonword's previous appearance (e.g., "ANGILE"). In other words, there are two possible solutions for each nonword (i.e., two different letters whose deletion will produce a word from the remaining letters), but only one of these two solutions is "legal" or permissible for any given trial. Reingold therefore argues that there is no conscious incentive to refer back to preceding trials when responding to the current stimulus. This is the case because any deliberate attempt to reference a previously encountered stimulus will only further impede the individual's performance (i.e., reduce speed and/or accuracy of responding), given the incompatibility of the two trials for that stimulus. It should be further added that Reingold's task is an indirect memory test in the sense that participants may not even be consciously aware that any of the previously presented items are related to those appearing subsequently. This is because the earlier trials do not constitute a distinct study phase but instead appear to belong to the same block of trials. If Reingold is correct in asserting that his task and other indirect interference paradigms are process pure, then one might expect that whatever small age differences present in facilitation priming studies would disappear altogether in the present investigation of interference priming. Such a claim would rest on the assumption that interference priming reflects only automatic uses of memory and that such automatic influences are age-invariant. In summary, accounts that attribute age differences in memory to impairments of deliberate recollection would predict age-equivalence (i.e., in speed and accuracy) in the interference conditions of Reingold's letter deletion task.

Alternatively, one might predict that the magnitude of interference priming would increase with age. This prediction rests on the impaired inhibition account of Hasher and Zacks (1988), which claims that older adults have more difficulty overcoming the effects of no-longer-appropriate information than do younger adults. If it is true that older adults have deficient inhibitory mechanisms that would make them less able to suppress no-longer-appropriate information, then older adults should show declines in the interference conditions of Reingold's task (in speed and/or accuracy) relative to younger adults. In other words, the older adults should show a greater increase in RT from the control to interference conditions than the younger adults, and/or a greater decrease in accuracy across these conditions. If older adults are indeed more prone to interference in Reingold's task and this age difference is attributed to declines in retrieval inhibition, then one must conclude either that inhibition need not always be consciously controlled or that frontal processes may sometimes operate in indirect tests (see, e.g., Winocur, Moscovitch, & Stuss, 1996).

A prediction of reduced interference priming with age was not seriously entertained. This prediction might be suggested by negative priming studies (e.g., Tipper & Cranston, 1985) in which younger but not older adults are slower in responding to a target that previously served as a distractor (e.g., Hasher et al., 1991). That is, negative priming is paradoxically greater for younger than older adults because the young better suppress the previously *irrelevant* information. However, Reingold's interference paradigm differs from the negative priming procedure in that it reflects the adverse effects of previously *relevant* information. To put it another way, the one paradigm is the "flip side" of the other. In Tipper's paradigm, the initially inappropriate response subsequently becomes appropriate (e.g., the stimulus to be ignored on one trial becomes the stimulus to respond to on the next trial). In contrast, in Reingold's paradigm, the initially appropriate response subsequently becomes inappropriate (e.g., the letter to be deleted upon a stimulus' first occurrence cannot be deleted later on). Moreover, Reingold's interference effects appear to persist over several minutes, whereas Tipper's effects typically occur over consecutive experimental trials. The likelihood of different underlying mechanisms for Tipper's task and Reingold's task suggests that the pattern of age differences should differ in the two paradigms. This analysis suggests that age differences should favor the old in Tipper's task (where older adults activate, rather than suppress, the initially inappropriate information) while favoring the young in Reingold's task (where older adults fail to overcome the activation of previously appropriate information).

In summary, it was unclear whether the pattern of age differences on the interference conditions of Reingold's (1995) task would better resemble indirect facilitation conditions (i.e., little or no age difference; see, e.g., Light & La Voie, 1993) or direct interference conditions (i.e., reliable age decrements; see, e.g., Winocur & Moscovitch, 1983). The extant literature is of little help in resolving this issue: Although interference has been measured within the same selective attention paradigm that has been used to measure inhibition, interference here is typically assessed in terms of performance costs produced by concurrent distraction, rather than prior learning. Moreover, even if the effects of concurrent distraction are relevant to the present investigation, conflicting results have been found: Some studies have found an age-related increase in interference (e.g., McDowd & Oseas-Kreger, 1991), whereas other studies have not (e.g., Kane et al., 1994). Empirical support for either of the two alternative predictions discussed above would therefore help to distinguish between impaired control versus impaired inhibition as an accurate account of age differences in memory.

Experiment 3

Experiment 3 was a simple study of age differences in interference priming using Reingold's (1995) letter deletion task. The study was therefore an extension of Reingold's third experiment in the sense that his basic procedures were employed, albeit with an additional age group: older adults. The present experiment focused on Reingold's interference condition. The facilitation condition of Reingold's third experiment was omitted, based on the suggestion that including a facilitation condition in a within-subjects design might produce a conscious incentive for participants to deliberately refer to prior trials (E. Reingold, personal communication, September 1996). Under such circumstances, consciously controlled retrieval might intrude upon an otherwise automatic retrieval situation. In other words, if the present study employed both facilitation and interference conditions and an age difference in interference was obtained, then this finding might be attributable to an age difference in deliberate recollection. Therefore, the present study included only an interference condition and a control condition, in order to ensure that the present results would not be vulnerable to the problem of explicit contamination (Jacoby, 1991).

Methods

Design. Figure 3.1 outlines the design and procedure for Experiments 3, 4, and 5. The experimental design for the present experiment consisted of one between-subjects factor and one within-subjects factor. The between-subjects factor was age (younger vs. older adults), and the within-subjects factor was experimental condition (Interference vs. Control). The Interference condition involved the use at test of a previously encountered nonword, albeit with a change in which letters were candidates for deletion. In contrast, the Control or baseline condition involved the use at test of a completely novel nonword.

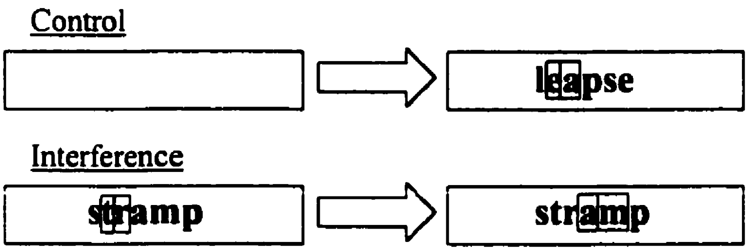
Participants. Demographic information on the 48 younger and 48 older adults is presented in Table 3.1. Older adults were healthy volunteers, the majority of whom were tested at a suburban campus of the University of Toronto and were offered $5 for their participation in a 30- to 45-min. session; the remaining older adults were tested at the downtown university campus and were offered reimbursement for their travel expenses. Younger adults were psychology undergraduates who received course credit. On a 20-item version of the Mill Hill Vocabulary Scale (Raven, 1965), the older adults scored higher than the younger adults, $t(94) = 2.02$, $p < .05$. In contrast, on a perceptual speed measure – the WAIS Digit Symbol Substitution Task (Wechsler, 1981) – the younger adults outperformed the older adults, $t(94) = 12.78$, $p < .001$. The finding of an age-related increase in vocabulary, despite a decrease in perceptual speed, was reported for the previous two experiments. The two age groups were equivalent in years of formal education, $t < 1$, and in self-reported health, $ts < 1.53$, $ps > .12$, as assessed for two time frames: the last two to three months, and today.

Materials. The present study employed the same stimuli previously used in Reingold's (1995) Experiment 3. For the critical stimulus set, five-letter words were paired according to one of two rules, as described by Reingold (1995). The first rule required that if the letters of the

Figure 3.1. Experiments 3, 4, and 5: Overview of design and procedure

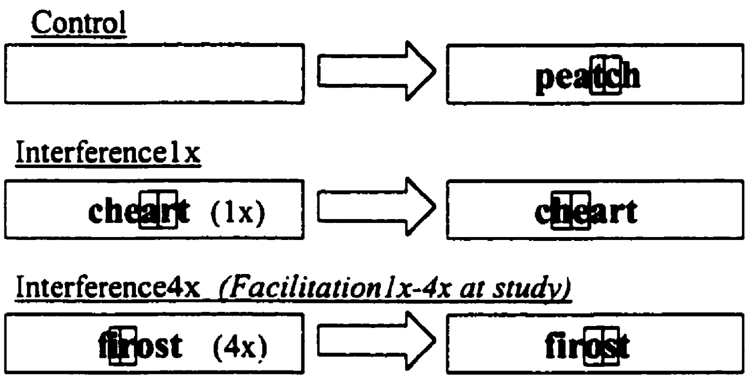
Study phase **Test phase** **Description**

Experiment 3: Interference priming



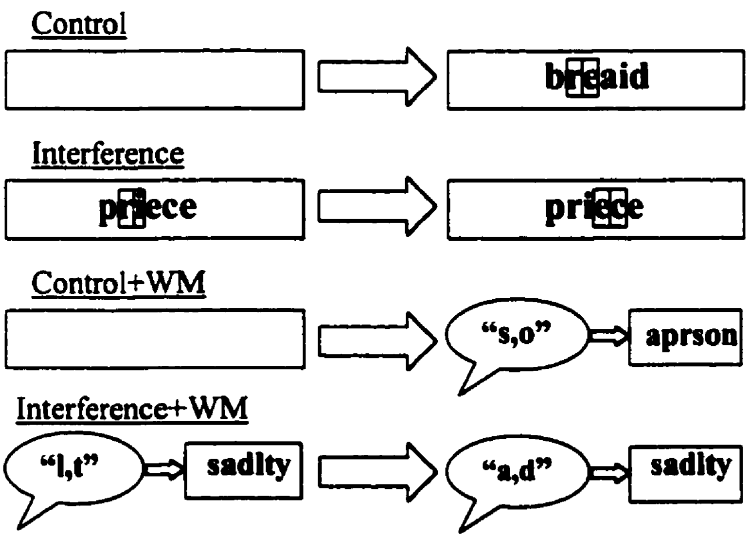
At test, two types of item appeared: previously unstudied items (*Control* condition) and previously studied items, albeit with a change in which letters were candidates for deletion (*Interference* condition).

Experiment 4: Facilitatory versus interfering effects of repetition



At test, three types of item appeared: unstudied items (*Control* condition), items studied once only (*Interference1x* condition), and items studied four times (*Interference4x* condition). In the interference conditions, study items reappeared, albeit with a change in the candidate letters for deletion. The repetition of study items for the *Interference4x* condition yielded further, *facilitation* conditions at study.

Experiment 5: Interference with and without a working memory load



The standard *Control* and *Interference* conditions from Experiment 3 were employed, along with two conditions that imposed a working memory (WM) load: *Control+WM* and *Interference+WM*. In the WM conditions, no boxes appeared to specify the candidate letters for deletion. Instead, the nonword was preceded by two auditorily presented letters that the participant repeated aloud and committed to memory. Then the nonword appeared and the participant decided whether the first or second of the two memorized letters could be deleted to form a word.

Table 3.1
Experiment 3: Mean Participant Characteristics

Variable	**Age Group**				**t**
	Young		Old		
	M (SD)	n	M (SD)	n	
Age	21.8 (4.0)	48	68.9 (6.1)	48	45.13***
Age range	18-35		60-84		-------
Gender:					-------
Female		36		39	
Male		12		9	
Handedness:					-------
Right-handed		42		43	
Left-handed		4		3	
Ambidextrous		2		2	
Education (years)	14.8 (2.3)		14.4 (2.3)		0.88
Self-reported health:					
Last 2 to 3 months	2.0 (0.8)		1.7 (0.8)		1.53
Today	1.9 (0.8)		1.7 (0.8)		1.33
Mill Hill Vocabulary Scale	14.3 (2.4)		15.2 (2.2)		2.02*
Digit Symbol Substitution Task	72.6 (9.7)		46.2 (10.5)		12.78***

Note. For gender, $\chi^2_{(1, N=96)} = .55$, $p > .30$. For handedness, $\chi^2_{(2, N=96)} = .15$, $p > .90$. Self-reported health was assessed for two time frames and was rated on a 5-point scale, ranging from 1 (excellent) to 5 (poor). Mill Hill Vocabulary Scale maximum = 20. Digit symbol score is the number of correct completions in 90 sec (maximum = 93). t values are for younger versus older adults ($df = 94$); * = $p < .05$, ** = $p < .01$; *** = $p < .001$.

first word are symbolized by ABCDE (e.g., LAPSE) then the letters of the second word should be AXBDE (e.g., LEASE) where X ≠ C. The pair of words is then combined into the string AXBCDE (e.g., LEAPSE), which can serve as a test cue for ABCDE if the second and third letters are underlined – AXBCDE, or for AXBDE if the fourth and fifth letters are underlined – AXBCDE (e.g., LEAPSE for LAPSE, LEAPSE for LEASE). Twenty-four pairs of five-letter words corresponding to this rule, a left-letter deletion rule, were selected and combined to create 24 pairs of test cues. The second rule required that if the letters of the first word are symbolized by ABCDE (e.g., STAMP), then the letters of the second word should be ABXCE (e.g., STRAP) where X ≠ C. The pair of words is then combined into the string ABXCDE, which can serve as a test cue for ABCDE if the second and third letters are underlined – ABXCDE, or for ABXCE if the fourth and fifth letters are underlined – ABXCDE (e.g., STRAMP for STAMP, STRAMP for STRAP). Twenty-four pairs of five-letter words corresponding to this rule, a right-letter deletion rule, were selected and combined to create 24 pairs of test cues. The mean frequency of the 96 potential solution words was 54 (SD = 92) (Francis & Kucera, 1982). An additional 24 buffer stimuli were employed, which were also previously devised by Reingold (1995). Like the critical stimuli, the buffer stimuli were 6-letter nonwords from which either of two letters could be deleted to form a 5-letter word from the remaining letters. However, only one of the two possible solutions was valid in any given trial. The stimuli for Experiments 3, 4, and 5 and their solution words are listed in Appendix 3.1.

Four experimental formats were generated in order to rotate each critical stimulus through the four conditions produced by crossing study status (studied vs. unstudied) with the underline pattern employed (either the second and third letters underlined at test, or the fourth and fifth letters). However, rather than employing underlines, as in Reingold's (1995) study, the present study surrounded the candidate letters with boxes in order to further delineate the nature of the task for the older adults, who sometimes have difficulty understanding experimenter instructions (see, e.g., Graf & Komatsu, 1994). Each participant was therefore presented with 24 study trials, in which 12 stimuli had the second and third letters surrounded by boxes and the other 12 had the fourth and fifth letters surrounded by boxes. Moreover, for each surround position, six stimuli required a left button response and six required a right button response. Study trials were preceded and followed by 12 buffer trials, which included an equal number of left and right button responses and the two surround patterns. The test phase consisted of 48 trials, 24 of which corresponded to the Interference condition and 24 to the Control condition. In the Interference trials of the test phase, previously encountered nonwords reappeared, albeit with different letters surrounded by boxes than in the nonword's previous appearance. For example, if a nonword's second and third letters were surrounded by boxes in the study phase, then the fourth and fifth letters would be surrounded in the test phase. The Control or baseline trials represented the 24 test cues that were not used in the study trials. The version of test cue selected for baseline trials was randomly determined with the same constraints as the study trials. A different random ordering of the study and test stimuli was employed for each experimental format.

Apparatus. The experiment was controlled by a 386 PC compatible microcomputer with a 14-in. color VGA monitor. Response times were recorded from the "1" and "2" keys of the computer keyboard. Participants initiated each trial by pressing the "ENTER" key of the keyboard. All programming was done using the Micro Experimental Laboratory (MEL) Professional Version 2.0 (Schneider, 1995).

Procedure. The procedure was virtually identical to that employed in the interference condition of Reingold's (1995) Experiment 3. The main difference between the present study and Reingold's is that boxes, rather than underlines, were used to specify the candidate letters for deletion on each trial. Participants were seated approximately 45 cm in front of a computer and completed a demographic questionnaire. Participants then received instructions on the letter deletion task and were presented with sheets of paper containing two sample letter deletion stimuli. Participants were informed that on each trial they would see a six-letter nonword on the computer screen, with two letters surrounded by boxes. They were to indicate whether the left or right surrounded letter could be deleted from the nonword to form a word by pressing the "1" or "2" key of the keyboard, respectively. Participants were further informed that they should use the index and middle fingers of their dominant hand to select between these two alternative keys on each trial. Moreover, participants were to press the "ENTER" key of the keyboard with their non-dominant hand to initiate each trial. Therefore, right-handed participants used the "1" and "2" keys of the numeric keypad and the "ENTER" key of the alphanumeric layout, whereas left-handed participants used the "1" and "2" keys of the alphanumeric layout and the "ENTER" key of the numeric keypad. Participants were further instructed that they should respond as quickly as they could without making errors. The experimenter's instructions to participants in Experiments 3 and 4 are shown in Appendix 3.2.

Participants solved the two sample stimuli and then completed the letter deletion trials. The sequence of 12 buffer trials, 24 study trials, 12 additional buffer trials, and 48 test trials was presented as a single block of 96 trials. There was no break between study and test. Participants then completed short tests of vocabulary and perceptual speed before being debriefed. The experiment took no more than 45 min. per individual session.

In each trial, a six-letter nonword was presented in black lower-case letters on a white background in the center of the computer screen. Stimuli were presented in a large MEL font (Rome39) of less than 36x36 pixels per character. In each stimulus, two adjacent letters (either the second and third letters, or the fourth and fifth letters) were surrounded by boxes that were approximately 36x36 pixels in size. The deletion of one of the surrounded letters, but not the other, created a word.

The trial sequence was as follows: A fixation cross was presented in the center of the screen. Participants initiated the trial by pressing the "ENTER" key on the computer keyboard. Approximately 200 ms after the "ENTER" key was pressed, an experimental stimulus replaced the fixation cross in the center of the computer screen. The stimulus was displayed until a response was made. The screen was then blanked for 500 ms followed by the reappearance of the fixation cross.

Results

Test phase performance. Figure 3.2 shows the proportion of correct test trials as a function of age group and experimental condition. A 2x2 (Age x Condition) ANOVA for accuracy failed to reveal any effects of age or condition or their interaction, $Fs < 1.01$, $ps > .31$. In other words, accuracy did not differ across the two age groups, nor did it differ across studied and unstudied items in the test phase.

Figure 3.3 shows the mean and median reaction times (RTs) for correct test trials by age group and experimental condition. A 2x2 (Age x Condition) ANOVA for correct mean RTs revealed main effects of age, $F(1,94) = 51.60$, $p < .001$, $MSE = 3747000$, and condition, $F(1,94) = 17.24$, $p < .001$, $MSE = 224782$, but no interaction of these two variables, $F < 1$. These effects revealed that the mean RT was greater for older than younger adults and for Interference than Control trials. Subsidiary analyses for mean RTs revealed that the absolute difference in RTs between the Interference and Control conditions was numerically larger for the older adults (M = 327 ms, SD = 819) than for the younger adults (M = 242 ms, SD = 478), but that this was not a statistically reliable difference, $t(94) < 1$.

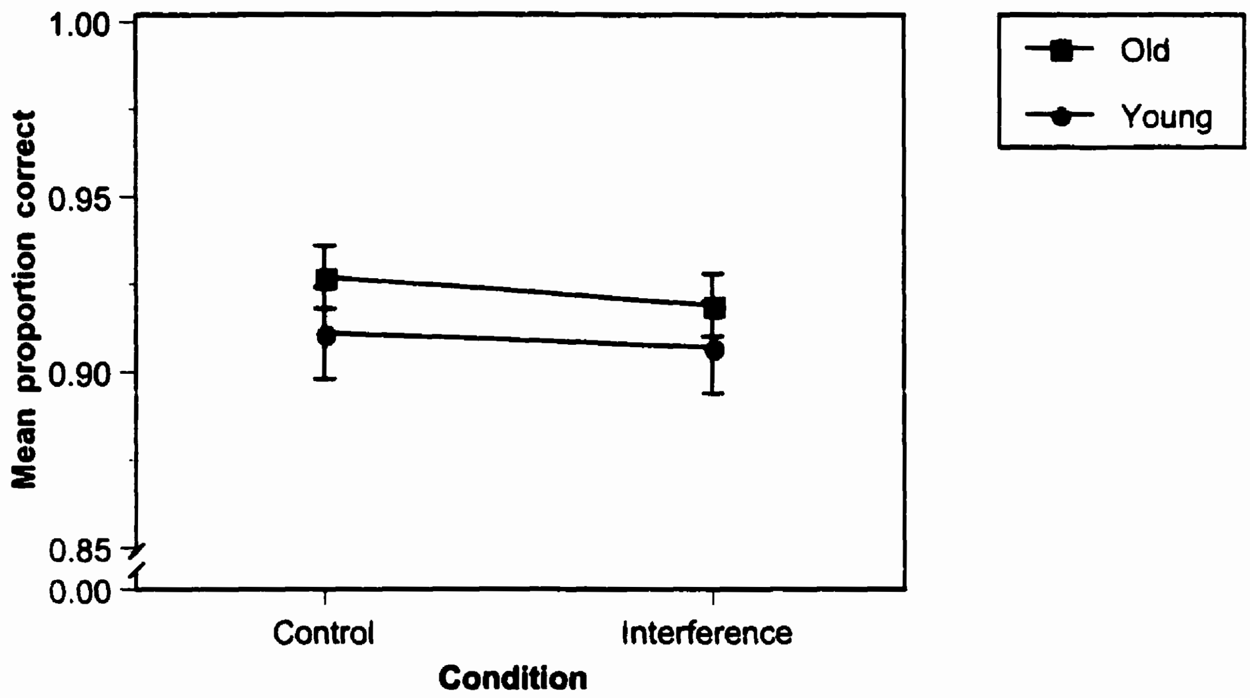
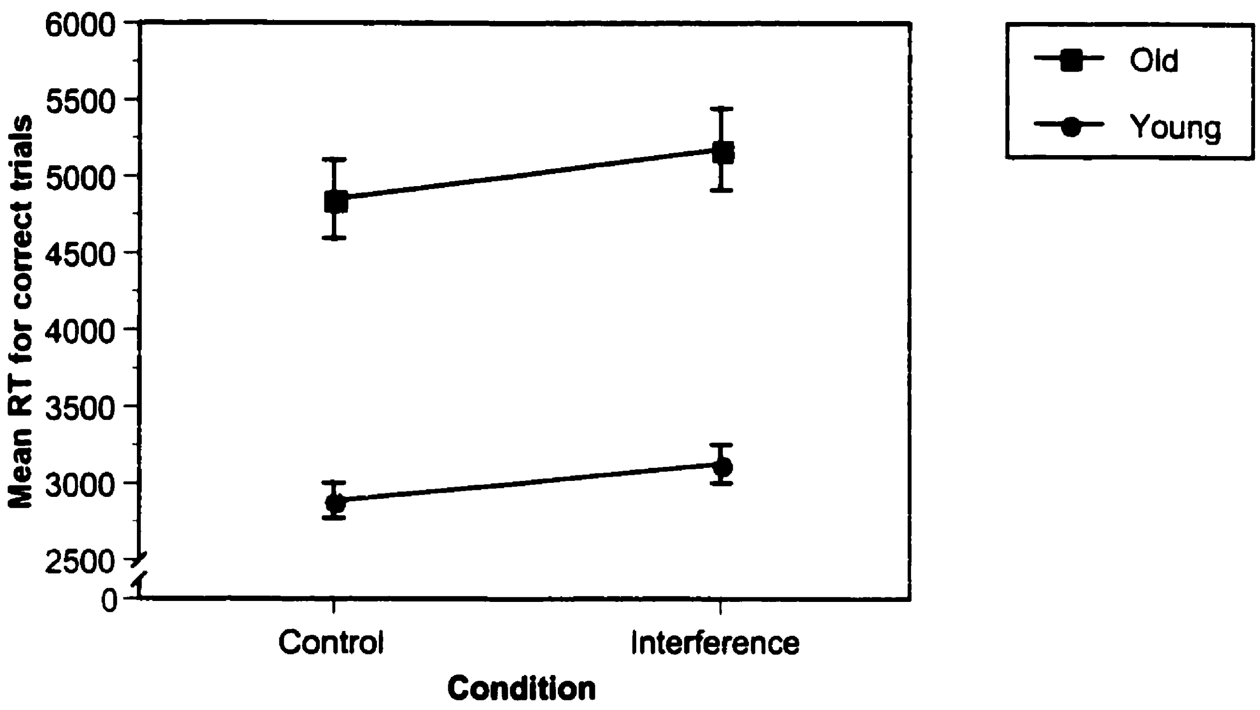


Figure 3.2. Experiment 3: Mean proportion correct test trials, by age group and condition. Error bars of one standard error are plotted for each point.

A. Mean RTs



B. Median RTs

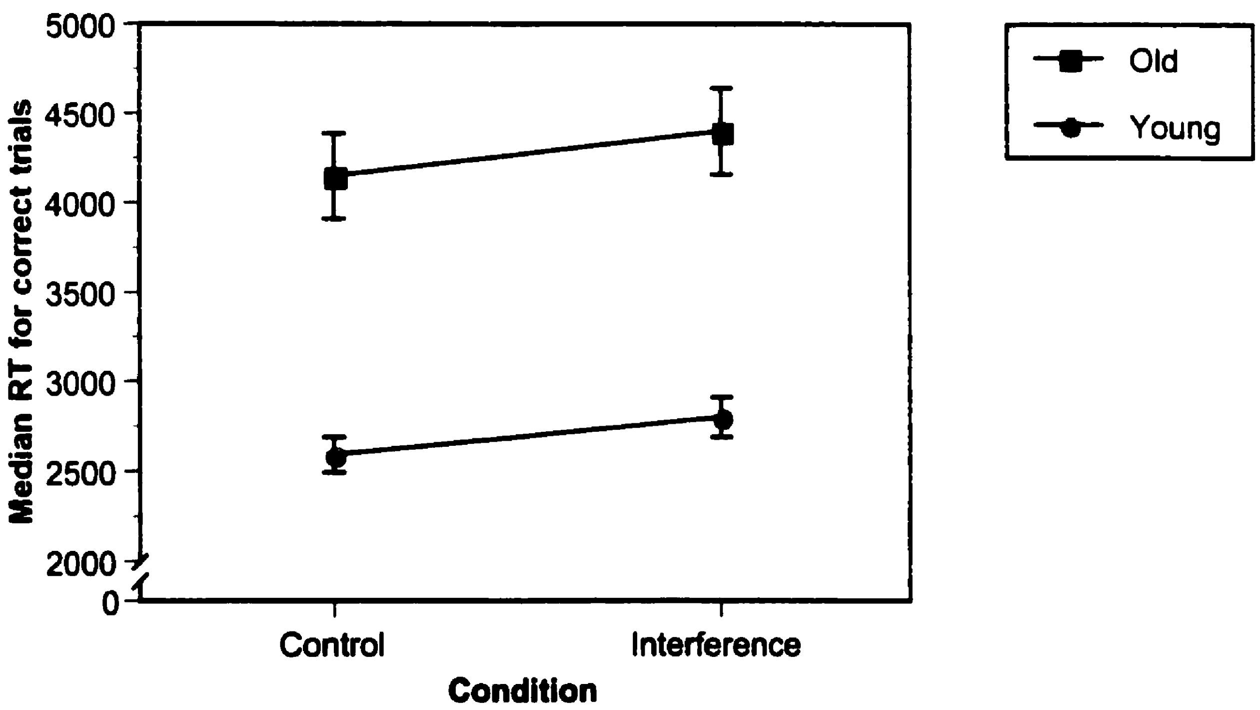


Figure 3.3. Experiment 3: Mean and median RTs for correct test trials, by age group and condition. Error bars of one standard error are plotted for each point.

An Age x Condition ANOVA for median correct RTs revealed effects similar to those found for mean RTs: Namely, there were effects of age, $F(1,94) = 39.87$, $p < .001$, $MSE = 3001532$, and condition, $F(1,94) = 9.11$, $p < .01$, $MSE = 279211$, in the same directions as before, but no two-way interaction, $F < 1$. Subsidiary analyses for median RTs revealed that the absolute magnitude of interference priming was numerically larger for the older adults ($M = 251$ ms, $SD = 920$) than for the younger adults ($M = 209$ ms, $SD = 521$), but that this was not a statistically reliable difference, $t(94) < 1$.

Further analyses examined whether an age difference emerged when the magnitude of interference priming was assessed in relative, rather than absolute terms. In other words, because older adults' baseline RTs were generally higher than those of younger adults, a more conservative way to depict within-participant differences from baseline was to calculate proportional slowing scores. Therefore, each participant's absolute priming score, or the average difference between the interference and baseline conditions, was adjusted by dividing by that participant's average baseline RT. These proportional slowing scores were calculated for both mean and median RTs. The analyses for relative differences in mean RTs revealed that the older adults showed as large an increase in RT from the Control to Interference conditions ($M = .083$, $SD = .173$) as did the younger adults ($M = .092$, $SD = .153$), $t(94) < 1$. Similarly, the analyses for relative differences in median RTs revealed that the older adults showed as large an increase in RT in the Interference condition ($M = .082$, $SD = .210$) as did the younger adults ($M = .097$, $SD = .199$), $t(94) < 1$.

Correlational analyses. Table 3.2 shows selected correlations among variables for younger and older adults. The correlations of interest concerned the relation between interference priming (operationalized as the difference between mean or median RTs in the Interference and Control conditions) with the various demographic measures collected. It was

Table 3.2
Experiment 3: Selected Correlations for Younger and Older Adults

Younger Adults

Memory measure		***Demographic Measures***		
	Age	Education	Vocabulary	Digit Symbol
Interference				
Difference between mean RTs	-.24	-.19	**-.47***	.04
Difference between median RTs	**-.36***	**-.24^{+}**	**-.29***	.08

Older Adults

Memory measure		***Demographic Measures***		
	Age	Education	Vocabulary	Digit Symbol
Interference				
Difference between mean RTs	-.03	.09	-.05	-.03
Difference between median RTs	-.15	.01	-.01	.06

Note. The interference effect is the increase in mean or median RT from the Control to the Interference conditions of the letter deletion task. N = 48 for each age group; $^{+}$ = $p < .10$, * = $p < .05$, ** = $p < .01$, *** = $p < .001$.

speculated that for both age groups, and especially the older adults, interference priming would be positively correlated with chronological age and negatively correlated with the remaining three variables of interest (education, vocabulary, and perceptual speed).

For the younger adults, four correlations were either reliable or marginally reliable. First, vocabulary was negatively correlated with interference priming, regardless of whether it was operationalized as the difference between mean RTs, $r(46) = -.47$, $p < .05$; or the difference between median RTs, $r(46) = -.29$, $p < .10$. These negative correlations imply that highly verbal adults were less likely to exhibit interference than were their less verbal peers. It is possible that the more verbal younger people processed the stimuli in a more global way than their less verbal peers did. In other words, if the highly verbal younger adults considered multiple solutions on each trial, rather than the one legal solution for each trial, then they may have been more likely to show facilitation than interference. Secondly, education was negatively correlated with interference when it operationalized as the difference between two median RTs, $r(46) = -.24$, $p < .10$; as was chronological age, $r(46) = -.36$, $p < .05$. It is possible that high levels of education and greater life experience made the younger adults less prone to functional fixedness, or single-minded consideration of one and only one solution on each trial. However, the roles of age and education appear to have been somewhat weak, given the failure to correlate these two variables with the mean RT difference between Interference and Control conditions. Moreover, the correlation between education and the median RT difference was unduly influenced by an outlier: One highly educated younger adult showed substantial facilitation, rather than interference. When this individual was omitted from the analysis, the correlation was no longer reliable, $r(45) = -.11$, $p = .46$. No correlations of interest were reliable for the older adults.

Discussion

In summary, the present experiment replicated Reingold's (1995) finding of interference priming, or an increase in RT from the Control to Interference conditions, in the letter deletion task. In addition, accuracy remained constant, regardless of experimental condition or age group. Moreover, the older adults were slower than younger adults when performing the letter deletion trials, as might be predicted from a large literature showing age-related slowing effects (see, e.g., Salthouse, 1996, for a review). However, there was no age difference in the magnitude of interference priming. This was true when the effect was assessed in terms of differences between mean RTs, as well as median RTs. Moreover, the age-invariance remained regardless of whether interference priming was assessed in absolute or relative terms. Finally, accuracy did not vary across age groups or conditions in the present study.

Although the present study replicated Reingold's (1995) basic finding of interference priming, the present sample of younger adults showed lesser interference priming ($\underline{M}$ = 242 ms) than did the younger adults in Reingold's third experiment ($\underline{M}$ = 426 ms).[3.1] Moreover, the present sample of younger adults had a faster mean baseline RT ($\underline{M}$ = 2887 ms) than did Reingold's younger adults ($\underline{M}$ = 3346 ms). Furthermore, accuracy was higher in the present study ($\underline{M}$ = .92) than in Reingold's study ($\underline{M}$ = .88). It is possible that the use of boxes, rather than underlines, to specify the candidate letters in the present experiment may have reduced the task's difficulty level. In other words, the boxed letters may have segregated participants' visual attention in such a way as to increase their accuracy and speed of responding and to reduce the degree of interference observed. Alternatively, the participants in the present study may have been superior to Reingold's participants on some individual-difference variable such as verbal

---

[3.1] Reingold (1995) only tested younger adults in his letter deletion studies.

ability; however, because Reingold did not report such information about his sample, this hypothesis must remain speculative in nature.

The present finding of age invariance in interference priming suggests that the interference condition of Reingold's (1995) task "behaves" more like an indirect facilitation paradigm (i.e., little or no age difference) than like a direct interference paradigm (i.e., a larger age difference favoring younger adults). The present data provide little evidence for explicit contamination or deliberate recollection of previously presented information, which should favor younger over older adults. Moreover, the failure to find age differences in this ostensibly automatic memory test further supports the view that age declines in memory are due to impairments of conscious control rather than impaired inhibition. However, these preliminary data fail to address two important issues. First, it remains to be seen whether the facilitation conditions of Reingold's paradigm show evidence of explicit contamination. Secondly, it is unclear whether age differences in an indirect interference task would emerge when the initial learning is strengthened or the processing demands are increased at the time of retrieval.

Experiment 4

Experiment 4 was partly an attempt to replicate the result of age-equivalence in interference priming in the previous experiment. However, an additional manipulation involved the effects of repetition on age differences in interference priming. More specifically, two researchers (F. I. M. Craik & E. Reingold, personal communications, March 1997) independently suggested that employing multiple presentations of a nonword, prior to a change in the candidate letters for deletion, might produce a larger age difference in interference than would a single prior presentation of the same nonword. The present experiment therefore varied the number of presentations (0 vs. 1 vs. 4) of a nonword prior to a switch in the candidate letters for deletion. However, this manipulation concerned only the critical conditions presented during the "test phase" (the final block of trials, separated from the initial "study phase" by a block of buffer trials that appeared only once during the experiment). The above design, which corresponds to an analysis of interference priming, also yielded facilitation conditions in the sense that participants were expected to become faster in responding to the repeated stimuli during the "study phase." Therefore, the magnitude of facilitation priming could be assessed for repeated items over the course of the four repetitions. In line with suggestions by Craik and Reingold, it was expected that the two age groups should show equal amounts of interference when the nonword appeared once previously, but that the older adults should show more interference when the nonword appeared four times previously. In other words, if older adults have impaired inhibitory mechanisms (Hasher & Zacks, 1988), then this age group should have particular difficulty overcoming prepotent influences that are strengthened through repetition. In contrast, the two age groups should show roughly equivalent facilitation across the four presentations of the study phase; however, if any age differences materialized in the facilitation conditions, then these might be attributed to explicit contamination, or consciously controlled uses of memory on the part of the younger adults.

Methods

Design. The following two-way factorial design was employed for examining test performance: Age (younger vs. older adults) x Condition (Interference1x vs. Interference4x vs. Control). Age was a between-subjects factor, and presentation condition was a within-subjects factor. The Interference1x and Interference4x conditions correspond to nonwords that appeared either one or four times at study and then re-appeared at test, albeit with different letters specified as candidates for deletion. The Control condition corresponds to test cues that were not studied previously. However, this design corresponded only to the critical conditions of the test phase. The above design also yielded facilitation conditions in the sense that participants were expected to become faster across the four presentations of each stimulus in the study phase. The facilitation conditions reflect the following two-way factorial design: Age (young vs. old) x Presentation number of the repeated study items (1 vs. 2 vs. 3 vs. 4). Age was a between-subjects factor, and presentation number was a within-subjects factor.

Participants. Demographic information on the 48 younger and 48 older adults is presented in Table 3.3. None of the participants had previously participated in Experiment 3. Older adults were healthy volunteers, most of whom were tested at a suburban campus of the University of Toronto and were offered $5 for their participation in a 45- to 60-min. session; the remaining older adults were tested at the downtown university campus and were offered reimbursement for their travel expenses. Younger adults were psychology undergraduates who received either course credit or $8 for their participation. On a 20-item version of the Mill Hill Vocabulary Scale (Raven, 1965), the older adults scored higher than the younger adults, $t(94) = 5.25$, $p < .001$. In contrast, on a perceptual speed measure – the WAIS Digit Symbol Substitution Task (Wechsler, 1981) – the younger adults outperformed the older adults, $t(94) = 10.13$, $p < .001$. The finding of an age-related increase in vocabulary, despite a decrease in perceptual speed, was reported previously. The two age groups were equivalent in years of formal education, $t < 1$, and in self-reported health, as assessed over the last two to three months, $t = 1.56$, $p > .12$. However, when health status was assessed for the day of testing, the older adults reported being in better health than did the younger adults, $t(94) = 2.53$, $p < .05$.

Materials. The critical stimuli consisted of 72 test cues: the 48 critical test cues from Experiment 3, plus an additional 24 test cues that were generated by the experimenter. For the additional test cues, pairs of five-letter words with three letters in common were combined to form a six-letter nonword (see also Reingold, 1995). Twelve pairs of words were combined according to one rule described by Reingold (i.e., either the second or fourth letters of the

Table 3.3
Experiment 4: Mean Participant Characteristics

Variable	Age Group				t
	Young		Old		
	M (SD)	n	M (SD)	n	
Age	21.9 (2.7)	48	72.8 (5.5)	48	57.26***
Age range	19-32		61-83		-------
Gender:					-------
Female		38		42	
Male		10		6	
Handedness:					-------
Right-handed		44		45	
Left-handed		3		1	
Ambidextrous		1		2	
Education (years)	15.2 (1.4)		15.5 (2.9)		0.47
Self-reported health:					
Last 2 to 3 months	2.3 (0.9)		2.0 (0.9)		1.56
Today	2.2 (0.8)		1.8 (0.7)		2.53*
Mill Hill Vocabulary Scale	14.0 (2.5)		16.6 (2.3)		5.25***
Digit Symbol Substitution Task	71.5 (14.6)		44.8 (10.8)		10.13***

Note. For gender, $\chi^2_{(1, N=96)} = 1.20$, $p > .20$. For handedness, $\chi^2_{(2, N=96)} = 1.45$, $p > .30$. Self-reported health was assessed for two time frames and was rated on a 5-point scale, ranging from 1 (excellent) to 5 (poor). Mill Hill Vocabulary Scale maximum = 20. Digit symbol score is the number of correct completions in 90 sec (maximum = 93). t values are for younger versus older adults (df = 94); * = $p < .05$, ** = $p < .01$; *** = $p < .001$.

nonword could be deleted to form a five-letter word), whereas an additional twelve pairs of words were combined according to a second rule (i.e., either the third or fifth letters of the nonword could be deleted to form a word). The 24 pairs of additional test cues and their solution words are shown in Appendix 3.1. Forty-three of the 48 additional solution words were listed in the Francis and Kucera (1982) word frequency corpus; the mean frequency of these words was 127 (SD = 376). The buffer stimuli consisted of the 24 buffer items from Experiment 3. Like the critical stimuli, the buffer stimuli were 6-letter nonwords from which either of two letters could be deleted to form a 5-letter word from the remaining letters. However, only one of the two possible solutions was valid in any given trial.

Six experimental formats were generated in order to rotate each critical stimulus through the six conditions produced by crossing study status (0, 1, or 4 presentations at study) with the underline pattern employed (either the second and third letters underlined at test, or the fourth and fifth letters). However, rather than employing underlines, as in Reingold's (1995) study, the present study surrounded the candidate letters with boxes in order to facilitate comparisons with Experiment 3. The study items consisted of 24 nonwords that were presented one time each and another 24 nonwords that were presented four times each with the same letters surrounded by boxes at each stimulus presentation. Each participant was therefore presented with 120 study trials, with 24 trials corresponding to the once-presented items and 96 trials corresponding to the four presentations of the repeated items. The ordering of the study stimuli was random, with the additional constraint that no fewer than three intervening trials separated successive presentations of the same stimulus. As in the previous expeirment, half of the stimuli had the second and third letters surrounded by boxes and the other half had the fourth and fifth letters surrounded by boxes. Moreover, for each surround position, half of the stimuli required a left button response and the remaining half required a right button response. Study trials were preceded and followed by 12 buffer trials, which included an equal number of left and right button responses and the two surround patterns. The test phase consisted of 72 trials: 24 trials in which test cues were studied once previously, 24 trials in which test cues were studied four times previously, and 24 trials with unstudied test cues. The interference trials were the 48 trials in which nonwords were seen previously: For these trials, the previously encountered nonwords reappeared, albeit with different letters surrounded by boxes than in the nonword's previous appearance. The control or baseline trials were the 24 test trials in which the test cues were not seen previously. The version of test cue selected for baseline trials was randomly determined with the same constraints as the study trials. A different random ordering of the test stimuli was employed for each experimental format.

Procedure. In most important respects, the procedure was identical to that employed in Experiment 3. Participants received sample stimuli and instructions that were virtually identical to those used in the previous experiment. Participants then completed the letter deletion trials. The sequence of 12 buffer trials, 120 study trials, 12 additional buffer trials, and 72 test trials was presented as a single block of 216 trials. There was no break between study and test. Participants then completed short tests of vocabulary, perceptual speed, and color- and word-naming (the blocked Stroop task from Experiments 1 and 2) before being debriefed. The experiment took no more than 60 min. per individual session.

Results

Study phase performance. Figure 3.4 shows the mean proportion of correct study trials by age group and presentation number for repeated items. A 2x4 (Age x Presentation number) ANOVA was conducted for the accuracy of items presented a total of four times at study: This analysis revealed as its only reliable effect a main effect of presentation number, $F(3,282) = 2.70$, $p < .05$, $MSE = 0.002$. Post-hoc Tukey HSD tests revealed that this main effect occurred

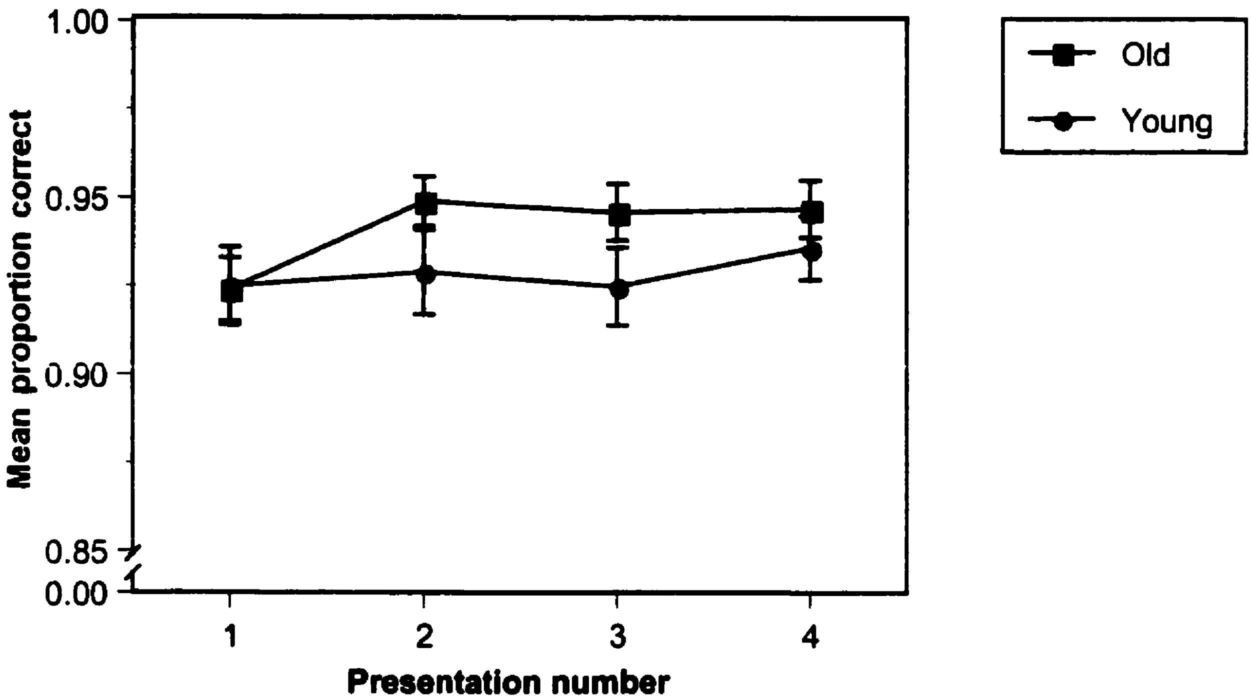


Figure 3.4. Experiment 4: Mean proportion correct responses to repeated study items, by age group and presentation number. Error bars of one standard error are plotted for each point.

because accuracy was reliably greater during an item's fourth presentation than during its first presentation; no other pairwise differences were statistically reliable.

Figure 3.5 shows the mean and median RTs for correct study trials by age group and presentation number. A 2x4 (Age x Presentation number) ANOVA was conducted for the mean correct RTs of items repeated four times at study: This analysis revealed main effects of age, $\underline{F}(1,94) = 85.79$, $\underline{p} < .001$, $\underline{MSE} = 3719447$, and presentation number, $\underline{F}(3,282) = 234.094$, $\underline{p} < .001$, $\underline{MSE} = 129581$, but no interaction of these two variables, $\underline{F} < 1$. The main effect of age indicated that mean RTs were larger for the older adults than for the younger adults. Post-hoc Tukey HSD tests revealed that the main effect of presentation number occurred because mean RTs decreased from the first through fourth presentations of the same stimulus. Moreover, although all pairwise differences were reliable at the $\underline{p} < .05$ level, visual inspection of Figure 3.5 suggests a curvilinear relationship in which the rate of improvement (i.e., reduction of mean RT) decreases with successive stimulus presentations.

With further regard to the mean RT analyses, facilitation priming scores were calculated for each age group by subtracting the mean RT for the fourth presentation from the mean RT for the first presentation. This subsidiary analysis revealed that the older adults showed the same decrease in mean RT from the first to fourth presentations of a stimulus ($\underline{M} = 1260$ ms, $\underline{SD} = 768$) as did the younger adults ($\underline{M} = 1271$ ms, $\underline{SD} = 473$), $\underline{t}(94) < 1$. Pairwise comparisons between successive presentation numbers provided a means of further decomposing this overall priming effect for each age group. Specifically, the older adults showed the same decrease in mean RT from the first to second presentations of a stimulus ($\underline{M} = 700$ ms, $\underline{SD} = 683$) as did the younger adults ($\underline{M} = 824$ ms, $\underline{SD} = 434$), $\underline{t}(94) = 1.06$, $\underline{p} > .29$. Similarly, the older adults showed the same decrease in mean RT from the second to third presentations of a stimulus ($\underline{M} = 354$ ms, $\underline{SD} = 525$) as did the younger adults ($\underline{M} = 327$ ms, $\underline{SD} = 268$), $\underline{t} < 1$. Finally, the older

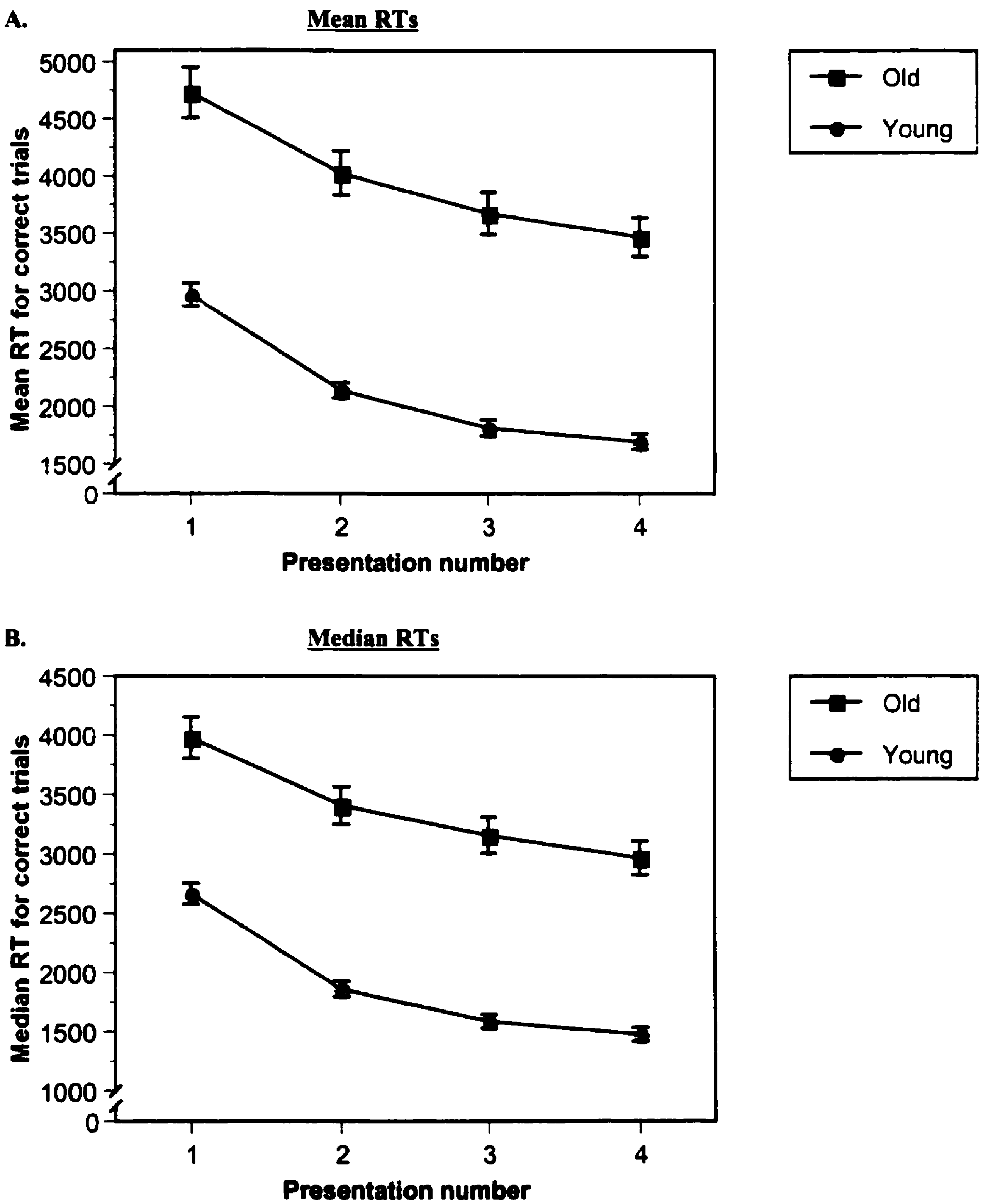


Figure 3.5. Experiment 4: Mean and median correct RTs for repeated study items, by age group and presentation number. Error bars of one standard error are plotted for each point but are only visible when greater in magnitude than the size of the symbol.

adults showed the same decrease from the third to fourth presentations of a stimulus (M = 207 ms, SD = 421) as did their younger counterparts (M = 120 ms, SD = 175), $t(94) = 1.32$, $p > .18$. In summary, there was no reliable age difference in the magnitude of facilitation priming when facilitation was assessed based on the absolute difference between mean RTs in two or more conditions. Moreover, the relation between repetition and performance appeared to be a curvilinear one for both the older and younger adults: Specifically, as the total number of stimulus presentations increased, there was lesser benefit from additional presentations.

However, a different story emerged when the decrease in mean RTs with increasing presentations was assessed in relative, rather than absolute, terms. A more conservative way to examine priming is to divide each participant's absolute priming score, or the average difference between the experimental and baseline conditions, by that participant's average baseline RT. This relative priming score takes into consideration the fact that RT slows quite reliably, and almost universally, with age (e.g., Salthouse, 1996). For the purposes of the present analyses, the condition with the lower presentation number served as a baseline condition and the condition with the higher presentation number served as an experimental condition. At the overall level, the older adults showed a smaller proportion decrease in mean RT from the first to fourth presentations of a stimulus (M = .261, SD = .127) than did the younger adults (M = .425, SD = .114), $t(94) = 6.66$, $p < .001$. This age-related decline in facilitation was replicated in most of the pairwise comparisons that assessed relative decreases in mean RT across successive presentation numbers. Namely, the older adults showed a smaller proportion decrease in mean RT from the first to second presentations of a stimulus (M = .140, SD = .126) than did the younger adults (M = .269, SD = .110), $t(94) = 5.35$, $p < .001$. Similarly, the older adults showed a smaller proportion decrease in mean RT from the second to third presentations of a stimulus (M = .086, SD = .109) than did the younger adults (M = .156, SD = .126), $t(94) = 2.89$, $p < .01$. However, the older adults showed the same proportion decrease in mean RT from the third to fourth presentations of a stimulus ($\underline{M}$ = .051, $\underline{SD}$ = .097) as did the younger adults ($\underline{M}$ = .061, $\underline{SD}$ = .093), $\underline{t} < 1$. In summary, unlike the previous absolute priming analyses that found no age differences, the relative priming analyses for mean RTs revealed a decline in facilitation with age. However, consistent with the absolute priming analyses, the relative priming analyses suggested a curvilinear relation between repetition and performance: That is, with increasing numbers of stimulus presentations, both age groups showed a smaller proportion decrease in mean RT across successive presentations.

A 2x4 (Age x Presentation number) ANOVA was conducted for the median correct RTs of repeated items at study: This analysis revealed main effects of age, $\underline{F}(1,94) = 84.80$, $\underline{p} < .001$, $\underline{MSE} = 2486918$, and presentation number, $\underline{F}(3,282) = 201.32$, $\underline{p} < .001$, $\underline{MSE} = 113158$, and the interaction of these two variables, $\underline{F}(3,282) = 2.98$, $\underline{p} < .05$, $\underline{MSE} = 113158$. The main effect of age indicated that median RTs were larger for the older adults than for the younger adults. Post-hoc Tukey HSD tests revealed that the main effect of presentation number occurred because median RTs decreased from presentation numbers 1 through 4. Moreover, although all pairwise differences were reliable at the $\underline{p} < .05$ level, there were numeric trends toward lesser improvement (i.e., reductions in median RT) with successive stimulus presentations. Tests of simple effects were conducted to assess the nature of the two-way interaction. First, pairwise comparisons assessing age differences for each presentation number revealed reliably longer RTs for older than younger adults in each cell, all $\underline{t}$s > 6.69, $\underline{p}$s < .001. Secondly, one-way ANOVAs assessed the effect of presentation number separately for each age group: These analyses revealed a main effect for the younger adults, $\underline{F}(3,141) = 183.05$, $\underline{p} < .001$, $\underline{MSE} = 75960$, as well as the older adults, $\underline{F}(3,141) = 61.28$, $\underline{p} < .001$, $\underline{MSE} = 150356$. Tukey post-hoc tests assessed the nature of these simple effects: Both the younger adults alone and the older adults alone showed a reliable decrease in median RT from presentation numbers 1 to 3, followed by a non-reliable decrease from presentation numbers 3 to 4. All other pairwise differences were reliable at the $p < .05$ level for each age group. Finally, analyses of priming effects (described below) suggested that the two-way interaction occurred because the two age groups showed equivalent decreases in median RTs across successive stimulus presentations, with one exception: The younger adults showed a larger drop in median RT from the first to second presentations of a stimulus than did the older adults, $t(94) = 2.44$, $p < .05$.

With further regard to the median RT analyses for the study phase, facilitation priming effects were assessed in both absolute and relative terms. First, at the overall level, the older adults showed the same difference between the mean RTs for the first and fourth stimulus presentations ($M = 1009$ ms, $SD = 717$) as did the younger adults ($M = 1192$ ms, $SD = 504$), $t(94) = 1.45$, $p > .15$. This overall priming effect was then further decomposed by pairwise comparisons that examined the difference in median RT between successive presentation numbers. First, in a departure from the age-invariance reported for the overall priming effect, the older adults showed a smaller decrease in median RT from the first to second presentations of a stimulus ($M = 569$ ms, $SD = 495$) than did the younger adults ($M = 807$ ms, SD 459), $t(94) = 2.44$, $p < .05$. However, the older adults showed the same decrease in median RT from the second to third presentations of a stimulus ($M = 250$ ms, $SD = 424$) as did the younger adults ($M = 274$ ms, $SD = 252$), $t < 1$. Similarly, the older adults showed the same decrease in median RT from the third to fourth presentations of a stimulus ($M = 190$ ms, $SD = 419$) as did their younger counterparts ($M = 111$ ms, $SD = 214$), $t(94) = 1.16$, $p > .24$. Therefore, with one exception, there was no age difference in facilitation as measured by the absolute difference between median RTs in two or more conditions. Once again, a curvilinear relation between repetition and performance was suggested by a decelerating trend toward shorter RTs with increasing stimulus presentations.

However, a different result was obtained when the decrease in median RTs across presentation conditions was assessed in relative, rather than absolute, terms. Consistent with the relative priming analyses for mean RTs, the older adults showed a smaller proportion decrease in median RT from the first through fourth presentations of a stimulus ($\underline{M}$ = .246, $\underline{SD}$ = .151) than did the younger adults ($\underline{M}$ = .439, $\underline{SD}$ = .132), $\underline{t}(94) = 6.67$, $\underline{p} < .001$. This finding of a smaller relative priming effect with age was replicated in most of the pairwise comparisons assessing relative decreases in median RT across successive presentation numbers. Namely, the older adults showed a smaller proportion decrease in median RT from the first to second presentations of a stimulus ($\underline{M}$ = .138, $\underline{SD}$ = .114) than did the younger adults ($\underline{M}$ = .292, $\underline{SD}$ = .138), $\underline{t}(94) = 5.95$, $\underline{p} < .001$. Similarly, the older adults showed a smaller proportion decrease in median RT from the second to third presentations of a stimulus ($\underline{M}$ = .071, $\underline{SD}$ = .121) than did the younger adults ($\underline{M}$ = .143, $\underline{SD}$ = .118), $\underline{t}(94) = 2.95$, $\underline{p} < .001$. However, the older adults showed the same proportion decrease from the third to fourth presentations of a stimulus ($\underline{M}$ = .053, $\underline{SD}$ = .118) as did their younger counterparts ($\underline{M}$ = .066, $\underline{SD}$ = .127), $\underline{t} < 1$.

Subsequent analyses examined the extent to which participants showed a generalized practice effect in this experiment -- or improvement over the course of the 120 study trials -- as opposed to a facilitation effect that is based solely on repeated exposure to specific items. The once-presented study items served as a baseline for assessing improvement in the absence of item repetition. More specifically, the 24 once-presented study items in each experimental format were subdivided into quadrants or quartiles, based on their positioning within the study list. For example, the first six once-presented items that occurred in the study list comprised the first list quartile, whereas the last six once-presented items comprised the fourth list quartile.

Figure 3.6 shows the mean proportion correct responses to once-presented study items, by age group and list quartile; whereas Figure 3.7 shows the mean and median RTs for correct study trials as a function of these two variables.

A 2x4 (Age x List quartile) ANOVA for the accuracy of once-presented study items failed to reveal any reliable effects, $\underline{F}s < 1.72$, $\underline{p}s > .16$. However, an Age x List quartile ANOVA for mean correct RTs revealed main effects of age, $\underline{F}(1,94) = 53.20$, $\underline{p} < .001$, $\underline{MSE} = 5431804$, and list quartile, $\underline{F}(3,282) = 3.47$, $\underline{p} < .05$, $\underline{MSE} = 768612$, as well as a marginally reliable two-way interaction, $\underline{F}(3,282) = 2.16$, $\underline{p} < .10$, $\underline{MSE} = 768612$. The main effect of age revealed that the older adults were slower than the younger adults. Post-hoc Tukey HSD tests revealed that the main effect of list quartile occurred because the mean RT was reliably greater for the fourth list quartile than for the third quartile; no other pairwise differences were reliable. Tests of simple effects assessed the nature of the Age x List quartile interaction. First, pairwise comparisons assessing age differences for each list quartile revealed reliably longer RTs for older than younger adults in each cell, all $\underline{t}s > 5.64$, $\underline{p}s < .001$. Secondly, one-way ANOVAs assessed the effect of list quartile separately for each age group: These analyses revealed a main effect for the older adults, $\underline{F}(3,141) = 3.33$, $\underline{p} < .05$, $\underline{MSE} = 1184283$, but not for the younger adults, $\underline{F}(3,141) = 1.09$, $\underline{p} > .35$, $\underline{MSE} = 352940$. Tukey post-hoc tests assessed the nature of the simple effect for older adults: These pairwise comparisons revealed that mean RT was higher in the fourth quartile than in the other three quartiles, $\underline{p}s < .05$; no other pairwise difference were reliable. Finally, an 2x4 (Age x List quartile) ANOVA for median correct RTs revealed a main effect of age in the same direction as before, $\underline{F}(1,94) = 60.84$, $\underline{p} < .001$, $\underline{MSE} = 3493953$, but no effect of list quartile or a two-way interaction, $\underline{F}s < 1.71$, $\underline{p}s > .16$.

In summary, analyses of the once-presented items failed to identify a generalized practice effect or monotonic trend toward improved speed or accuracy across the four quartiles of the

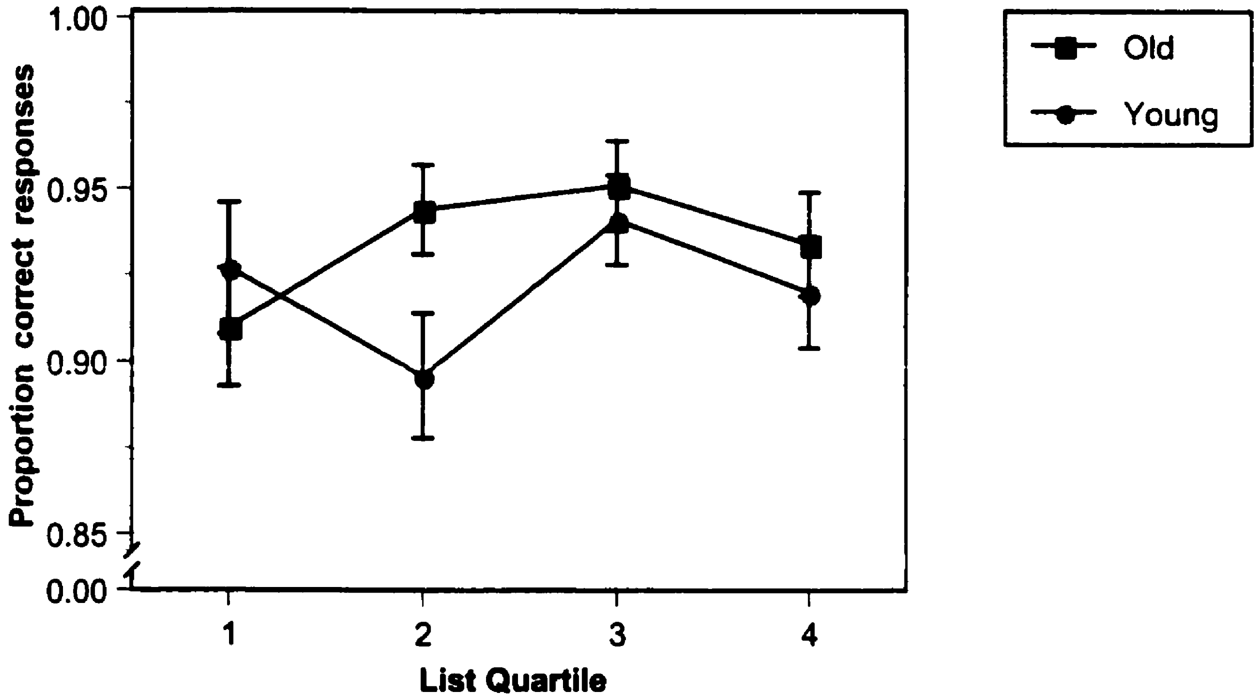


Figure 3.6. Experiment 4: Mean proportion correct responses to once-presented study items, by age group and study list quartile. Error bars of one standard error are plotted for each point.

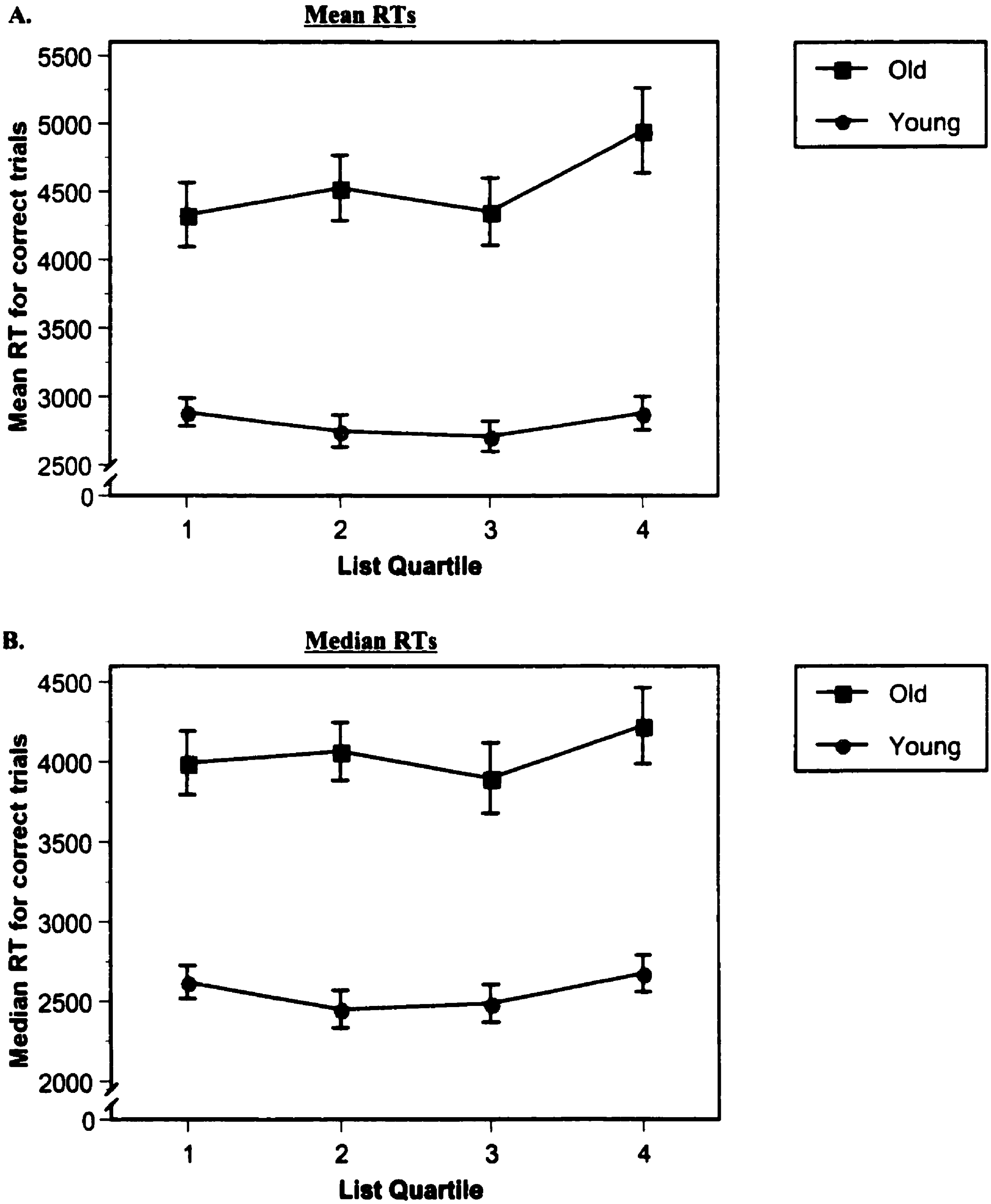


Figure 3.7. Experiment 4: Mean and median correct RTs for once-presented study items, by age group and study list quartile. Error bars of one standard error are plotted for each point.

study list. In fact, there was some evidence for a fatigue effect – in other words, speed and accuracy generally remained invariant across conditions, except for a finding of an increase in RT for older (but not younger) adults in the fourth list quartile.

Test phase performance. Figure 3.8 shows the mean proportion of correct test trials by age group and experimental condition. A 2x3 (Age x Condition) ANOVA for the accuracy of test trials revealed main effects of age, $F(1,94) = 4.18$, $p < .05$, $MSE = 0.014$, and condition, $F(2,188) = 3.70$, $p < .05$, $MSE = 0.002$, but no interaction of these two variables, $F < 1$. The main effect of age revealed that the older adults were more accurate overall in responding to the test trials than were the younger adults. Post-hoc Tukey HSD tests revealed that the main effect of condition occurred because accuracy was higher for items that were presented one time at study (Interference1x condition) than for items presented four times at study (Interference4x condition), $p < .01$; no other pairwise differences were reliable.

Figure 3.9 shows the mean and median RTs for correct test trials by age group and condition. A 2x3 (Age x Condition) ANOVA for mean correct RTs in the test phase revealed main effects of age, $F(1,94) = 65.69$, $p < .001$, $MSE = 3235428$, and condition, $F(2,188) = 31.34$, $p < .001$, $MSE = 188629$, and a reliable two-way interaction, $F(2,188) = 3.04$, $p = .05$, $MSE = 188629$. The main effect of age revealed that mean RTs were larger for the older adults than for the younger adults. Post-hoc Tukey HSD tests revealed that the main effect of condition occurred because mean RTs rose monotonically from the Control condition to the Interference1x condition to the Interference4x condition. Moreover, all pairwise differences were reliable, $ps < .01$. Tests of simple effects were conducted to assess the nature of the two-way interaction. First, pairwise comparisons assessing age differences for each presentation number revealed reliably longer RTs for older than younger adults in each cell, all $ts > 7.21$, $ps < .001$. Secondly, one-way ANOVAs assessed the effect of condition separately for each age group: These

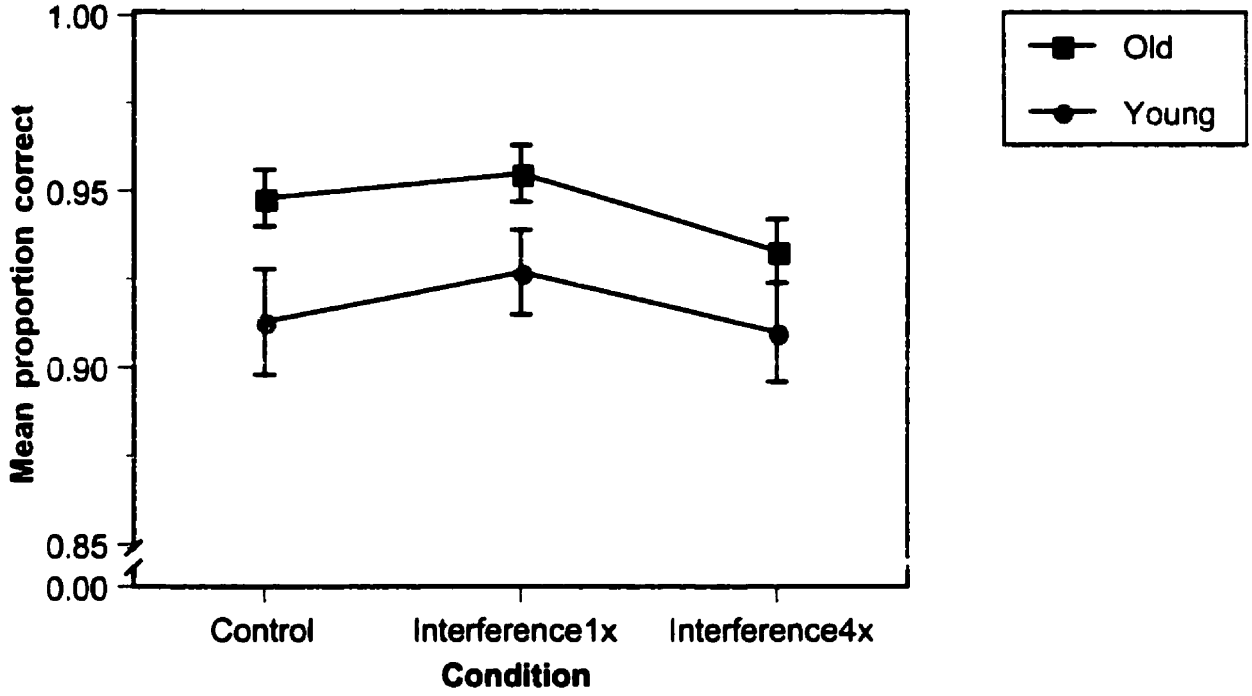


Figure 3.8. Experiment 4: Mean proportion correct test trials, by age group and condition. Error bars of one standard error are plotted for each point.

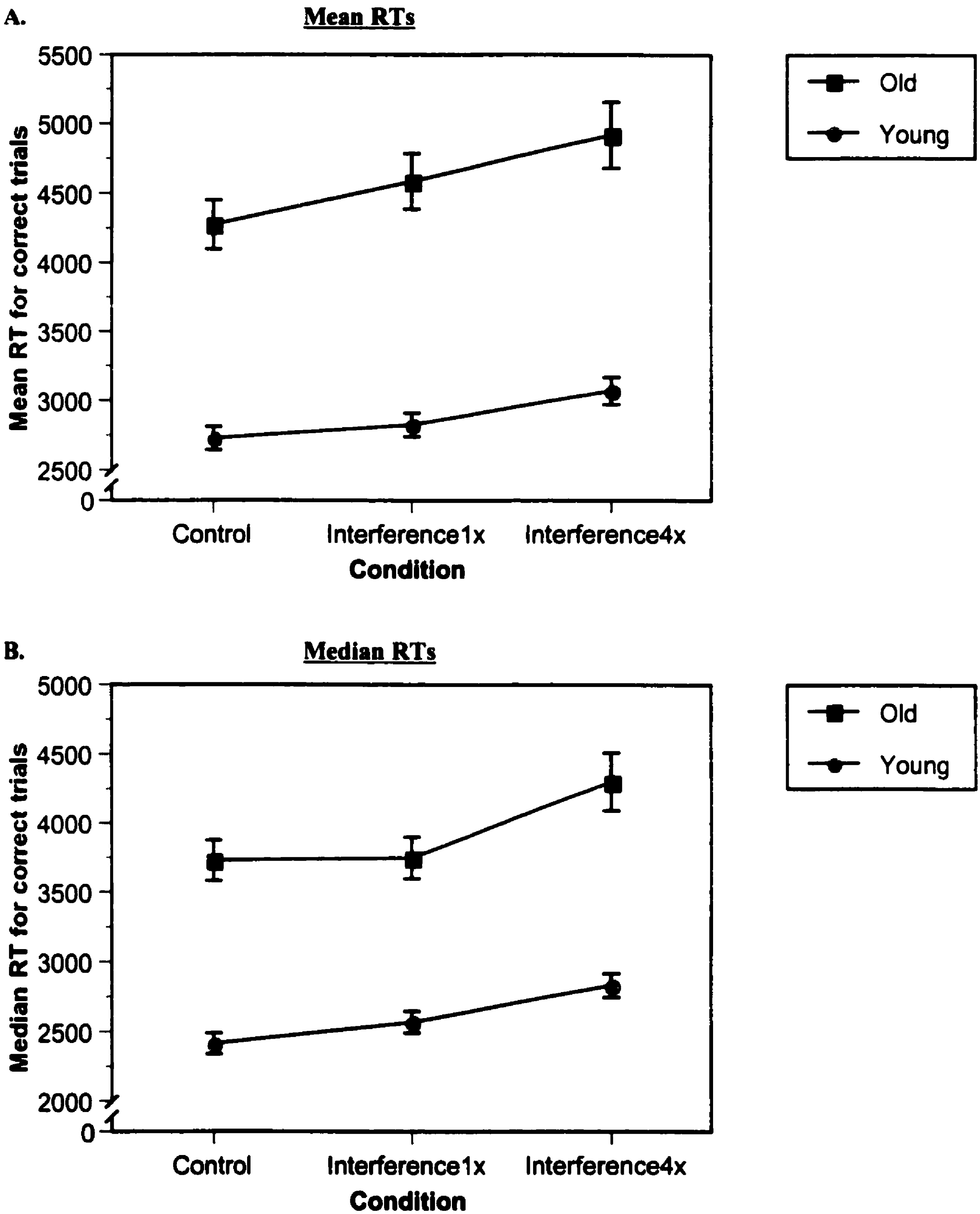


Figure 3.9. Experiment 4: Mean and median RTs for correct test trials, by age group and condition. Error bars of one standard error are plotted for each point.

analyses revealed a main effect for the younger adults, $\underline{F}(2,94) = 22.35$, $\underline{p} < .001$, $\underline{MSE} = 67322$, as well as the older adults, $\underline{F}(2,94) = 16.07$, $\underline{p} < .001$, $\underline{MSE} = 309937$. Tukey post-hoc tests assessed the nature of these simple effects: For the younger adults, the mean RT in the Interference4x condition was greater than that in either the Interference1x or Control conditions, $\underline{p}$s < .01; however, there was no reliable difference between the Interference1x and Control conditions. In contrast, for the older adults, all pairwise differences were reliable: Specifically, the mean RT in the Interference4x condition was greater than that in either the Interference1x condition ($\underline{p} < .05$) or the Control condition ($\underline{p} < .01$); moreover, the mean RT was greater in the Interference1x condition than in the Control condition ($\underline{p} < .05$).

With further regard to the mean RT analyses, interference priming scores were calculated for each age group by subtracting the mean RT in the Control condition from the mean RTs in the Interference1x and Interference4x conditions, respectively. These subsidiary analyses revealed that the increase in mean RT from the Control to Interference1x conditions was marginally greater for the older adults ($\underline{M}$ = 310 ms, $\underline{SD}$ = 685) than for the younger adults ($\underline{M}$ = 98 ms, $\underline{SD}$ = 357), $\underline{p} < .06$. Similarly, the increase in mean RT from the Control to Interference4x conditions was reliably greater for the older adults ($\underline{M}$ = 644 ms, $\underline{SD}$ = 861) than for the younger adults ($\underline{M}$ = 344 ms, $\underline{SD}$ = 386), $\underline{p} < .05$. However, when interference priming was analyzed in proportionate rather than absolute terms, a different result was obtained. (The proportional slowing analyses divide the absolute difference in average RTs between an experimental condition and a baseline condition by the average RT for the baseline condition.) First, the older adults showed the same proportion increase in mean RT from the Control condition to the Interference1x condition ($\underline{M}$ = .078, $\underline{SD}$ = .149) as did the younger adults ($\underline{M}$ = .045, $\underline{SD}$ = .125), $\underline{t}(94) = 1.19$, $\underline{p} > .23$. Secondly, the older adults showed as large an increase in mean RT from the Control to Interference4x conditions ($\underline{M}$ = .154, $\underline{SD}$ = .178) as did their younger counterparts ($\underline{M}$ = .133, $\underline{SD}$ = .138), $\underline{t} < 1$.

A 2x3 (Age x Condition) ANOVA for median correct RTs in the test phase revealed main effects of age, $\underline{F}(1,94) = 56.77$, $\underline{p} < .001$, $\underline{MSE} = 2216240$, and condition, $\underline{F}(2,188) = 39.42$, $\underline{p} < .001$, $\underline{MSE} = 171213$, and a marginally reliable two-way interaction, $\underline{F}(2,188) = 2.90$, $\underline{p} < .06$, $\underline{MSE} = 171213$. The main effect of age revealed that median RTs were larger for the older adults than for the younger adults. Post-hoc Tukey HSD tests revealed that the main effect of condition occurred because median RTs were longer in the Interference4x condition than in either the Interference1x or Control conditions, $\underline{p}s < .01$; however, there was no reliable difference between the Interference1x and Control conditions. Tests of simple effects were conducted to assess the nature of the two-way interaction. First, pairwise comparisons assessing age differences for each presentation number revealed reliably longer RTs for older than younger adults in each cell, all $\underline{t}s > 6.52$, $\underline{p}s < .001$. Secondly, one-way ANOVAs assessed the effect of condition separately for each age group: These analyses revealed a main effect for the younger adults, $\underline{F}(2,94) = 26.51$, $\underline{p} < .001$, $\underline{MSE} = 81792$, as well as the older adults, $\underline{F}(2,94) = 19.48$, $\underline{p} < .001$, $\underline{MSE} = 260635$. Tukey post-hoc tests assessed the nature of these simple effects: For the younger adults, all pairwise differences were reliable. Specifically, the median RT in the Interference4x condition was greater than that in either the Interference1x or Control conditions, $\underline{p}s < .01$; moreover, the median RT was greater in the Interference1x condition than in the Control condition, $\underline{p} < .05$. In contrast, for the older adults, the median RT in the Interference4x condition was greater than that in the other two conditions, $\underline{p}s < .01$, but there was no reliable difference between the Interference1x and Control conditions.

With further regard to the median RT analyses, the magnitude of interference priming was assessed in both absolute and relative terms for the Interference1x and Interference4x conditions. The analyses for absolute interference priming scores revealed that the increase in median RT from the Control to Interference1x conditions was numerically smaller for the older adults (M = 20 ms, SD = 506) than for the younger adults (M = 154 ms, SD = 364); however, this was not a statistically reliable difference, $t(94) = 1.50$, $p > .13$. In contrast, the increase in median RT from the Control to Interference4x conditions was numerically greater for the older adults (M = 573 ms, SD = 767) than for the younger adults (M = 420 ms, SD = 418), but was once again a non-reliable difference, $t(94) = 1.21$, $p > .22$. In summary, the analyses for absolute priming scores failed to yield any reliable age differences. However, when interference priming was analyzed in proportionate rather than absolute terms, a slightly different result was obtained. First, the older adults showed a proportionally smaller increase in median RT from the Control condition to the Interference1x condition (M = .013, SD = .146) than did the younger adults (M = .076, SD = .159), $t(94) = 2.02$, $p < .05$. Secondly, the older adults showed as large an increase in median RT from the Control to Interference4x conditions (M = .154, SD = .208) as did their younger counterparts (M = .190, SD = .188), $t < 1$.

Stroop task performance. Table 3.4 depicts participants' performance on the blocked Stroop task for Experiments 4 and 5, by age group and condition. In the Stroop task, there were two baseline conditions (word-reading and color-naming) and one interference (incongruent color-naming) condition. For the baseline conditions, the older adults were slower than the younger adults, both when reading words in black ink, $t(94) = 3.82$, $p < .001$, and when naming the colors of rectangles, $t(94) = 4.46$, $p < .001$. Both age groups showed perfect accuracy in the reading baseline and were equally accurate in the color-naming baseline, $t(94) < 1$, with near-perfect accuracy in this condition. For the interference condition, the older adults were slower than the younger adults, $t(94) = 10.15$, $p < .001$, but were as accurate as the younger adults, $t(46) = 1.65$, $p > .10$, in naming color words presented in incompatible ink colors. With further regard

Table 3.4

Experiments 4 and 5: Stroop Task Performance

Experiment and Measure	**Age Group**				t
	Young		Old		
	M	(SD)	M	(SD)	
Experiment 4 Stroop Task:					
Word-Reading Baseline					
RT	5.80	(1.11)	6.80	(1.44)	3.82***
Accuracy	1.00	(0.00)	1.00	(0.00)	-----
Color-Naming Baseline					
RT	8.07	(2.44)	10.15	(2.11)	4.46***
Accuracy	0.995	(0.02)	0.99	(0.02)	0.73
Interference *(Incongruent)*					
RT	11.98	(2.71)	20.49	(5.14)	10.15***
Accuracy	0.97	(0.05)	0.95	(0.07)	1.65
Experiment 5 Stroop Task:					
Word-Reading Baseline					
RT	5.55	(0.81)	6.21	(1.34)	2.94**
Accuracy	0.999	(0.01)	0.997	(0.00)	0.58
Color-Naming Baseline					
RT	7.56	(1.21)	8.99	(2.99)	3.06**
Accuracy	0.995	(0.02)	1.00	(0.00)	-----
Interference *(Incongruent)*					
RT	12.50	(2.73)	20.64	(6.31)	8.21***
Accuracy	0.97	(0.04)	0.95	(0.07)	1.86$^+$

Note. Reaction times (RTs) for the Stroop task are the number of seconds taken to either name 16 colors or read 16 words. t values are for younger versus older adults ($df = 94$); $^+ = p < .10$, $* = p < .05$, $** = p < .01$, $*** = p < .001$.

to the Stroop interference effect, the older adults showed a larger increase in mean RT from the color-naming baseline to the incongruent color-naming condition (M = 10.34 sec, SD = 4.55) than did the younger adults (M = 3.91 sec, SD = 2.32), $t(94) = 8.72$, $p < .001$. When this interference effect was assessed in proportionate terms, the older adults showed a greater proportion increase in mean RT across these two conditions (M = 1.05, SD = 0.46) than did their younger counterparts (M = 0.53, SD = 0.35), $t(94) = 6.21$, $p < .001$. These findings for the Stroop task are virtually identical to those reported for Experiment 2.

Correlational analyses. Table 3.5 shows selected correlations among variables for younger and older adults. In these correlational analyses, three priming effects were of principal theoretical interest. The first effect was the magnitude of facilitation priming, operationalized as the difference in mean or median RT between the first and fourth presentations of a repeated stimulus. The second effect was the magnitude of interference priming for previously once-presented items; here, priming was operationalized as the difference in mean or median RT between the Interference1x and Control conditions. The third effect was the magnitude of interference priming for repeated items; here, priming was the difference in RT between the Interference4x and Control conditions. It was of some theoretical interest whether the three priming effects correlated with any of the demographic variables discussed previously, or with a further measure of interference: the Stroop task.

For the younger adults, there were four correlations that were either reliable or marginally reliable. First, age was negatively correlated with two priming effects, as assessed in terms of differences between mean RTs: facilitation priming, $r(46) = -.29$, $p < .05$; and interference priming for repeated items, $r(46) = -.29$, $p < .05$. These correlations suggest that, with greater chronological age, the younger adults showed a lesser magnitude of both facilitation priming and interference priming. Speculatively, the older members of this age group may have processed

Table 3.5
Experiment 4: Selected Correlations for Younger and Older Adults

Younger Adults

Memory measure		***Demographic and Neuropsychological Measures***			
	Age	Education	Vocabulary	Digit Symbol	Stroop Interference
Facilitation					
Mean RT difference	**-.29***	-.21	-.12	.04	-.04
Median RT difference	-.16	-.14	-.09	-.04	.01
Interference (1x)					
Mean RT difference	.00	.04	.00	.17	-.17
Median RT difference	-.13	-.11	.10	.03	.06
Interference (4x)					
Mean RT difference	**-.29***	-.07	-.17	.16	**-.27˙**
Median RT difference	-.17	-.02	-.07	.12	**-.32***

Older Adults

Memory measure		***Demographic and Neuropsychological Measures***			
	Age	Education	Vocabulary	Digit Symbol	Stroop Interference
Facilitation					
Mean RT difference	-.12	.03	**-.54*****	-.11	-.12
Median RT difference	-.10	-.12	**-.40****	-.15	-.13
Interference (1x)					
Mean RT difference	.14	.23	-.08	-.03	.08
Median RT difference	.00	.09	.19	-.08	-.11
Interference (4x)					
Mean RT difference	.10	.08	-.11	-.18	-.03
Median RT difference	.20	.07	-.12	-.19	.14

Note. The facilitation effect is the decrease in mean or median RT from the first to the fourth presentations of a stimulus. The interference effect is the increase in mean or median RT from the Control condition to either the Interference1x condition or the Interference4x condition. Stroop interference is the difference in mean RT between the incongruent (interference) condition and color-naming baseline of the blocked Stroop test. $N = 48$ for each age group; ˙ = $p < .10$, * = $p < .05$, ** = $p < .01$, *** = $p < .001$.

stimuli more globally than their younger peers and therefore may have shown less of an effect (either benefit or cost) from instantiation of a particular solution at each stimulus presentation. However, the correlation between age and interference priming for repeated items was unduly influenced by two outliers: Two younger adults in their thirties showed facilitation, rather than interference, in this condition. When these two individuals were omitted from the analysis, the correlation was no longer reliable, $r(44) = .03$, $p > .84$. Secondly, there was a negative correlation between Stroop interference and interference priming for repeated items, regardless of whether this interference effect was assessed in terms of differences between mean RTs, $r(46) = -.27$, $p < .07$, or median RTs, $r(46) = -.32$, $p < .05$. These correlations, surprisingly, suggested that the two interference measures (Stroop and letter deletion) bore an inverse relation to one another. However, there was one clearly visible outlier: One younger adult showed no interference on the Stroop task but showed substantial interference in the repeated-items condition of the letter deletion task. When this individual was omitted from consideration, the latter two correlations ceased to be reliable, $rs > -.24$, $ps > .11$.

For the older adults, there were two reliable correlations: Vocabulary was negatively correlated with facilitation priming, regardless of whether this effect was operationalized as the difference between mean RTs, $r(46) = -.54$, $p < .001$; or as the difference between median RTs, $r(46) = -.40$, $p < .001$. These correlations suggest that the highly verbal older adults may have performed better at one stimulus presentation than did their less verbal peers. Consequently, the highly verbal elderly may have had less room for improvement before they reached a functional ceiling.

Discussion

The present study yielded a more complex pattern of results than that obtained in Experiment 3. Namely, the type of conclusions drawn depends partly on the specific analyses employed. First, in the study phase, accuracy differed only slightly across presentation conditions for repeated items: Although accuracy levels were higher for the fourth presentation of a stimulus than for the first stimulus presentation, no other pairwise comparisons among conditions yielded reliable differences. Moreover, accuracy did not differ as a function of age group, nor was there an interaction of age and presentation condition. In contrast, an analysis of RT for the study phase revealed a decline in RT across each of the four presentation conditions; this was true, regardless of whether the absolute differences in RTs were analyzed for means or medians. However, there was some evidence for a curvilinear function for both age groups: Analyses of facilitation priming revealed that, although there was always a reliable benefit (i.e., a decrease in RT) across successive presentations of the same stimulus, the magnitude of that benefit decreased with increasing numbers of presentations. Further analyses assessed whether the facilitation effect had its locus on item-specific processing or general improvement over time. When the trials for once-presented trials were divided into four quartiles, based on trial sequence, there was no monotonic decrease in RT across the four quartiles. This result suggests that the facilitation effect reflected sustained activation of a representation for a particular item, rather than a general practice effect. In fact, there was evidence of a fatigue effect, or increase in RT in the fourth quartile of the once-presented items, for the older adults.

With regard to age differences in the study conditions, the older adults were slower in overall mean and median RT than were the younger adults, a finding consistent with Experiment 3 and with a large literature suggesting general slowing with age (see, e.g., Salthouse, 1996). However, when facilitation priming was assessed in terms of absolute RT differences for either means or medians, there was no age difference. On the other hand, when facilitation priming was assessed in proportionate terms, or relative to the baseline RT, the older adults generally showed a smaller facilitation effect than did their younger counterparts. The finding of age-equivalence in the absolute RT analyses of the facilitation conditions is consistent with the view that indirect facilitation paradigms largely reflect automatic processes, which are age-invariant. However, to the extent that there are age differences favoring the young in the proportionate analyses, this finding may be taken as some evidence for intrusions of deliberate recollection on the part of younger, but not older, adults (see, e.g., Light and La Voie, 1993).

The results for the test phase were similarly complex. First, accuracy was greater for previously once-presented items (the Interference1x condition) than for previously repeated items (the Interference4x condition) and for older than for younger adults; however, there was no interaction of age and condition. There was therefore some evidence for a decline in accuracy with repetition of previously learned, but no longer appropriate, information. Moreover, older adults appeared to have adopted a slightly more conservative response criterion than did the younger adults. Secondly, in the RT analyses for the test phase, the older adults had longer overall mean and median RTs than did the younger adults. However, in the analyses of absolute differences in mean RTs, the older adults showed more interference than did the younger adults. This was true, regardless of whether items were presented one or four times prior to a change in candidate letters for deletion. On the other hand, when the interference effect was analyzed in terms of absolute differences in median RTs, or in terms of relative changes in mean or median RTs, there was no evidence for an age difference favoring the younger adults. In fact, the median RT analyses provided some evidence for a small decrease in interference with age, at least for items that were presented once prior to a change in candidate letters.

Although the above results appear somewhat complicated at first, they do yield a coherent picture of age differences in memory. Namely, the analyses of absolute differences in mean RTs suggest that interference priming increases with age, whereas the analyses of median RTs and of relative differences in RTs suggest that interference priming is age-invariant. The former analyses are highly susceptible to the influences of outlying observations, or very long RTs in the right tail of the RT distribution, whereas the latter analyses are less susceptible to these outliers. Therefore, it would appear that any age-related increase in interference is due to an increase in the number of very long RTs for the older adults, but not for the younger adults. However, when the influence of these extremely long RTs is minimized through the use of more robust statistical methods (e.g., medians and trimmed means), the age difference is eliminated (see, e.g., Ratcliff, 1993, for a review of these statistical methods).[3.2]

A similar result was recently reported by Faust (1998), who found different patterns of age differences in a negative priming study, depending on whether negative priming was operationalized in terms of differences between mean RTs or median RTs. Specifically, when Faust included attended repetition trials (use of the same target across successive displays) in his negative priming procedure, he found an age-related increase in negative priming (i.e., greater priming for older than younger adults) when mean RTs were considered. However, he found age-equivalence in negative priming when median RTs were considered. Faust suggested that episodic retrieval might be playing some role in his paradigm over and above inhibition (see also, e.g., Neill & Valdes, 1992, for a similar explanation of negative priming). Similarly, this mechanism might also enter into the present paradigm. Speculatively, the inclusion of a repetition condition may make participants more "test-aware" and therefore prompt them to deliberately retrieve previously presented solutions, at least in the study phase. Faust suggested that the processes underlying negative priming may affect the shape of the RT distributions, rather than simply shifting them. Although the present interference paradigm is in many respects

[3.2] In fact, when a trimming procedure was performed on the present data, the results of the median analyses were largely replicated. This procedure, which eliminated observations more than 2.5 standard deviations from the mean for each age group, found equivalent interference priming effects for the two age groups. These analyses of trimmed means are omitted in order to simplify the description of the results.

the converse of the negative priming paradigm (i.e., the target on one trial becomes the distractor on the next trial, rather than the other way around), there is nonetheless some parallelism between the two sets of results which suggests that similar mechanisms might be at play. This is particularly true, given the inclusion in both studies of a facilitation condition. Faust suggests that episodic retrieval might be playing some role in his paradigm over and above inhibition (see also, e.g., Neill & Valdes, 1992, for a similar explanation of negative priming). Similarly, this mechanism might also enter into the present paradigm. Speculatively, the inclusion of a repetition condition may make participants more "test-aware" and therefore prompt them to deliberately retrieve previously presented solutions, at least in the study phase. Moreover, it is possible that this strategy is initially adaptive and that the younger adults are superior at this intentional retrieval of prior solutions. Furthermore, it is possible that, in the interference trials, the younger adults learn that this strategy is no longer effective, whereas the older adults are less aware of the change in experimental conditions. However, this explanation does not seem a very parsimonious one. On a further note, Faust's conclusions suggest that an ex-Gaussian type of analysis might be appropriate to examine age differences in the shapes of the RT distributions for the present study. However, it is somewhat unclear whether such an analysis could yield meaningful results in the present study, given the relatively few observations available per individual participant (personal communication, E. Reingold, April 1998).

One surprising aspect of the present results is that a particular condition (i.e., items that appeared once prior to test) which appeared in both Experiments 3 and 4 yielded an age difference in the mean RT analyses of the present study, but not in the previous study. One possible explanation for this discrepant finding is that the present study employed more trials leading up to the test phase than had the previous study (144 vs. 48 trials, respectively). It is possible that the additional trials produced a fatigue effect in older, but not younger, adults that increased the older age group's susceptibility to interference. In support of this theory, the older adults, but not the younger adults, showed an increase in RT in the fourth quartile of the once-presented study items. The possibility that fatigue and old age interacted to impair participants' inhibitory functioning is nicely mirrored in Circadian arousal research which has shown that age declines in inhibition are especially pronounced at non-optimal times of day (see, e.g., Yoon, May, & Hasher, 1998, for a review of this research).

A final point regarding the present results is that the age-related increase in interference does not appear to reflect a speed-accuracy trade-off (SATO). Although it is true that, overall, the older adults performed more accurately and more slowly than the younger adults, this apparent SATO does not fully account for all aspects of the data. Namely, the older and younger adults showed the same magnitude decrease in accuracy from the Interference1x condition to the Interference4x condition, and yet the older adults showed a greater *increase* in mean RT across these two conditions than did the younger adults. In other words, the classification of this result as a speed-accuracy correspondence misses the mark because the absolute costs in mean RT across these two conditions are larger for older than younger adults, despite equivalent costs in accuracy for the two age groups.

Experiment 5

Experiment 5 was a further attempt to assess the boundary conditions under which age-equivalence in interference might be obtained. Specifically, Denise Park (personal communication, January 1998) suggested that imposing a working memory load within the letter deletion task might be one means of producing reliable age differences in interference. This suggestion was in line with her previous argument that environment support might be most successful (i.e., in eliminating age differences in performance) when it reduces the demands on working memory (Park & Shaw, 1992). Conversely, one might predict age differences to be greater under high, as opposed to low, working memory loads. Therefore, Experiment 5 examined age differences in interference priming under two conditions: 1) the standard, "boxed" version of the letter deletion task employed previously, in which the candidate letters for deletion are surrounded by boxes; and 2) a working-memory (WM) version of the task, in which the two candidate letters for deletion are presented to the participant before each trial, rather than surrounded by boxes in the trial itself. In other words, the WM trials started off with the experimenter saying the two candidate letters aloud (e.g., "n,g") and then requesting the participant to repeat these two letters aloud before initiating the presentation of the nonword (without boxes) on the computer screen. As in Experiment 3, interference priming was assessed for each task by subtracting the RT in the Control condition (no prior presentations of a stimulus) from that in the Interference condition (one stimulus presentation, before a change in candidate letters for deletion). In line with Park's suggestion, a three-way interaction (Age x Task x Condition) interaction was expected in which the magnitude of interference priming was equivalent for the two age groups in the boxed or standard version of the task but was greater for the older adults in the WM version.

Methods

Design. The following four-way factorial design was employed for examining test performance, with each factor having two levels: Age (younger vs. older adults) x Task Order (the Boxed task vs. the WM task as the first task presented) x Task (WM vs. Boxed) x Condition (Interference vs. Control). Age and task order were between-subjects factors, whereas task and condition were within-subjects factors. Task order refers to whether the standard, boxed version of the letter deletion task was the first or the second of the two tasks presented to a participant; this variable was a "nuisance variable" and, as such, was not of theoretical significance. Task refers to the use of either the boxed or standard version of the letter deletion task (used in

Experiments 3 and 4), or the working-memory (WM) version in which the candidate letters for deletion are maintained in WM in the absence of boxes. Finally, condition refers to whether a test item was previously studied, albeit with a change in which letters were candidates for deletion (Interference condition), or was previously unstudied (Control condition).

Participants. Demographic information on the 48 younger and 48 older adults is presented in Table 3.6. None of the participants had previously participated in Experiments 3 or 4. Older adults were healthy volunteers, most of whom were tested at a suburban campus of the University of Toronto and were offered $5 for their participation in a 45- to 60-min. session; the remaining older adults were tested at the downtown university campus and were offered reimbursement for their travel expenses. Younger adults were psychology undergraduates who received either course credit or $8 for their participation. On a 20-item version of the Mill Hill Vocabulary Scale (Raven, 1965), the older adults scored higher than the younger adults, $t(94) = 6.54$, $p < .001$. In contrast, on a perceptual speed measure – the WAIS Digit Symbol Substitution Task (Wechsler, 1981) – the younger adults outperformed the older adults, $t(94) = 10.93$, $p < .001$. The finding of an age-related increase in vocabulary, despite a decline in perceptual speed, was reported in each of the four previous experiments. The two age-groups were equivalent in years of formal education and in self-reported health, as assessed for two time frames: the last two to three months, and today, ts < 1.

Materials. The critical stimuli consisted of 96 test cues: the 72 critical test cues from Experiment 4, plus an additional 24 test cues that were newly generated by the experimenter. For the additional test cues, pairs of five-letter words with three letters in common were combined to form a six-letter nonword (see also Reingold, 1995). Twelve new pairs of words were combined according to one rule described by Reingold (i.e., either the second or fourth letters of the nonword could be deleted to form a word), whereas an additional twelve pairs of

Table 3.6
Experiment 5: Mean Participant Characteristics

Variable	Age Group				t
	Young		Old		
	M (SD)	n	M (SD)	n	
Age	22.4 (3.9)	48	70.4 (5.5)	48	49.32***
Age range	19-35		60-81		-------
Gender:					-------
Female		42		27	
Male		6		21	
Handedness:					-------
Right-handed		43		38	
Left-handed		5		6	
Ambidextrous		0		4	
Education (years)	14.9 (1.8)		14.8 (3.1)		0.14
Self-reported health:					
Last 2 to 3 months	2.0 (0.9)		2.1 (0.8)		0.59
Today	2.0 (1.0)		1.9 (0.8)		0.40
Mill Hill Vocabulary Scale	13.2 (2.0)		15.9 (2.1)		6.54***
Digit Symbol Substitution Task	71.3 (11.7)		46.3 (10.7)		10.93***

Note. For gender, $\chi^2_{(1, N=96)} = 11.59$, $p < .001$. For handedness, $\chi^2_{(2, N=96)} = 4.40$, $p > .10$. Self-reported health was assessed for two time frames and was rated on a 5-point scale, ranging from 1 (excellent) to 5 (poor). Mill Hill Vocabulary Scale maximum = 20. Digit symbol score is the number of correct completions in 90 sec (maximum = 93). t values are for younger versus older adults ($df = 94$); * = $p < .05$, ** = $p < .01$; *** = $p < .001$.

words were combined according to a second rule (i.e., either the third or fifth letters of the nonword could be deleted to form a word). The 24 pairs of newly generated test cues and their solution words are shown in Appendix 3.1. As was the case in Experiment 4, 43 of the 48 new solution words were listed in the Francis and Kucera (1982) word frequency corpus; the mean frequency of these newly generated words was 94 (SD = 263). The buffer stimuli consisted of 72 six-letter nonwords, including 21 of the buffer items from Experiment 4, plus an additional 51 buffer items that were generated by the experimenter. Thirty-six of the 72 buffer stimuli were 6-letter nonwords from which either of two letters could be deleted to form a 5-letter word. However, only one of the two possible solutions was valid in any given trial. The remaining 36 buffer stimuli were 6-letter nonwords from which only one letter could be deleted to form a 5-letter word.

Eight experimental formats were generated in order to rotate each critical stimulus through the eight conditions produced by crossing task (WM vs. Boxed) with study status (studied vs. unstudied) and the surround pattern employed (either the second and third letters serving as candidate letters, or the fourth and fifth letters). Each task corresponded to a different block of trials, for a total of two blocks per participant. In the block corresponding to the Boxed task condition, the two candidate letters in each trial were surrounded by boxes. However, in the block corresponding to the WM task condition, the nonword in each trial appeared without boxes. In each block there were 24 study trials in which a study stimulus appeared. As in the previous experiment, for half of the study stimuli the second and third letters served as candidates for deletion, and for the remaining stimuli the fourth and fifth letters served as candidates for deletion. Moreover, for each surround position, half of the stimuli required a left button response and the remaining half required a right button response. Study trials were preceded and followed by 18 buffer trials, which included an equal number of left and right button responses and the two surround patterns. The test phase consisted of 48 trials, of which 24 corresponded to the Interference condition and 24 to the Control condition. In the Interference trials of the test phase, previously encountered nonwords reappeared, albeit with different letters serving as candidates for deletion than in the nonword's previous appearance. For example, in the Interference+WM condition, if a nonword's second and third letters were auditorily presented in the study phase, then the fourth and fifth letters of that nonword would be auditorily presented in the test phase. The Control or baseline trials represented the 24 test cues that were not used in the study trials. The version of test cue selected for baseline trials was randomly determined with the same constraints as the study trials. A different random ordering of the study and test stimuli was employed for each experimental format.

Procedure. Participants received two blocks of trials, with one block corresponding to the Boxed or standard version of the letter deletion task and one block corresponding to the WM version of the task. Half of the participants received the Boxed task first and the remaining half received the WM task first. For the Boxed task, the procedure was largely identical to that employed in Experiment 3; namely, participants indicated which of two boxed letters (either the left or right one) could be deleted from a 6-letter nonword in order to form a 5-letter word. For the WM task, no boxes appeared; instead, the two candidate letters for deletion were presented auditorily before each trial and were to be retained in memory. Then participants saw a nonword on the computer screen and had to indicate which of the two previously presented letters could be deleted to form a word. Specifically, participants indicated whether they wished to delete the first or the second of the previously announced letters by pressing either the "1" or "2" key of the keyboard, respectively. As in the previous experiments, participants were instructed to use the index and middle fingers of their dominant hand to select between the two alternative keys on each trial. Moreover, participants were to press the "ENTER" key of the keyboard with their non-dominant hand to initiate each trial. Furthermore, participants were instructed to respond as quickly as they could without making errors. The experimenter's instructions to participants are shown in Appendix 3.3.

Before starting each block of trials, participants received sample stimuli and instructions that were quite similar to those used in the previous two experiments. Then participants completed the letter deletion trials. Each block consisted of 108 trials in the following sequence: 18 buffer trials, 24 study trials, 18 additional buffer trials, and 48 test trials. There was no break between the study and test trials of a given block.

The trial sequence for the Boxed task condition was identical to that employed for Experiments 3 and 4. The trial sequence for the WM task condition was as follows: A fixation cross was presented in the center of the screen. The experimenter then read two letters aloud (e.g., "s,o") from a sheet of paper. The experimenter spoke in a loud, low voice and clearly enunciated the two letters presented for each trial. The participant was to repeat the two letters aloud and commit these two letters to memory for the upcoming trial. Any errors in repetition were immediately corrected by the experimenter; however, such errors rarely occurred in the experiment. Participants then initiated the trial (i.e., the presentation of the nonword and the recording of RTs) by pressing the "ENTER" key on the computer keyboard. Approximately 200 ms after the "ENTER" key was pressed, a nonword replaced the fixation cross in the center of the computer screen. The nonword was displayed until a response was made. The screen was then blanked for 500 ms followed by the reappearance of the fixation cross.

After completing the two blocks of letter deletion trials, participants completed short tests of vocabulary, perceptual speed, and color- and word-naming (the blocked Stroop task from Experiment 4). Participants then completed a survey assessing frequency of use of word puzzles and word games in daily life before being debriefed. This survey assessed whether participants used a number of word puzzles or games, including crossword puzzles, word jumbles, word searches, the board game *Scrabble*, and the television game show *Wheel of Fortune*. In addition, participants who indicated that they used a particular game or puzzle were asked about their frequency of usage (e.g., daily, weekly, or monthly). The experiment took no more than 60 min. per individual session.

Results

Test phase performance. Figure 3.10 shows the proportion of correct test trials as a function of age group and experimental condition. A 2x2x2x2 (Age x Task Order x Task x Condition) ANOVA for accuracy revealed main effects of age, $F(1,92) = 4.19$, $p < .05$, $MSE = 0.026$, and task, $F(1,92) = 8.55$, $p < .01$, $MSE = 0.007$. These main effects indicated that accuracy was higher for older adults, as compared to younger adults, and for the Boxed task, as compared to the WM task. In addition, there was a marginally reliable Age x Order interaction, $F(1,92) = 3.17$, $p < .08$, $MSE = 0.026$. The two-way interaction occurred because there was an age difference in accuracy favoring older adults when the WM task appeared first, $t(46) = 2.97$, $p < .01$, but no age difference in accuracy when the Boxed task appeared first, $t < 1$. There were also two reliable three-way interactions: Age x Order x Task, $F(1,92) = 5.43$, $p < .05$, $MSE = 0.007$, and Age x Order x Condition, $F(1,92) = 4.73$, $p < .05$, $MSE = 0.003$. Secondary analyses assessed the nature of each three-way interaction: Specifically, separate two-way ANOVAs were conducted for each level of the task ordering variable (i.e., Boxed task first, and WM task first). First, the Age x Order x Task interaction occurred because there was a reliable Age x Task interaction when the WM task appeared first, $F(1,46) = 9.68$, $p < .01$, $MSE = 0.002$, but not when the Boxed task appeared first, $F < 1$. More specifically, when the WM task appeared first, there was an age difference in accuracy favoring older adults in the WM task, $t(46) = 3.50$, $p < .01$, but not in the Boxed task, $t(46) = 1.63$, $p = .11$. Secondly, the Age x Order x Condition

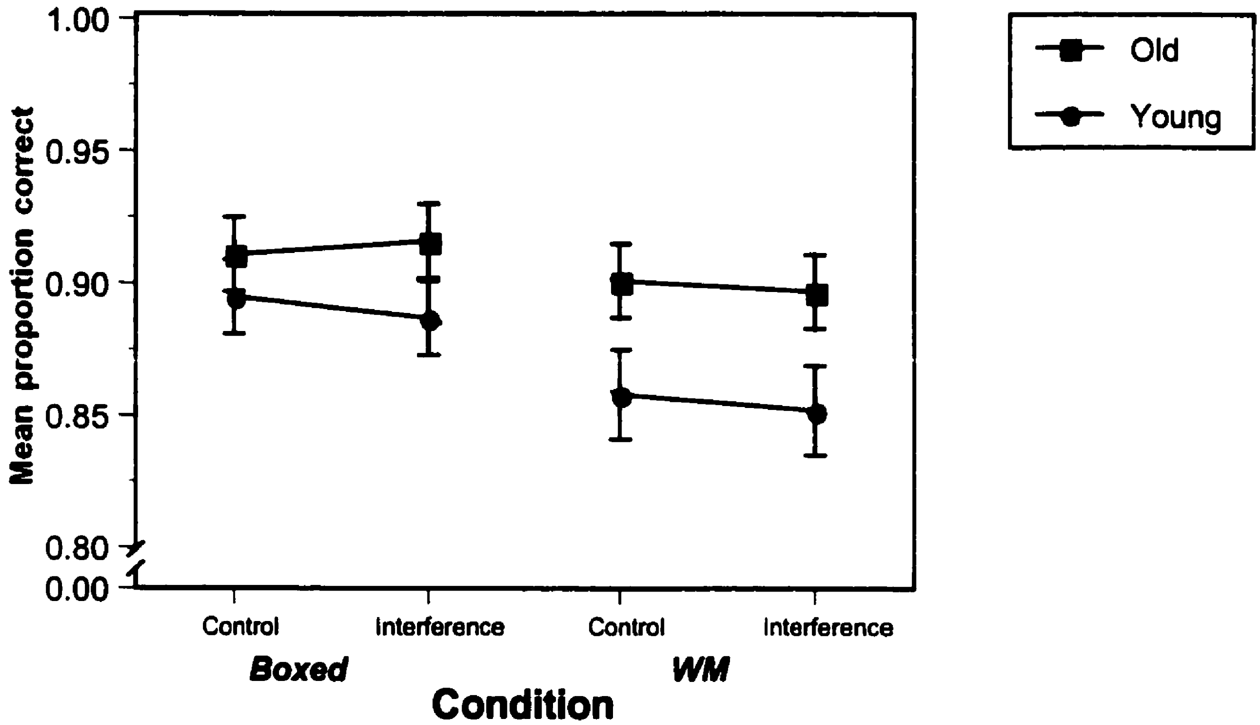


Figure 3.10. Experiment 5: Mean proportion correct test trials, by age group and condition. Error bars of one standard error are plotted for each point.

interaction occurred because there was a marginally reliable Age x Condition interaction when the WM task appeared first, $F(1,46) = 3.55$, $p < .07$, $MSE = 0.002$, but not when the Boxed task appeared first, $F(1,46) = 1.34$, $p > .25$, $MSE = 0.001$. More specifically, when the WM task appeared first, there was a reliable age difference in accuracy favoring older adults in both presentation conditions, ts > 2.11, ps < .05; however, this age difference was numerically greater in the Interference condition (M difference = .08) than in the Control condition (M difference = .05).

Figure 3.11 shows the mean and median RTs for correct test trials by age group and condition. A 2x2x2x2 (Age x Task Order x Task x Condition) ANOVA for mean correct RTs revealed main effects of age, $F(1,92) = 63.85$, $p < .001$, $MSE = 6484228$, task order, $F(1,92) = 4.31$, $p < .05$, $MSE = 6484228$, task, $F(1,92) = 26.56$, $p < .001$, $MSE = 708255$, and condition, $F(1,92) = 16.37$, $p < .001$, $MSE = 258985$. These main effects indicated that mean correct RTs were higher when participants were older adults, as compared to younger adults, and when the Boxed task appeared first, as compared to when the WM task appeared first. Moreover, mean RTs were higher in the WM task than in the Boxed task, and were higher in the Interference condition than in the Control condition. There were also three reliable interactions: Age x Task, $F(1,92) = 26.56$, $p < .01$, $MSE = 708255$, Age x Order x Task, $F(1,92) = 7.44$, $p < .01$, $MSE = 708255$, and Age x Condition, $F(1,92) = 4.95$, $p < .05$, $MSE = 258985$. Tests of simple effects assessed the nature of the interactions. First, the Age x Task interaction occurred because the increase in mean RT from the Boxed to WM tasks was reliable for the older adults, $t(47) = 4.69$, $p < .001$, but was only marginally reliable for the younger adults, $t(47) = 1.86$, $p < .07$. Secondly, the Age x Order x Task interaction occurred because the Age x Task interaction was reliable when the Boxed task appeared first, $F(1,46) = 13.99$, $p < .01$, $MSE = 461113$, but not when the WM task appeared first, $F < 1$. More specifically, when the Boxed task appeared first,

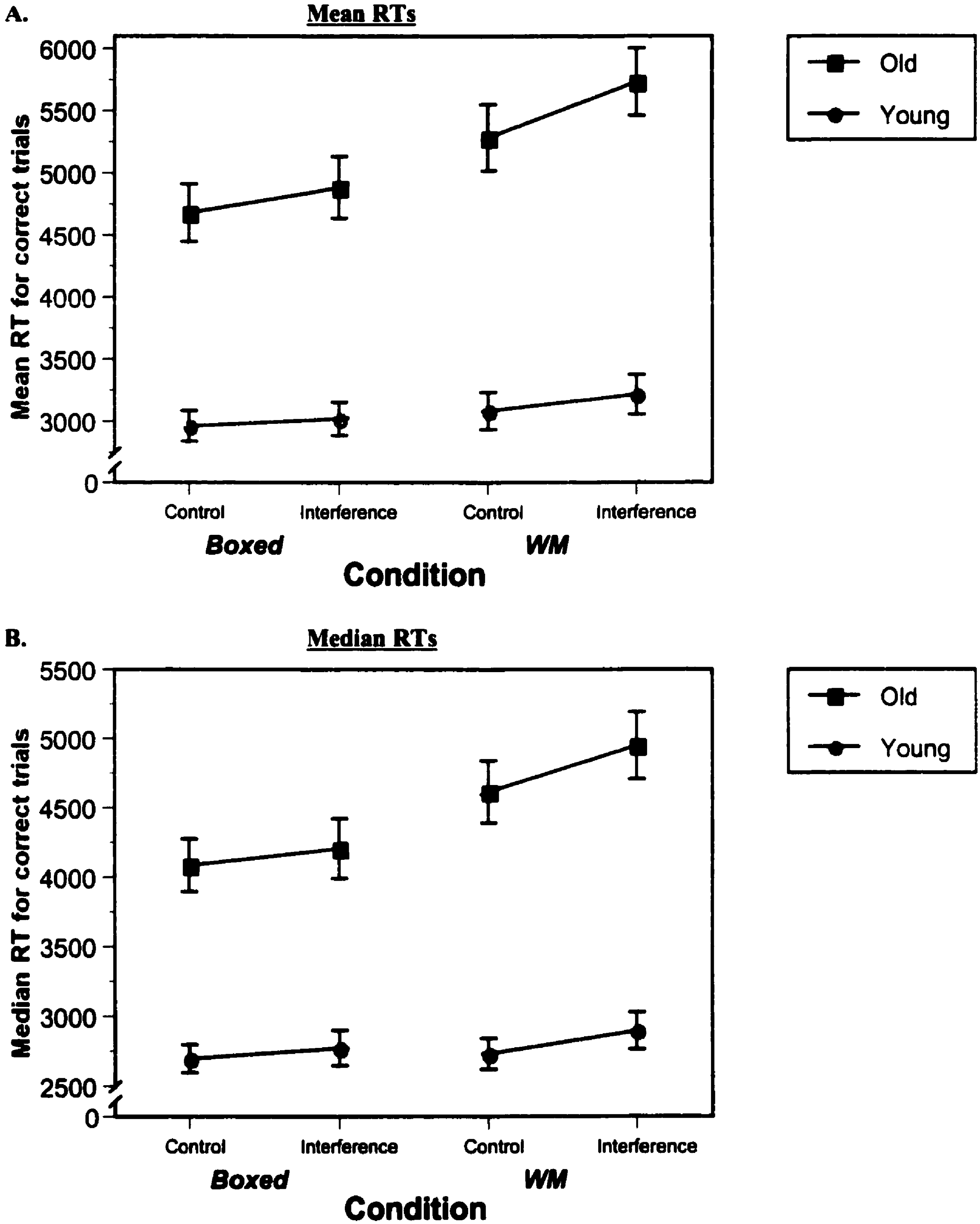


Figure 3.11. Experiment 5: Mean and median RTs for correct test trials, by age group and condition. Error bars of one standard error are plotted for each point.

the older adults showed a reliable increase in mean RT from the Boxed task to the WM task, $\underline{t}(23) = 3.66$, $\underline{p} < .01$, whereas the younger adults showed equivalent mean RTs across the two tasks, $\underline{t} < 1$. Thirdly, and most importantly for the hypotheses of interest, the Age x Condition interaction occurred because, although both age groups were reliably faster in the Control condition than in the Interference condition ($\underline{t}$s > 2.21, $\underline{p}$s < .05), the difference between conditions was numerically greater for older than younger adults.

With further regard to the mean RT analyses, interference priming scores were calculated for each age group by subtracting the mean RT in each Control condition from the mean RT in the corresponding Interference condition. First, overall absolute priming scores were calculated, collapsing across task type. This analysis revealed that, at the overall level, the increase in mean RT from the Control to Interference conditions was reliably greater for the older adults ($\underline{M} = 326$ ms, $\underline{SD} = 652$) than for the younger adults ($\underline{M} = 95$ ms, $\underline{SD} = 296$), $\underline{t}(94) = 2.24$, $\underline{p} < .05$. Secondly, absolute priming scores were calculated separately for each task type, due to the theoretical significance of these task-specific scores. In the Boxed task, the increase in mean RT from the Control to Interference conditions was numerically greater for the older adults ($\underline{M} = 206$ ms, $\underline{SD} = 1035$) than for the younger adults ($\underline{M} = 59$ ms, $\underline{SD} = 438$), but this age difference was not statistically reliable, $\underline{t}(94) < 1$. However, in the WM task, the older adults showed a marginally greater increase in mean RT from the Control to Interference conditions ($\underline{M} = 445$ ms, $\underline{SD} = 1021$) than did the younger adults ($\underline{M} = 130$ ms, $\underline{SD} = 475$), $\underline{t}(94) = 1.94$, $\underline{p} < .06$. Although these subsidiary analyses suggest a three-way interaction (Age x Task x Condition) in which the greatest interference is shown by older adults in the WM task, this interaction failed to attain statistical significance in the four-way ANOVA reported above, $\underline{F}(1,92) = 0.44$, $\underline{p} > .50$, $\underline{MSE} = 382822$. This failure to obtain a reliable three-way interaction may be due in part to the high degree of variability shown by participants.

When interference priming was analyzed in proportionate rather than absolute terms, a slightly different result was obtained. (The proportional slowing analyses divide the absolute difference in average RTs between an experimental condition and a baseline condition by the average RT for the baseline condition.) First, at the overall level, the older adults showed a larger proportion increase in mean RT from the Control condition to the Interference condition (M = .075, SD = .118) than did the younger adults (M = .028, SD = .098), $t(94) = 2.09$, $p < .05$. However, at the level of the specific tasks (Boxed and WM), this proportionate age difference was only partly corroborated. In the Boxed task, the older adults showed a numerically larger increase in mean RT from the Control to Interference conditions (M = .059, SD = .191) than did the younger adults (M = .021, SD = .150), but this age difference was not statistically reliable, $t(94) = 1.07$, $p > .28$. Similarly, in the WM task, the older adults (M = .105, SD = .185) showed a numerically greater relative priming score than did the younger adults (M = .050, SD = .160), but this age difference was not reliable, $t(94) = 1.57$, $p > .12$. These subsidiary analyses for proportionate changes in mean RT therefore largely confirm the Age x Condition interaction in the four-way ANOVA reported above – that is, the older adults show more interference priming than the younger adults, in both relative and absolute terms. Moreover, there is once again no evidence for the hypothesized three-way interaction in these analyses.

A 2x2x2x2 (Age x Task Order x Task x Condition) ANOVA was conducted for median RTs for correct test trials. This analysis revealed main effects of age, $F(1,92) = 62.90$, $p < .001$, MSE = 4362093, task order, $F(1,92) = 5.88$, $p < .05$, MSE = 4362093, task, $F(1,92) = 18.43$, $p < .001$, MSE = 668110, and condition, $F(1,92) = 13.07$, $p < .001$, MSE = 222071. These main effects indicated that median correct RTs were higher when participants were older adults, as compared to younger adults, and when the Boxed task appeared first, as compared to when the WM task appeared first. Moreover, median RTs were higher in the WM task than in the Boxed task, and were higher in the Interference condition than in the Control condition. There were also two reliable interactions: Age x Task, $F(1,92) = 11.19$, $p < .01$, $MSE = 668110$, and Age x Order x Task, $F(1,92) = 8.46$, $p < .01$, $MSE = 668110$. Tests of simple effects assessed the nature of these two interactions. First, the Age x Task interaction occurred because the increase in median RT from the Boxed to WM tasks was reliable for the older adults, $t(47) = 4.02$, $p < .001$, but not for the younger adults, $t(47) = 1.14$, $p > .26$. Secondly, the Age x Order x Task interaction occurred because the Age x Task interaction was reliable when the Boxed task appeared first, $F(1,46) = 13.37$, $p < .01$, $MSE = 488642$, but not when the WM task appeared first, $F < 1$. More specifically, when the Boxed task appeared first, the older adults showed a reliable increase in median RT from the Boxed task to the WM task, $t(23) = 3.58$, $p < .01$, whereas the younger adults performed similarly across the two tasks, $t < 1$.

With further regard to the median RT analyses, interference priming scores were calculated for each age group by subtracting the median RT in each Control condition from the median RT in the corresponding Interference condition. First, overall absolute priming scores were calculated, collapsing across task type. This analysis revealed that, at the overall level, the increase in median RT from the Control to Interference conditions was numerically greater for the older adults ($M = 227$ms, $SD = 524$) than for the younger adults ($M = 121$ ms, $SD = 409$), but this age difference was not statistically reliable, $t(94) = 1.11$, $p > .27$. Secondly, absolute priming scores were calculated separately for each task type, which confirmed this basic result. Namely, in the Boxed task, the increase in median RT from the Control to Interference conditions was numerically greater for the older adults ($M = 120$ ms, $SD = 959$) than for the younger adults ($M = 76$ ms, $SD = 573$), but this age difference was not reliable, $t(94) < 1$. Similarly, in the WM task, the increase in median RT from the Control to Interference conditions was numerically greater for the older adults ($M = 334$ ms, $SD = 859$) than for the younger adults

(M = 166 ms, SD = 547), but was not reliably so, $t(94) = 1.14$, $p > .25$. These subsidiary analyses corroborate the lack of reliable interactions for median RTs, both for the Age x Condition interaction and for the Age x Task x Condition interaction.

When the magnitude of change in median RTs was assessed in relative, rather than absolute terms, the same age-invariance in interference priming was observed. First, at the overall level, the older adults showed the same proportion increase in median RT from the Control to Interference conditions (M = .039, SD = .151) as did the younger adults (M = .059, SD = .117), $t(94) < 1$. Secondly, at the level of specific tasks, both age groups showed the same relative priming effect. In the Boxed task, the older adults showed the same proportion increase in median RT (M = .031, SD = .216) as did the younger adults (M = .042, SD = .205), $t < 1$. Similarly, in the WM task, the older adults showed the same proportion increase (M = .064, SD = .187) as did the younger adults (M = .088, SD = .181), $t < 1$.

Performance in only the first-presented task. In the previous four-way ANOVAs for accuracy, mean RTs, and median RTs, there were several main effects and interactions involving task order. The presence of order effects makes interpretation of the results somewhat less straightforward. Therefore, additional analyses examined only the first task completed by each participant. These analyses were not susceptible to the complicating effects of task order because only one task type – either the WM task, or the Boxed task – was examined for each participant. The following analyses therefore conform to a three-way (Age x Task x Condition) factorial design, with age and task as between-subjects factors and with condition as a within-subjects factor.

Figure 3.12 shows the proportion of correct test trials in only the first-presented task by age group and experimental condition. A 2x2x2 (Age x Task x Condition) ANOVA for accuracy in the first-presented task revealed a main effect of age, $F(1,92) = 7.11$, $p < .01$, MSE = 0.020,

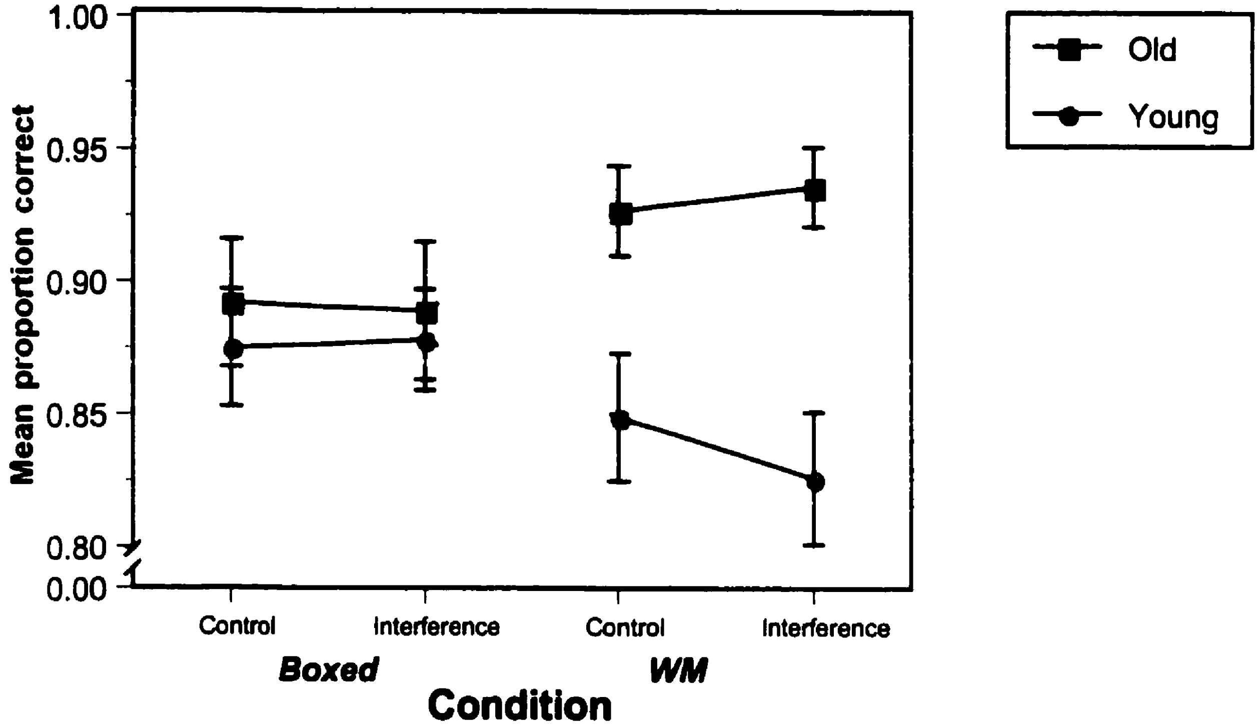


Figure 3.12. Experiment 5: Mean proportion correct test trials *in the first-presented task only*, by age group and condition. Error bars of one standard error are plotted for each point.

and a marginally reliable Age x Task interaction, $F(1,92) = 3.92$, $p < .06$, $MSE = 0.020$. The main effect of age occurred because accuracy was higher for older than for younger adults. The Age x Task interaction occurred because the age-related increase in accuracy was reliable in the WM task, $t(46) = 3.50$, $p < .01$, but was not reliable in the Boxed task, $t(46) < 1$.

Figure 3.13 shows the mean and median RTs for correct test trials in only the first-presented task by age group and condition. A 2x2x2 (Age x Task x Condition) ANOVA for mean correct RTs in the first-presented task revealed a main effect of age, $F(1,92) = 48.80$, $p < .001$, $MSE = 3338462$, and a marginally reliable effect of condition, $F(1,92) = 3.68$, $p < .06$, $MSE = 362715$. These main effects occurred because the mean RT was higher for older adults, as compared to younger adults, and for the Interference condition, as compared to the Control condition. These main effects are consistent with those previously reported in the four-way ANOVA for mean RTs. However, in a noticeable departure from the previous results for mean RTs, the Age x Condition interaction failed to attain statistical significance, $F(1,92) = 2.30$, $p > .13$, $MSE = 362715$. The non-significance of this interaction in the present analysis may be due to a lack of statistical power. In other words, there were fewer observations per cell in the present three-way ANOVA than in the previous four-way ANOVA, due to the removal of data for the second task completed by each participant.

With further regard to the mean RT analyses for the first-presented task, interference priming scores were calculated for each age group. First, at the overall level, the older adults showed a numerically larger increase in mean RT from the Control to Interference conditions (M = 298 ms, SD = 1119) than did the younger adults (M = 35 ms, SD = 420); however, this age difference was not statistically reliable, $t(94) = 1.53$, $p > .13$. Secondly, at the level of the two specific tasks, this basic result was corroborated: In the Boxed task, the older adults showed a numerically larger increase in mean RT from the Control to Interference conditions (M = 211

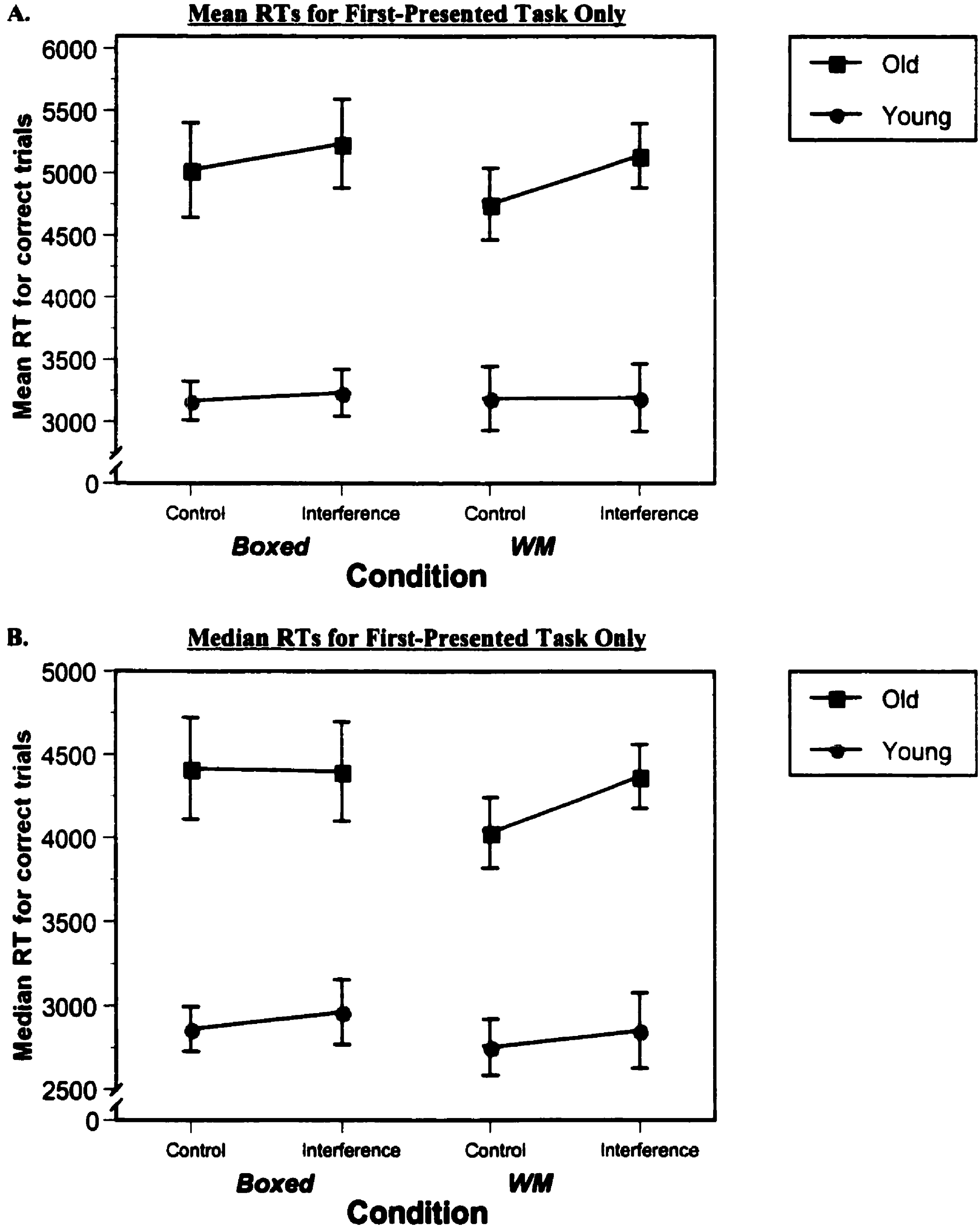


Figure 3.13. Experiment 5: Mean and median RTs for correct test trials *in the first-presented task only*, by age group and condition. Error bars of one standard error are plotted for each point.

ms, SD = 1097) than did the younger adults (M = 64 ms, SD = 520), but this age difference was not reliable, $t(46) < 1$. Similarly, in the WM task, the older adults showed a numerically larger increase in mean RT across conditions (M = 386 ms, SD = 1157) than did the younger adults (M = 6 ms, SD = 298), but this was not a reliable age difference, $t(46) = 1.56$, $p > .12$.

A slightly different result was obtained when the change in mean RTs for the first-presented task was assessed in relative, rather than absolute, terms. First, at the overall level, the older adults showed a larger relative increase in mean RT from the Control to Interference conditions (M = .091, SD = .209) than did the younger adults (M = .011, SD = .137), $t(94) = 2.22$, $p < .05$. However, at the level of the specific tasks, this age difference was only partly replicated. Namely, in the Boxed task, the older adults showed a numerically larger proportion increase in mean RT from the Control to Interference conditions (M = .071, SD = .205) than did the younger adults (M = .018, SD = .172), but this age difference was not statistically reliable, $t(46) < 1$. However, in the WM task, the older adults showed a reliably larger increase in mean RT across conditions (M = .111, SD = .217) than did the younger adults (M = .003, SD = .094), $t(46) = 2.23$, $p < .05$.

A 2x2x2 (Age x Task x Condition) ANOVA for median correct RTs in the first-presented task revealed a main effect of age, $F(1,92) = 48.44$, $p < .001$, MSE = 2077247, in the direction of higher RTs for older than younger adults. No other effects attained statistical significance, although the effect of condition approached marginal significance, $F(1,92) = 2.61$, $p = .11$, MSE = 2077247.

With further regard to the median RT analyses for the first-presented task, interference priming scores were calculated in absolute, and then relative, terms. First, at the overall level, the older adults showed a numerically larger increase in median RT (M = 161 ms, SD = 948) from the Control to Interference conditions than did the younger adults (M = 100 ms, SD = 601), but this age difference was not statistically reliable, $\underline{t}(94) < 1$. However, at the level of the specific tasks, this result was only partly confirmed. In the Boxed task, the older adults actually showed a slight decrease in median RT from the Control to Interference conditions ($\underline{M}$ = -17 ms, $\underline{SD}$ = 974), whereas the younger adults showed an increase in RT across these two conditions ($\underline{M}$ = 101 ms, $\underline{SD}$ = 644); however, there was no reliable age difference in this priming effect, $\underline{t}(46) < 1$. In contrast, in the WM task, the older adults showed a numerically larger increase in median RT from the Control to Interference conditions ($\underline{M}$ = 339 ms, $\underline{SD}$ = 905) than did the younger adults ($\underline{M}$ = 99 ms, $\underline{SD}$ = 569), but again this age difference was non-reliable, $\underline{t}(46) = 1.10$, $\underline{p} > .27$.

The above interference priming results for the first-presented task were largely confirmed when the change in median RT was assessed in relative, rather than absolute, terms. First, at the overall level, the older adults showed a numerically larger proportion increase in median RT from the Control to Interference conditions ($\underline{M}$ = .067, $\underline{SD}$ = .215) than did the younger adults ($\underline{M}$ = .029, $\underline{SD}$ = .204), but this age difference was not statistically reliable. In the Boxed task, the older adults showed the same proportion increase in median RT from the Control to Interference conditions ($\underline{M}$ = .021, $\underline{SD}$ = .210) as did the younger adults ($\underline{M}$ = .033, $\underline{SD}$ = .234), $\underline{t}(46) < 1$.[3.3] In contrast, in the WM task, the older adults showed a numerically greater relative increase in median RT across conditions ($\underline{M}$ = .112, $\underline{SD}$ = .215) than did the younger adults ($\underline{M}$ = .025, $\underline{SD}$ = .175), but this age difference was not reliable, $\underline{t}(46) = 1.53$, $\underline{p} > .13$.

[3.3] The secondary priming analyses for median RTs in the first-presented task yielded a seemingly contradictory result. Namely, in the absolute interference priming analyses, the older adults actually showed a slight decrease in median RT from the Control to Interference conditions. However, in the relative priming analyses, the older adults showed a slight increase in median RT across these two conditions. This difference might be explained by the fact both priming scores in question approached zero and by the fact that inter-individual variability was quite high. Therefore, it is likely that the two methods of averaging individual priming scores yielded numerically different results, which nonetheless did not differ significantly from zero.

Stroop task performance. In the Stroop task, there were two baseline conditions (word-reading and color-naming) and one interference (incongruent color-naming) condition. For the baseline conditions, the older adults were slower than the younger adults, both when reading words in black ink, $t(94) = 2.94$, $p < .01$, and when naming the colors of rectangles, $t(94) = 3.06$, $p < .01$. For the word-reading baseline, the two age groups were equally accurate, $t < 1$, with near-perfect accuracy in this condition. For the color-naming baseline, the older adults showed perfect accuracy whereas the younger adults showed near-perfect accuracy with a total of four errors committed by this age group. For the interference condition, the older adults were slower than the younger adults, $t(94) = 8.21$, $p < .001$, and were marginally less accurate than the younger adults, $t(46) = 1.86$, $p < .07$, in naming color words presented in incompatible ink colors. With further regard to the Stroop interference effect, the older adults showed a larger increase in mean RT from the color-naming baseline to the incongruent color-naming condition ($M = 11.65$ sec, $SD = 5.35$) than did the younger adults ($M = 4.93$ sec, $SD = 2.41$), $t(94) = 7.94$, $p < .001$. When this interference effect was assessed in proportionate terms, the older adults showed a greater proportion increase in mean RT across these two conditions ($M = 1.36$, $SD = 0.67$) than did their younger counterparts ($M = 0.67$, $SD = 0.34$), $t(94) = 6.43$, $p < .001$. These findings for the Stroop task largely corroborate those reported in the previous experiments.

Frequency of use of word games and word puzzles. Table 3.7 depicts participants' self-reported use of word puzzles and word games in daily life. Two types of analyses were conducted on participants' responses to a word game and word puzzle survey. First, the number of participants reporting use of each word game and word puzzle was tabulated for each age group. Chi-square tests then assessed whether there were group differences in the numbers of participants reporting use, versus those reporting non-use, of each game or puzzle. These analyses revealed no age differences in the numbers of younger versus older adults using each

Table 3.7
Experiment 5: Participants' Use of Word Puzzles and Word Games in Daily Life

Type of Word Puzzle or Game	Age Group: Young			Old			χ^2	t	
	Number using N	Frequency of use M	(SD)	Number using N	Frequency of use M	(SD)		Value	df
Wheel of Fortune	32	3.3	(0.9)	28	3.6	(1.0)	0.71	1.23	58
Crossword Puzzles	31	3.2	(0.8)	26	3.8	(1.0)	1.08	2.62*	55
Word Search	27	2.6	(0.7)	17	2.8	(1.1)	4.20*	1.01	42
Scrabble	25	2.4	(0.8)	25	2.4	(0.6)	0.00	0.00	48
Word Jumble	19	3.0	(0.9)	18	3.4	(1.2)	0.04	1.26	35
Other	11	3.0	(0.8)	5	2.6	(0.9)	2.95^{+}	0.92	14
Overall Frequency of Use:	45	2.9	(0.6)	45	3.2	(0.7)	0.00	2.33*	88
		Number used M	(SD)		Number used M	(SD)			
Mean Number of Puzzles and Games Used:	-----	3.0	(1.6)	-----	2.5	(1.5)	-----	1.69^{+}	94

Note. Number using each puzzle or game is out of 48 participants for each age group. Participants' responses to the frequency-of-use question were categorized into one of five frequency categories: 1=less than yearly; 2=at least yearly; 3=at least monthly; 4=at least weekly; 5=at least daily. Overall (mean) frequency of use was assessed for participants reporting use of at least one puzzle or game. Mean number of puzzles and games used was assessed for all 96 participants. Chi-square tests assess age differences in the numbers of participants reporting use, versus non-use, of each puzzle or game (df = 1; N = 96). t tests assess age differences in the frequency of game and puzzle use, with varying degrees of freedom; $^{+}$ = p < .10, * = p < .05, ** = p < .01; *** = p < .001.

game or puzzle, with two exceptions. First, there were more younger adults reporting use of word-search puzzles than there were older adults, $\chi^2_{(1, N=96)} = 4.20$, $p < .05$. In addition, there were marginally more younger adults reporting use of "other" games and puzzles – i.e., those not listed on the questionnaire – than there were older adults, $\chi^2_{(1, N=96)} = 2.95$, $p < .10$; however, the total number of such individuals was small for each age group (less than one-quarter of each group). These "other" games and puzzles included *Boggle* ($N = 7$), cryptograms ($N = 2$), and the games *Hangman*, *Pictionary*, and *Scattergories* ($N = 1$ each). A further analysis revealed that equivalently high numbers of older and younger adults ($N = 45$ in each age group, or 94% of the total) reported using at least one type of word game or word puzzle in daily life, $\chi^2_{(1, N=96)} = 0.00$, $p = 1.00$.

A second set of analyses assessed age differences in participants' responses to a frequency of use question – i.e., how often they used each word game or word puzzle in daily life. This question was asked only of participants after they indicated that they used a particular word game or word puzzle. Participants' responses to the frequency-of-use question were subsequently categorized by the experimenter into one of five frequency categories: 1 (less than yearly), 2 (at least yearly), 3 (at least monthly), 4 (at least weekly), or 5 (at least daily). T-tests then assessed whether there were age differences in participants' frequency of use of each game and puzzle, with the degrees of freedom corresponding to the total number of individuals using a particular game or puzzle, minus two. These analyses revealed no age differences in frequency of game and puzzle use, except for crossword puzzles, which the older adults used more often than did the younger adults, $t(55) = 2.62$, $p < .05$. A further analysis assessed overall (mean) frequency of use for participants using at least one puzzle or game: This analysis revealed that, at the overall level, older adults used word games and word puzzles more often than did the younger adults, $t(88) = 2.33$, $p < .05$. However, this age-related increase in the overall frequency of game and puzzle use may be largely attributable to the older adults' more frequent use of crossword puzzles. A final analysis assessed the mean number of puzzles and games used for all 96 participants: This analysis revealed that the younger adults used more games and puzzles, on average, than did the older adults, $\underline{t}(94) = 1.69$, $\underline{p} < .10$.

In summary, the analyses for self-reported word game and word puzzle use revealed relatively few age differences. First, the vast majority (94%) of older and younger adults reported using at least one type of word game or word puzzle in daily life. Secondly, the numbers of younger and older adults reporting use of each game or puzzle were equivalent in most cases, except for word searches, which were used by more younger adults than older adults. Thirdly, the two age groups used most individual word games and puzzles with equal frequency, except for crossword puzzles, which were used more regularly by the older adults than by the younger adults. Fourthly, at the overall level, the older adults used puzzles and games with somewhat greater regularity than did the younger adults. Finally, despite the younger adults' less frequent game and puzzle use, this age group used a greater number of puzzles and games than did the older adults.

Correlational analyses. Table 3.8 shows selected correlations among variables for younger and older adults. The memory measures of interest were the interference priming effects in the Boxed condition, as well as the WM condition. For each task condition, interference priming was operationalized as the difference in mean or median RT between the Interference and Control conditions for that task. The demographic and neuropsychological measures were the same ones employed in Experiment 4; however, in addition, there were two measures of word puzzle and word game usage in daily life. These were the total number of word games and puzzles used by each participant and, for those participants reporting use of at least one word game or puzzle, the mean frequency of word game and puzzle usage. The latter

Table 3.8
Experiment 5: Selected Correlations for Younger and Older Adults

Younger Adults

Memory measure	Age	Education	Vocabulary	Digit Symbol	Stroop Interference	Total Puzzles/ Games Used	Frequency Puzzle/Game Use
			Demographic and Neuropsychological Measures				
Interference (Boxed)							
Mean RT difference	-.04	-.02	-.05	-.09	-.05	.02	.13
Median RT difference	.04	.11	-.06	-.10	.11	-.09	-.04
Interference+WM							
Mean RT difference	.09	.00	.03	-.22	**.36***	**-.25^{+}**	-.01
Median RT difference	.21	.09	.11	-.24	.18	**-.36***	-.01

Older Adults

Memory measure	Age	Education	Vocabulary	Digit Symbol	Stroop Interference	Total Puzzles/ Games Used	Frequency Puzzle/Game Use
			Demographic and Neuropsychological Measures				
Interference (Boxed)							
Mean RT difference	-.13	**-.27^{+}**	**-.26^{+}**	-.07	**.30***	**.40****	.15
Median RT difference	-.20	-.09	**-.30***	.05	.17	**.36***	.01
Interference+WM							
Mean RT difference	-.22	**-.30***	-.15	.02	.10	-.02	-.09
Median RT difference	.05	-.24	-.08	-.05	.17	-.10	-.08

Note. The interference effect is the increase in (mean or median) RT from the Control condition to the Interference condition within each task (Boxed or WM). Stroop interference is the difference in mean RT between the incongruent (interference) condition and color-naming baseline of the blocked Stroop test. "Total puzzles/games used" refers to the mean number of word games and word puzzles that each participant reported using in daily life. "Frequency puzzle/game use" refers to participants' self-reported mean frequency of playing word games or word puzzles, ranging from 1 (less than yearly) to 5 (at least daily). N = 48 for each age group, except for puzzle/game frequency, where N = 45 for each age group; $^{+}$ = $p < .10$, * = $p < .05$, ** = $p < .01$, *** = $p < .001$.

measure was subsequently scored on a five-point frequency scale, ranging from 1 (less than yearly) to 5 (at least daily). It was thought that greater use of word puzzles and word games in daily life would moderate the magnitude of interference priming observed in both the younger and older adults.

For the younger adults, there were three reliable correlations. First, the number of puzzles and games used was negatively correlated with the interference priming effect in the WM task: This was true regardless of whether interference was operationalized as the difference between mean RTs, $r(46) = -.25$, $p < .05$; or as the difference between median RTs, $r(46) = -.36$, $p < .05$. This result suggests that the more avid younger game and puzzle players processed the nonwords more globally (i.e., considered more solutions on each trial) than did the non-players and that this global processing strategy reduced the players' susceptibility to interference. However, this conclusion should be interpreted cautiously, given the existence of an outlier for the correlation involving the mean RT difference. Namely, one younger adult who played one word game showed substantial interference. When this individual was removed from the analysis, the correlation involving the mean RT difference was no longer reliable, $r(45) = -.15$, $p = .32$. Finally, the degree of Stroop interference was positively correlated with the degree of interference priming on the WM task, $r(46) = .36$, $p < .05$, but only when interference was assessed as the difference between mean RTs. This result suggests that the younger adults who had difficulty overcoming prepotent responses (e.g., word reading) in the Stroop task had similar difficulties in the WM condition of the letter deletion task.

For the older adults, there were seven correlations that were either reliable or marginally reliable. First, education was negatively correlated with interference priming in the Boxed task, $r(46) = -.27$, $p < .10$, and in the WM task, $r(46) = -.30$, $p < .05$, when these priming effects were operationalized as the mean RT difference between the Interference and Control conditions.

These correlations suggest that the highly educated older adults showed lesser interference than did their less educated peers. It is possible that greater education was associated with a more global processing strategy (i.e., consideration of multiple solutions on each trial) that decreased participants' susceptibility to interference. However, each of these correlations appear to have been unduly influenced by outliers: For the mean RT difference, there was one highly educated older adult who showed strong facilitation; and for the median RT difference, there was one less educated older adult with substantial interference. When each outlier was omitted from its corresponding analysis, neither correlation remained reliable, $rs(45) = -.22$, $ps = .14$. Secondly, vocabulary was negatively correlated with interference priming in the Boxed task, regardless of whether this effect was operationalized as the difference between mean RTs, $r(46) = -.26$, $p < .10$; or as the difference between median RTs, $r(46) = -.30$, $p < .05$. These correlations suggest that the highly verbal older adults were less susceptible to interference, and perhaps more likely to use a global processing strategy, than were the less verbal older adults. However, these correlations also appear to be unduly influenced by one outlier: When one older adult with low vocabulary and high interference was omitted from the analyses, neither correlation remained reliable, $rs(45) > -.21$, $ps > .17$. Thirdly, Stroop interference was positively correlated with interference priming (the mean RT difference) in the Boxed task, $r(46) = .30$, $p < .05$. However, one older adult showed substantial interference on both the Stroop and letter deletion tasks. When this individual was omitted from the analysis, the correlation was no longer reliable, $r(45) = .13$, $p = .37$. Finally, the number of word puzzles and word games used was positively correlated with interference priming in the Boxed task, regardless of whether interference was operationalized as the difference between mean RTs, $r(46) = .40$, $p < .01$; or as the difference between median RTs, $r(46) = .36$, $p < .05$.[3.4] Surprisingly, these latter two correlations suggest that, with greater exposure to word games and word puzzles, the older adults actually became more susceptible to interference effects. It is possible that, for the older age group, expertise is associated with use of the same strategy, either automatic or otherwise, from one word puzzle to the next. This is in contrast to the younger experts in this study, whose lesser susceptibility to interference may have been related to greater flexibility in problem-solving. This reliance on a smaller number of "foolproof" strategies by older experts than by younger experts has been noted by Charness (1981, 1985) in his work with older chess players. It is therefore possible that he older "expert" game players in the present study were using a strategy (i.e., considering the same solution across presentations of the same word puzzle) that was ordinarily adaptive, but that proved costly in the context of the present interference conditions.

Discussion

The present results revealed several effects and interactions involving task order. However, because this variable was not of particular theoretical interest, these effects will be largely omitted from the present discussion. Apart from these complicating effects of order, there were several findings of interest. First, at the overall level, accuracy was higher for older than for younger adults and for the Boxed task than for the WM task. Once again, the older adults appeared to have adopted a more conservative response criterion than their younger counterparts. Secondly, mean RTs were higher for older adults, as compared to younger adults, and for the WM task, as compared to the Boxed task. Moreover, in each task, mean RTs were higher for the Interference condition than for the Control condition. Furthermore, the mean RT

[3.4] Although visual inspection of the scatterplots for the last two correlations appeared to reveal outliers, this suspicion was not borne out statistically. These correlations reflect the fact that most older adults showing facilitation in the boxed interference condition reported using two or fewer word games or word puzzles, whereas most older adults showing interference reported using three or more games or puzzles.

analyses revealed two interactions – Age x Task, and Age x Condition – which suggested that the older adults were more penalized by increased processing demands than were the younger adults. The first interaction revealed that the older adults showed a larger increase in mean RT from the Boxed task to the WM task than did the younger adults. The second interaction revealed that the older adults showed a greater increase in mean RT from the Control to Interference conditions than did the younger adults. In other words, the interference priming effect increased with age when it was operationalized as the absolute difference between mean RTs. However, the hypothesized three-way interaction was not obtained in these analyses, so the age-related increase in interference priming was reliable in both the Boxed task and the WM task. On the other hand, analyses of interference priming scores revealed a trend toward a greater age difference in interference in the WM task, as compared to the Boxed task. It was speculated that the failure to obtain a reliable three-way interaction was partly due to the high degree of variability shown by participants.

The analyses for absolute differences in median RTs and relative differences in mean and median RTs yielded somewhat different results. First, the analyses for absolute differences in medians revealed main effects of age and task and an Age x Task interaction in the same directions as before. However, there was no Age x Condition interaction, which suggested that the two age groups showed equivalent increases in median RTs from the Control to Interference conditions. In other words, this analysis yielded no age difference in interference. Secondly, analyses of relative differences in means and medians generally showed age-invariance in interference with one exception: At the overall level, the older adults showed a larger proportion increase in mean RT from the Control to Interference conditions than did the younger adults. However, this overall relative increase in mean RTs was not corroborated at the level of the individual tasks employed, nor was it confirmed in any of the analyses for relative differences in median RTs.

A further set of analyses examined the results for the first-presented of the two tasks as an attempt to eliminate the confounding effects of order. These analyses, which employed task (Boxed vs. WM) as a between-subjects variable, partly corroborated the previous analyses including task order. The highlights of these analyses include the now-standard findings of greater RTs for older adults, as compared to younger adults, and for the Interference condition, as compared to the Control condition. However, there were no reliable age differences in interference, except in the overall comparison of relative differences in mean RTs for the two age groups. Nonetheless, there was a trend toward greater interference priming in the absolute mean analyses that likely would be reliable with greater statistical power.

The present results therefore largely corroborate the results of Experiment 4. In other words, there was an age-related increase in interference priming when this effect was operationalized as the absolute increase in mean RTs across conditions. However, most other statistical methods revealed a pattern of age-invariance in interference. This discrepant result suggests that, to the extent that older people showed more interference than younger people, this greater interference was caused by an age-related increase in the number of very long RTs. This is the case because the median and proportionate analyses, which revealed no age difference, tend to dampen the effects of outliers relative to the mean analyses, which are more susceptible to outliers. This is further corroboration for differential change in the shape of the RT distributions for the two age groups – more specifically, for increasing positive skew with age.

Once again, it is somewhat surprising to have identified a condition (i.e., the Boxed task) that yielded an age difference in the present experiment but not in Experiment 3. However, it is once again likely that this discrepant result is due to a fatigue effect, or age-related increase in susceptibility to interference that is tied to the greater number of trials preceding the test phase. Namely, there were 60 trials preceding the test phase in Experiment 5, as compared to 48 in Experiment 3. Moreover, there were two blocks of trials in Experiment 5, as compared to one block in Experiment 3. It seems somewhat unlikely that episodic retrieval is the source of the age-related increase in interference in the present experiment (see, e.g., Faust, 1998), given the lack of facilitation trials. If some stimuli had been precisely reinstated, as was true in Experiment 4, then there would have been more of an incentive for participants to deliberately recollect prior instances of the same stimulus. However, the absence of such facilitation trials in the present experiment should have eliminated any such incentive.

It is somewhat unlikely that the present findings can be attributed to a simple speed-accuracy trade-off. Although the older adults were somewhat slower and more accurate at the overall level than were the younger adults, this fact alone cannot fully account for the disproportionate slowing shown by older adults in the interference conditions of the mean RT analysis. In other words, there was an Age x Condition interaction for mean RT, but not for accuracy.

Finally, although task ordering entered into several interactions in the four-way ANOVAs for speed and accuracy, the majority of these interactions were not of particular theoretical interest. The one exception to this general rule is in the accuracy ANOVA, where an Age x Order x Condition interaction was obtained. This three-way interaction occurred because the Age x Condition interaction was reliable when the WM task appeared first, but not when the Boxed task appeared first. More specifically, when the WM task appeared first, there was an age difference in accuracy favoring older adults in the Control condition but an even larger age difference in the Interference condition. However, there was not a corresponding Age x Order x

Condition interaction for mean or median RTs, which suggests that the disproportionate slowing showed by older adults in the interference conditions is not a SATO related to task ordering.

General Discussion

Three experiments employed Reingold's (1995) letter deletion task to investigate age differences in interference priming. It was previously unclear whether the pattern of age differences in this indirect interference paradigm would resemble that found in most indirect facilitation paradigms (little or no age decrement; see, e.g., Light & La Voie, 1993) or that found in most direct interference paradigms (a larger age decrement; see, e.g., Winocur & Moscovitch, 1983). The former result would have corroborated the view that age differences in memory primarily reflect age decrements in recollection (e.g., Jacoby, 1991), whereas the latter result would have supported the view that age differences in memory reflect impaired inhibitory mechanisms (e.g., Hasher & Zacks, 1988).

Experiment 3 appeared to support the former hypothesis: That is, the older and younger adults showed equivalent degrees of interference, regardless of whether the effect was assessed in terms of differences in mean or median RTs; moreover, age-equivalence was obtained for both the absolute and relative priming analyses. This result appeared to support the impaired-recollection hypothesis: That is, there appeared to be no age difference in memory when consciously controlled uses of memory were largely precluded by the retrieval task. Moreover, the two age groups appeared equally capable of inhibiting no-longer-appropriate responses under these automatic retrieval conditions. However, this initial conclusion was revised slightly based on the results of Experiments 4 and 5: In these two studies, an age-related increase in interference emerged only in the analyses for absolute differences in mean RTs. In all other analyses (absolute differences in median RTs, and relative differences in means and medians), no age differences were obtained (see, e.g., Faust, 1998, for a similar finding). This result suggests that, in the latter two experiments, there was an age-related increase in the number of very long RTs in the interference conditions. In other words, when the influence of outliers was minimized by the use of more robust statistical methods, such as medians or trimmed means (see, e.g., Ratcliff, 1993), the age difference in interference "disappeared."

Surprisingly, the age-related increase in interference for the mean RT analyses remained constant across two manipulations that were thought to disproportionately affect the older adults: frequency of prior presentation (i.e., 1 vs. 4 prior presentations in Experiment 4) and the presence/absence of a working memory load (Experiment 5). Moreover, the latter two experiments showed an age-related increase in interference under conditions (i.e., the standard, "boxed" version of the procedure from Experiment 3) that had previously shown no age difference. It was surmised that the emergence of an age difference for this condition was linked to the greater numbers of trials in Experiments 4 and 5, as compared to Experiment 3. In other words, it was thought that the older adults showed a fatigue effect in the latter two experiments that selectively impaired their inhibitory functioning. As evidence for this speculation, the older adults of Experiment 4, but not the younger adults, showed a reliable increase in mean RT from the third to fourth quartiles of the study phase. The hypothesis that inhibitory functioning showed the greatest impairments under conditions of fatigue and old age is consistent with the finding that age declines in inhibition are especially pronounced at non-optimal times of day (see, e.g., Yoon et al., 1998). However, the older adults' inhibitory impairment was not manifest in every trial: Rather, this decrement appeared to exert its effect by increasing the number of trials with very long RTs for this age group. In other words, there appeared to be greater positive skewing of the RT distribution for the older adults (i.e., a longer right tail of the distribution), as compared to the younger adults.

With further regard to the interference analyses of Experiments 4 and 5, it is true that neither manipulation of interest (stimulus repetition and working memory load) reliably increased the observed age difference in interference. Nonetheless, there were numerical trends in the predicted directions for these two experiments. For example, in Experiment 4 the older adults showed a larger numeric increase in mean RT from the Interference1x condition to the Interference4x condition (M increase = 334 ms) than did the younger adults (M increase = 246 ms). In addition, analyses of mean RTs for Experiment 5 revealed that the addition of a working memory load produced a bigger increase in interference priming in the older adults (M increase = 239 ms) than in the younger adults (M increase = 71 ms). The failure to obtain statistically reliable interactions in these two experiments may be due in part to the high degree of variability shown by participants on the letter deletion task.

Only one of the three studies (Experiment 4) included facilitation conditions: Namely, RT reliably decreased across the four presentations of each repeated stimulus in the study phase. Although the analyses of absolute differences in RTs suggested equivalent degrees of facilitation for the two age groups, the analyses of proportionate differences in RTs suggested an age-related decrement in facilitation. The former result was consistent with a large literature showing little or no age difference in repetition priming (see, e.g., La Voie & Light, 1994, for a review), which is commonly interpreted as evidence for age-invariance in automatic uses of memory. However, to the extent that age differences appear in the proportionate analyses, this discrepant result may reflect greater "test awareness", or intrusions of deliberate recollection, on the part of the younger adults as compared to the older adults (e.g., Light & La Voie, 1993).

In summary, some results of the present investigation support the impaired-inhibition hypothesis, whereas other results support the impaired-recollection hypothesis. On the one hand, there is an age-related increase in interference priming in Experiments 4 and 5, when this effect is defined as the absolute increase in mean RT from the Control to Interference conditions. On the other hand, there is age-equivalence in interference priming in all of the analyses for Experiment 1 and in the proportionate and median RT analyses of Experiments 4 and 5. These seemingly contradictory results suggest that, under conditions of fatigue, the older adults demonstrate a decrease in inhibitory functioning that manifests itself as increased variability in the right tail of the RT distribution. However, when the effects of fatigue and outlying observations are controlled for, either experimentally or statistically, the two age groups appear equivalent in inhibitory functioning. The interference results therefore provide some evidence for a mild inhibitory deficit with age: These results cannot be readily explained in terms of impaired recollection, because the interference conditions should largely preclude deliberate referencing of previously studied information. In other words, there should be little incentive to deliberately recollect prior instances of a stimulus because such conscious uses of memory can only detract from performance. In contrast, the facilitation priming results for Experiment 4 provide some evidence for age-related decrements in recollection. In these facilitation conditions, there is a greater incentive to deliberately recollect the previous occurrence of a stimulus because such consciously controlled retrieval can further bolster the automatic retrieval produced by repetition.

The present results suggest that cognitive aging researchers should not limit their RT analyses to only one statistical method. Rather, a consideration of multiple methods (e.g., Ratcliff, 1993) may be necessary to gain a full understanding of age-related changes in cognitive processes. Moreover, it appears that age differences in memory may reflect age-related changes in multiple mechanisms (i.e., both conscious recollection and inhibition) and that the efficiency of these mechanisms may vary across experimental conditions. Perhaps the most parsimonious explanation for the present findings is that aging produces a general decline in cognitive control, which may manifest itself as either a loss of explicit memory and/or as a lesser ability to suppress irrelevant information, depending on subject and task parameters (see, e.g., Hay & Jacoby, 1999, for a similar suggestion). For example, a sample of community-dwelling older adults might exhibit modest declines in both recollection and inhibition. However, the decline in recollection would not be apparent in the interference conditions of an indirect memory test, where there is little conscious incentive to refer to previously studied information. On the other hand, these interference conditions might reveal a mild inhibitory decline (i.e., an increase in the number of trials with disproportionately long RTs), but only when the high-functioning older adults are fatigued. In contrast, a group of institutionalized older adults would likely show more pervasive recollective and inhibitory declines (see, e.g., Moroz, 1995, for similar suggestions). For instance, these lower-functioning older adults might show elevated RTs on most interference trials, rather than on a small subset of trials.

CHAPTER FOUR:
GENERAL SUMMARY AND CONCLUSIONS

Five experiments examined age differences in memory under facilitation and interference conditions of both a direct and an indirect memory test. For the purposes of the dissertation, facilitation was defined as the reinstatement of previously learned associations and contexts, whereas interference was defined as the provision of different associations and/or contexts than those employed in prior learning. These experiments served to empirically differentiate between two competing accounts of aging and memory: impaired inhibition (e.g., Hasher & Zacks, 1988), versus impaired recollection (e.g., Jacoby, 1991). More specifically, the impaired-inhibition account suggested that age differences in memory should be reliable under the interference conditions, but not the facilitation conditions, of both direct and indirect memory tests. In contrast, the impaired-recollection account suggested that there should be reliable age differences on direct memory tests, regardless of condition (i.e., facilitation or interference), but little or no age difference on indirect tests.

Experiments 1 and 2 examined the facilitatory versus interfering effects of repetition in a direct memory test, based on Jacoby's (1991) process dissociation procedure or PDP (see also, e.g., Hay & Jacoby, 1996). In these two experiments, older and younger adults studied sentence completions that had been either the dominant or the non-dominant completion for that sentence in a prior training phase. It was expected that both age groups would find it easier to recall sentence completions that had appeared on three of the four training trials for a sentence (a facilitation condition), as compared to completions that had appeared on only one of the four training trials (an interference condition). Moreover, it was expected that the older adults would perform significantly worse than the younger adults when recalling once-trained completions; however, it was unclear whether reliable age differences would appear for thrice-trained completions. The first two experiments revealed that recall was better for younger than older adults and for thrice-trained than once-trained completions. Moreover, in both experiments, age differences in recall were numerically greater when sentence completions were once-trained, as compared to thrice-trained; however, this Age x Presentation frequency interaction was not reliable in either experiment. Furthermore, both experiments revealed an age-related decline in the PDP-derived estimate of conscious recollection, as well as an age-related increase in intrusions of non-target sentence completions. However, these age decrements in recollection and inhibition were reliable only in Experiment 2. In contrast, the PDP-derived estimate of habitual influences failed to reveal age differences in either experiment.

Of further interest in these first two experiments is the fact that across the critical and baseline conditions, the total number of target presentations across all experimental phases ranged from 0 to 4 presentations. Additional presentation frequency analyses bridging across the baseline and critical conditions revealed that age differences decreased with increasing stimulus presentations: Age differences were either non-reliable or marginally reliable at 3 or 4 target presentations. This result suggests that stimulus repetition can be an effective form of environmental support for older adults; however, this result should be viewed somewhat cautiously, given the trends toward ceiling in both experiments and the failure to eliminate age differences at higher presentation frequencies in Experiment 2.

In addition to varying presentation frequency, the first two experiments each varied a particular stimulus attribute that was expected to influence the magnitude of age differences in recall. In Experiment 1, the perceptual characteristics of stimuli were varied by presenting sentences in either the same voice or different voices across experimental phases. In Experiment 2, the extent to which participants could use prior knowledge to recall information was varied: Specifically, sentence completions were either medium-cloze (moderately predictable) or low cloze (relatively unpredictable). It was expected that repeatedly reinstating perceptual and conceptual contexts (in Experiment 1) and schema-consistent information (in Experiment 2) would benefit the older adults more than the younger adults. These predictions were based on Craik and Jennings' (1992) notions of environmental support: That is, precise reinstatement of meaningful study cues (see also Rabinowitz et al., 1982; and Shaw & Craik, 1989) and use of prior knowledge (see also Charness, 1985; and Hess, 1990) should minimize age differences in memory by reducing the need for self-initiated processes, in which older adults are deficient. Unfortunately, neither of these manipulations reliably reduced age differences in recall. Although age differences were negligible in Experiment 1 when voice context was precisely reinstated, the Age x Voice context interaction was not a reliable one. Moreover, in Experiment 2, the magnitude of age differences in recall was equivalent for low-cloze and medium-cloze completions. On the other hand, there was some evidence in this experiment that the rather sizable age difference in the Once-trained/Low-cloze condition was reduced by increasing either the presentation frequency or the cloze value of sentence completions. However, employing both frequently presented *and* schema-congruent completions did not further reduce age differences.

It is somewhat surprising that age differences were almost eliminated under highly supportive conditions in Experiment 1, but not in Experiment 2. For example, older and younger adults were almost equivalent in their recall of thrice-trained, voice-congruent sentence completions in the first experiment, but not in the second experiment. Moreover, it was surprising that the addition of schema-congruent sentences in Experiment 2 was of no further benefit to the older adults. It is possible that the participants sampled in these two experiments differed along some individual-difference variable that affected their ability to utilize environmental support. Namely, the older adults were more highly educated in Experiment 1 than in Experiment 2. The possibility that subject parameters such as education can influence the nature of Age x Condition interactions has been suggested elsewhere (e.g., Craik et al., 1987; Craik & Jennings, 1992).

In summary, the first two studies examining age differences in facilitation and interference conditions of a direct memory test revealed modest evidence of age decrements in both recollection and inhibition. The finding of an age-related decrease in recall of repeated targets, together with an age-related increase in intrusions of repeated distractors, suggests that aging produces a general failure of cognitive control, rather than a specific inhibitory failure (see, e.g., Hay & Jacoby, 1999, for a similar suggestion).

The remaining three experiments examined age differences in an indirect memory test – Reingold's (1995) letter deletion task. In this task, older and younger adults were asked to remove one of two specified letters from a six-letter nonword in order to make a meaningful five-letter word from the remaining letters. In the facilitation trials, a previously seen nonword reappeared with the same letters specified as candidates for deletion as in the nonword's previous appearance. In the interference trials, a nonword reappeared, albeit with different letters specified as candidates for deletion than in the nonword's previous appearance. In other words, each word puzzle had two possible solutions, but only one solution was permissible on any given trial. Facilitation and interference priming were operationalized as changes in the time to solve a word puzzle (either decreases or increases in RT, respectively) as a function of previous exposure to that puzzle. It was expected that there should be little or no age difference in the facilitation priming conditions (Experiment 4), in accordance with previous work in this area (see, e.g., La Voie & Light, 1994, for a review). However, it was unclear whether the older adults should show as much interference priming (Experiments 3, 4, and 5) as the younger adults, or more interference priming. The former prediction is consistent with the impaired-recollection hypothesis (e.g., Jacoby, 1991): That is, age differences should be negligible under automatic retrieval conditions. The latter prediction is consistent with the impaired-inhibition hypothesis (e.g., Hasher & Zacks, 1988), which suggests that aging increases one's susceptibility to interference.

Experiment 3 provided preliminary support for the impaired-recollection hypothesis: In this experiment, which employed only interference and control conditions, the two age groups showed equivalent interference priming, regardless of the statistic used to measure the change in RT across conditions. However, a less straightforward result was obtained in the two subsequent experiments, which employed manipulations thought to disproportionately increase interference for older adults (i.e., stimulus repetition in Experiment 4, and a working memory load in Experiment 5). In these latter two experiments, the older adults showed larger increases in mean RTs from the control to interference conditions than did the younger adults. However, this age-related increase in interference priming was also observed in the standard (boxed) conditions, which had previously revealed no age difference (Experiment 3). To further complicate matters, a pattern of age-invariance in interference priming was observed when interference was assessed in terms of either the absolute increase in median RT or the proportionate increase in either mean or median RT. These seemingly contradictory results may be explained by the fact that the two studies which provided evidence for an age-related increase in interference (Experiments 4 and 5) employed more trials than did Experiment 3, which revealed no age difference. Moreover, in the final two experiments, there appears to have been an age-related increase in the number of interference trials with very long RTs (see also Footnote 3.2). It is therefore possible that under conditions of fatigue, the older adults demonstrate a decrease in inhibitory functioning that manifests itself as increased variability in the right tail of the RT distribution.[4.1]

With regard to facilitation effects, Experiment 4 was the only letter deletion study to include facilitation conditions. More specifically, some nonwords appeared a total of four times prior to a change in candidate letters for deletion. Older and younger adults showed equivalent absolute decreases in mean or median RT across the four presentations of each stimulus. However, the older adults showed lesser facilitation than the younger adults did when this effect was assessed in terms of proportionate changes in RT across conditions. The former result was consistent with previous results showing no age difference in automatic uses of memory (e.g., Jacoby et al., 1997; and La Voie & Light, 1994). However, to the extent that age differences appear in the proportionate analyses, this discrepant result may reflect intrusions of deliberate recollection, in which the older adults are deficient (e.g., Light & La Voie, 1993).

In summary, the three studies examining age differences in an indirect test provided some evidence for an age-related inhibitory decline: Namely, age differences in interference priming emerged in two studies that employed a fairly large number of trials (Experiments 4 and 5), but not in the one study that employed fewer trials (Experiment 3). Surprisingly, however, the age difference appeared to be caused by an age-related increase in very long RTs for a small subset of interference trials, rather than an increase in RT across most trials. Moreover, the two manipulations that were predicted to disproportionately increase interference for the older adults (i.e., stimulus repetition and a working memory load) did not yield the predicted result: In other

---

[4.1] The age-related increase in interference for the analyses of absolute differences in mean RTs, but not median RTs, suggests that an ex-Gaussian analysis be performed to examine the shape of the RT distribution. However, there do not appear to be enough observations per cell to warrant this type of analysis (E. Reingold, personal communication, April 1998).

words, both age groups were equivalently penalized by these demanding processing conditions. However, there were non-reliable trends in the predicted directions for the last two experiments. In addition, there was some evidence for an age decline in recollection, if it is assumed that the age decrement in facilitation priming (in the proportionate analyses of Experiment 4) was due to intrusions of conscious recollection, in which the older adults were impaired.

Taken as a whole, the five experiments suggest that age differences in memory reflect general failures of cognitive control, rather than a specific inhibitory failure (see, e.g., Hay & Jacoby, 1999, for a similar suggestion). However, it also appears that the specific nature of the age-related deficit depends substantially on subject and task parameters (see, e.g., Craik & Jennings, 1992, and Moroz, 1995, for similar suggestions). For example, reliable age decrements in recollection and inhibition emerged on a direct memory test when the older adults were less educated than the younger adults (Experiment 2), but not when the reverse was true (Experiment 1). Moreover, age-related inhibitory declines appeared on an indirect test when a fairly large number of trials were employed (Experiments 4 and 5), but not when fewer trials were employed (Experiment 3). This latter result is consistent with the finding that the magnitude of age-related inhibitory declines may fluctuate as a function of fatigue or Circadian arousal (see, e.g., Yoon et al., 1998).

Correlational analyses in the five experiments were undertaken in order to further identify the relations among individual-difference measures and memory performance measures. Although several of these correlations appear to have been unduly influenced by one or two outlying observations, a few results withstood further statistical scrutiny. For example, older adults with a high degree of frontal-lobe functioning, as assessed by verbal fluency and Stroop task performance, had better recall of once-studied items than did older adults with lesser frontal-lobe functioning (Experiments 1 and 2). This result suggests that recall of once-presented sentence completions is a strategic task for older adults which depends heavily on the frontal lobes (see also, e.g., Moscovitch & Winocur, 1992). In addition, further analyses suggested that older adults' susceptibility to interference on the letter deletion task was affected by the number of word games and word puzzles that they used in daily life (Experiment 5). However, this relation was not in the expected direction: Namely, the magnitude of the interference effect increased with greater exposure to games and puzzles in daily life. It is therefore possible that the older "expert" game players used a strategy (i.e., considering the same solution across presentations of the same word puzzle) that was adaptive in daily life, but that proved costly in the context of the present interference conditions.

Clearly, much work is still needed to identify the locus of age-related memory difficulties. Although age differences in memory were generally greater under the present interference conditions than under the facilitation conditions, this was not a hard and fast rule. However, the present results suggest that older adults may suffer from general failures of cognitive control, which may encompass difficulties in intentional remembering of target information and/or inhibition of irrelevant information. The magnitude of age decrements in recollection and inhibition may partly depend on sampling characteristics: For example, Moroz (1995) recently suggested that community-dwelling older adults may exhibit a decline in recollection, whereas institutionalized older adults may show declines in both recollection and inhibition. To take this analogy a step further, it is possible that both populations of older adults show an inhibitory deficit, but that this deficit manifests itself in different ways for the two groups: For example, on the letter deletion task, high-functioning elderly may show a fatigue-related increase in RT on a minority of the interference trials, whereas low-functioning elderly may show a disproportionate increase in RT on most interference trials. Therefore, an obvious future direction would be to adapt the present tasks for use in patient populations. In addition, it would be interesting to examine the correlations among the present memory measures and a wider array of neuropsychological tasks. For example, it is of theoretical interest to determine whether performance under indirect interference conditions is mediated more by frontal-lobe functioning, which is associated with consciously controlled processes such as inhibition, or by hippocampal and medial temporal functioning, which is associated with automatic memory processes (see also, e.g., Moscovitch, 1992, 1994; and Winocur, Moscovitch, & Bruni, 1996).

Future work should also seek to rehabilitate age-related memory deficits. Such efforts have already begun: For example, Jacoby, Jennings, and Hay (1996) have described a procedure that has been used to rehabilitate recollection in older adults. More specifically, Jacoby and colleagues used an interference or opposition condition in which participants were to say "yes" if test items had appeared once in a previous study list but were to say "no" if the test items were new or were "catch items" (i.e., items that appeared twice on the current study list). Older adults are more likely than younger adults to erroneously respond "yes" to catch words because aging produces a decline in recollection but spares automatic memory influences (see also Jennings & Jacoby, 1997). Jacoby et al. (1996) have increased recollection in older adults by using a training procedure in which participants learned to "exclude" repeated items over progressively longer test intervals. Similarly, Hay and Jacoby (1999) have recently reported that age declines in recollection can be "repaired" under supportive processing conditions. Specifically, older and younger adults were equally able to recollect words that were cued with distinctive associative contexts when more retrieval time was provided to the older age group. Unfortunately, the present manipulations of environmental support were not as effective at reducing age differences in recollection as were the manipulations employed by Jacoby and his colleagues. Nonetheless, the present experiments suggest that age differences in recollection can be reduced when information is repeated (Experiments 1 and 2). Moreover, although reinstatement of stimulus characteristics and prior knowledge structures did not specifically reduce age differences in recall, these manipulations were still beneficial to older adults. Furthermore, the present data suggest that age decrements in inhibition are minimal when conditions are not unduly fatiguing (Experiment 3) or when competing responses are relatively weak (Experiment 2; but see Experiment 4 for a conflicting result). Speculatively, it might be possible to more actively rehabilitate inhibitory processes in a manner similar to Jacoby et al.'s (1996) procedure for rehabilitating recollection: That is, participants might learn to inhibit progressively larger amounts of irrelevant information (e.g., concurrent distractors) over the course of several training sessions. However, this methodology might only be warranted for older adults exhibiting impaired inhibition but spared recollection (e.g., frontal patients). Alternatively, however, manipulations which emphasize automatic uses of memory (e.g., the method of spaced retrieval, as described by Camp & Schaller, 1989) might serve as complementary means of preserving memory in old age. Future research should therefore aim to devise interventions for the specific memory impairments observed in different populations of older adults.

REFERENCES

Anderson, N. D. (1997). The attentional demands of encoding and retrieval in younger and older adults. Unpublished doctoral dissertation, University of Toronto, Toronto, Canada.

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

Battig, W. F., & Montague, W. E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut Category Norms. Journal of Experimental Psychology Monograph, 80, 1-46.

Bloom, P. A., & Fischler, I. (1980). Completion norms for 329 sentence contexts. Memory and Cognition, 8, 631-642.

Borkowski, J. G., Benton, A. L., & Spreen, O. (1967). Word fluency and brain damage. Neuropsychologia, 5, 135-140.

Brown, A. S., Neblett, D. R., Jones, T. C., & Mitchell, D. D. (1991). Transfer of processing in repetition priming: Some inappropriate findings. Journal of Experimental Psychology: Learning, Memory, and Cognition, 17, 514-525.

Brown, S. C. (1994). Age and depth of processing: Relative contributions of automaticity and recollection. Unpublished master's thesis, University of Toronto, Toronto, Canada.

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

Butler, K., Zacks, R. T., & Henderson, J. M. (in press). Age comparisons on an antisaccade task. Memory and Cognition.

Camp, C. J., & Schaller, J. R. (1989). Epilogue: Spaced retrieval memory training in an adult day-care center. Educational Gerontology, 15, 641-648.

Charness, N. (1981). Search in chess: Age and skill differences. Journal of Experimental Psychology: Human Perception and Performance, 7, 467-476.

Charness, N. (1985). Age and expertise: Responding to Talland's challenge. Paper presented at the George A. Talland Memorial Conference on Aging and Memory, Cape Cod, MA.

Connelly, S. L., Hasher, L., & Zacks, R. T. (1991). Age and reading: The impact of distraction. Psychology and Aging, 4, 533-541.

Craik, F. I. M. (1983). On the transfer of information from temporary to permanent memory. Philosophical Transactions of the Royal Society of London, Series B 302, 341-359.

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

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.

Craik, F. I. M., Byrd, M., & Swanson, J. M. (1987). Patterns of memory loss in three elderly samples. Psychology and Aging, 2, 79-86.

Craik, F. I. M., & Jennings, J. M. (1992). Human memory. In F. I. M. Craik & T. A. Salthouse (Eds.), Handbook of aging and cognition (pp. 51-110). Hillsdale, NJ: Erlbaum.

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

Craik, F. I. M., Morris, R. G., & Gick, M. L. (1990). Adult age differences in working memory. In G. Vallar & T. Shallice (Eds.), Neuropsychological impairments of short-term memory (pp. 247-267). Cambridge, England: Cambridge University Press.

Faust, M. E. (1998, April). The effect of target repetition on negative priming in older and younger adults: Evidence from picture naming. Poster presented at the Seventh Cognitive Aging Conference, Atlanta.

Francis, W. N., & Kucera, H. (1982). Frequency analysis of English usage: Lexicon and grammar. Boston: Houghton Mifflin.

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

Golden, C. J. (1978). Stroop color and word test. Chicago: Stoelting.

Graf, P., & Komatsu, S. (1994). Process dissociation procedure: Handle with caution! European Journal of Cognitive Psychology, 6, 113-129.

Hamm, V. P., & Hasher, L. (1992). Age and the availability of inferences. Psychology and Aging, 7, 56-64.

Hartley, A. A. (1992). Attention. In F. I. M. Craik & T. A. Salthouse (Eds.), The handbook of aging and cognition (pp. 3-49). Hillsdale, NJ: Erlbaum.

Hartman, M., & Hasher, L. (1991). Aging and suppression: Memory for previously relevant information. Psychology and Aging, 6, 587-594.

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. Bower (Eds.), The psychology of learning and motivation (Vol. 22) (pp. 193-225). New York: Academic Press.

Hasher, L., Zacks, R. T., & May, C. P. (1999). Inhibitory control, circadian arousal, and age. In D. Gopher & A. Koriat (Eds.), Attention and performance XVII: Cognitive regulation of performance – Interaction of theory and application. (pp. 653-675). Cambridge, MA: MIT Press.

Hay, J. F., & Jacoby, L. L. (1996). Separating habit and recollection: Memory slips, process dissociations, and probability matching. Journal of Experimental Psychology: Learning, Memory, and Cognition, 22, 1323-1335.

Hay, J. F., & Jacoby, L. L. (1999). Separating habit and recollection in young and elderly adults: Effects of elaborative processing and distinctiveness. Psychology and Aging, 14, 122-134.

Hess, T. M. (1990). Aging and schematic influences on memory. In T. M. Hess (Ed.), Aging and Cognition: Knowledge Organization and Utilization (pp. 93-160). North Holland: Elsevier.

Howard, D. V. (1980). Category norms: A comparison of the Battig and Montague (1969) norms with the responses of adults between the ages of 20 and 80. Journal of Gerontology, 35, 225-231.

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

Jacoby, L. L., Jennings, J. M., & Hay, J. F. (1996). Dissociating automatic and consciously controlled processes: Implications for diagnosis and rehabilitation of memory deficits. In D. J. Herrmann, C. L. McEvoy, C. Hertzog, P. Hertel, & M. K. Johnson (Eds.), Basic and applied memory research: Theory in context (Vol. 1, pp. 161-193). Mahwah, NJ: Erlbaum.

Jacoby, L. L., Yonelinas, A. P., & Jennings, J. M. (1997). The relation between conscious and unconscious (automatic) influences. A declaration of independence. In J. Cohen & J. W. Schooler (Eds.), Scientific approaches to consciousness (pp. 13-47). Mahwah, NJ: Erlbaum.

Jenkins, J. J. (1979). Four points to remember: A tetrahedral model for memory experiments. In L. S. Cermak & F. I. M. Craik (Eds.), Levels of processing in human memory (pp. 429-446). Hillsdale, NJ: Erlbaum.

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

Jennings, J. M., & Jacoby, L. L. (1997). An opposition procedure for detecting age-related deficits in recollection: Telling effects of repetition. Psychology and Aging, 12, 352-361.

Kane, M. J., Hasher, L., Stoltzfus, E. R., Zacks, R. T., & Connelly, S. L. (1994). Inhibitory attentional mechanisms and aging. Psychology and Aging, 9, 103-112.

Kausler, D. H. (1994). Learning and Memory in Normal Aging. San Diego: Academic Press.

Kimberg, D. Y., D'Esposito, M., & Farah, M. J. (1997). Cognitive functions in the prefrontal cortex – Working memory and executive control. Current Directions in Psychological Science, 6, 185-191.

La Voie, D., & Light, L. L. (1994). Adult age differences in repetition priming: A meta-analysis. Psychology and Aging, 9, 539-553.

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

Light, L. L. (1992). The organization of memory in old age. In F. I. M. Craik & T. A. Salthouse (Eds.), Handbook of aging and cognition (pp. 111-165). Hillsdale, NJ: Erlbaum.

Light, L. L., & La Voie, D. L. (1993). Direct and indirect measures of memory in old age. In P. Graf & M. E. J. Masson (Eds.), Implicit memory: New directions in cognition, development, and neuropsychology (pp. 207-230). Hillsdale, NJ: Erlbaum.

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

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

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

McDowd, J. M., & Filion, D. L. (1995). Aging and negative priming in a location suppression task: The long and the short of it. Psychology and Aging, 10, 34-47.

McDowd, J. M., & Oseas-Kreger, D. M. (1991). Aging, inhibitory processes, and negative priming. Journal of Gerontology, 46, 340-345.

Moroz, T. (1995). Interference in young and old adults: Poor explicit memory? Unpublished master's thesis, University of Toronto, Toronto, Canada.

Moscovitch, M. (1992). A neuropsychological model of memory and consciousness. In L. R. Squire & N. Butters (Eds.), Neuropsychology of Memory (2nd ed.). New York: Guilford.

Moscovitch, M. (1994). Memory and working with memory: Evaluation of a component process model and comparisons with other models. In D. L Schacter & E. Tulving (Eds.), Memory Systems 1994 (pp. 269-310). Cambridge, MA: MIT Press/Bradford.

Moscovitch, M., & Winocur, G. (1992). The neuropsychology of memory and aging. In F. I. M. Craik & T. A. Salthouse (Eds.), Handbook of aging and cognition (pp. 315-372).

Moscovitch, M., & Winocur, G. (1995). Frontal lobes, memory, and aging. Annals of the New York Academy of Sciences, 769, 119-150.

Munoz, D. P., Broughton, J. R., Goldring, J. E., & Armstrong, I. T. (1998). Age-related performance of human subjects on saccadic eye movement tasks. Experimental Brain Research, 121, 391-400.

Neill, W. T., & Valdes, L. A. (1992). Persistence in negative priming: Steady state or decay? Journal of Experimental Psychology: Learning, Memory, and Cognition, 18, 565-576.

Newcombe, F. (1969). Missile wounds of the brain. London: Oxford University Press.

Park, D. C., & Shaw, R. (1992). Effect of environmental support on implicit and explicit memory in younger and older adults. Psychology and Aging, 7, 632-642.

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control: The Loyola symposium. In R. L. Solso (Ed.), Information processing and cognition (pp. 55-85). Hillsdale, NJ: Erlbaum.

Rabinowitz, J. C., Craik, F. I. M., & Ackerman, B. P. (1982). A processing resource account of age differences in recall. Canadian Journal of Psychology, 36, 325-344.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. Psychological Bulletin, 114, 510-532.

Raven, J. C. (1965). Mill Hill Vocabulary Scale. London: H. K. Lewis.

Reingold, E. (1995). Facilitation and interference in indirect/implicit memory tests and in the process dissociation paradigm: The letter insertion and the letter deletion tasks. Consciousness and Cognition, 4, 459-482.

Reitan, R. M. (1955). The relation of the Trail Making Test to organic brain damage. Journal of Consulting Psychology, 19, 393-394.

Reitan, R. M. (1958). Validity of the Trail Making Test as an indicator of organic brain damage. Perceptual and Motor Skills, 8, 271-276.

Roediger, H. L., III, & McDermott, K. B. (1993). Implicit memory in normal human subjects. In H. Spinnler & F. Boller (Eds.), Handbook of neuropsychology (pp. 63-131). Amsterdam: Elsevier.

Ruch, F. L. (1934). The differentiative effects of age upon human learning. Journal of General Psychology, 11, 261-286.

Salthouse, T. A. (1985). A theory of cognitive aging. Amsterdam: North-Holland.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. Psychological Review, 103, 403-428.

Salthouse, T. A., & Meinz, E. J. (1995). Aging, inhibition, working memory and speed. Journal of Gerontology: Psychological Sciences & Social Sciences, 50(B), P297-P306.

Schacter, D. L. (1987). Implicit memory: History and current status. Journal of Experimental Psychology: Learning, Memory, and Cognition, 13, 501-518.

Schneider, W. (1995). Micro Experimental Laboratory Professional Version 2.0. [Computer software.] Pittsburgh, PA: Psychology Software Tools.

Shaw, R. J., & Craik, F. I. M. (1989). Age differences in predictions and performance on a cued recall task. Psychology and Aging, 4, 131-135.

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.

Spieler, D. H., Balota, D. A., & Faust, M. E. (1996). Stroop performance in healthy younger and older adults and in individuals with dementia of the Alzheimer's type. Journal of Experimental Psychology: Human Perception and Performance, 22, 461-479.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. Journal of Experimental Psychology, 18, 643-662.

Tipper, S. P. (1992). Selection for action: The role of inhibitory mechanisms. Current Directions in Psychological Science, 1, 105-109.

Tipper, S. P., & Cranston, M. (1985). Selective attention and priming: Inhibitory and facilitatory effects of ignored primes. The Quarterly Journal of Experimental Psychology, 37A, 591-611.

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

Verhaeghen, P., & De Meersman, L. (1998). Aging and the Stroop effect: A meta-analysis. Psychology and Aging, 13, 120-126.

Wechsler, D. (1981). Manual for the Wechsler Adult Intelligence Scale – Revised. New York: The Psychological Corporation.

West, R. L. (1996). An application of prefrontal cortex function theory to cognitive aging. Psychological Bulletin, 120, 272-292.

West, R., & Baylis, G. C. (1998). Effects of increased response dominance and contextual disintegration on the Stroop interference effect in older adults. Psychology and Aging, 13, 206-217.

Winocur, G., & Moscovitch, M. (1983). Paired-associate learning in institutionalized and noninstitutionalized older people: An analysis of interference and context effects. Journal of Gerontology, 38, 455-464.

Winocur, G., Moscovitch, M., & Bruni, J. (1996). Negative transfer (AB-AC) learning in normal old people and in patients with unilateral temporal-lobe or frontal-lobe lesions: Heightened interference on conceptual, implicit tests of memory. Brain and Cognition, 59, 44-58.

Winocur, G., Moscovitch, M., & Stuss, D. T. (1996). Explicit and implicit memory in the elderly: Evidence for double dissociation involving medial temporal- and frontal-lobe functions. Neuropsychology, 10, 57-65.

Yoon, C., May, C. P., & Hasher, L. (1998). Aging, circadian arousal patterns, and cognition. In N. Schwarz, D. C. Park, B. Knauper, & S. Sudman (Eds.), Cognition, aging, and survey measurement (pp. 117-143). Philadelphia: Psychology Press.

Zacks, R., & 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., Radvansky, G. A., & Hasher, L. (1996). Studies of directed forgetting in older adults. Journal of Experimental Psychology: Learning, Memory, and Cognition, 22, 143-156.

Appendix 2.1
Experiment 1 Stimuli

144 Sentences and Their Alternative Experimental Completions

1. A fruit tray was prepared for a (birthday / wedding).
2. A man cut his (lip / chin) while shaving.
3. A new job was offered at the (bank / factory).
4. After the game, the coach bought the kids (ice cream / pizza).
5. Amy washes her hair every (morning / night).
6. Angie uses the stamp on her desk to stamp her (phone number / name).
7. Bears do not like the smell of burning (leaves / coal).
8. Cellists often complain about stiff (shoulders / necks).
9. Customers to the store often buy (pants / sweaters).
10. Dan is replacing the (tiles / carpet) on the floor.
11. David wanted some (oatmeal / pancakes) for breakfast.
12. Doris keeps seashells in her (bathroom / living room).
13. Douglas does not eat any (meat / fish).
14. Geena wishes she had (curly / straight) hair.
15. George wants to be a (monster / clown) for Halloween.
16. Grapes from the vineyard are used to make (champagne / sherry).
17. In her bathroom drawers, Donna keeps her (tweezers / nail cutter).
18. In the businessman's suitcase, there are many (documents / files).
19. Jack seasons his food with (pepper / mustard).
20. Jane puts pictures in her scrapbook with (tape / glue).
21. Jenny wants to paint her bedroom (black / purple).
22. Jimmy likes (cheddar / mozzarella) cheese in his sandwiches.
23. Joe's video collection has a variety of (comedies / dramas).
24. John is reading the (classifieds / sports) section of the newspaper.
25. Julie served her cake with some (coffee / tea).
26. Karen (runs / swims) a mile every other day.
27. Kids like to pitch (dimes / nickels) into the water fountain.
28. Marge uses the ladder in her kitchen to reach the (sugar / flour).
29. Michelle likes her bed sheets to made out of (linen / satin).
30. Mickie bought herself a new pair of (gloves / glasses).
31. Most young people are concerned about the (economy / environment).
32. Nancy was frightened by the story about (ghosts / vampires).
33. Neil ran late and missed his (train / plane).
34. Next, the chess player will move his (queen / bishop).
35. On the stove, John is heating up some (sauce / soup).
36. One of the dealers in the casino wants to be a (priest / rabbi).
37. One pet these farmers have is a (Dalmatian / poodle).
38. One tool needed to make the cabinets is a (drill / screwdriver).
39. Pat likes the smell of the new (cologne / perfume).
40. People travel to the beach by (foot / car).
41. Pierre's favorite card game is (poker / bridge).
42. Rescuers at the lake are trying to find a lost (boy / man).
43. Sammy was the (only / first) one to hear the secret.
44. Sarah's holiday decorations include (candles / wreaths).
45. Some of the fish in the river are young (salmon / mackerel).
46. Some sections of the Great Wall of China protect a (city / town).

Appendix 2.1 (continued)

47. Somebody threw a (rock / ball) at the house.
48. Steven keeps his old receipts in (bags / boxes).
49. Students should keep their homework in a (binder / folder).
50. The African artifact was discovered by a (native / prospector).
51. The artist is skilled at (sculpture / drawing).
52. The astronomer just discovered a new (comet / planet).
53. The average age of drivers in the town is (28/ 44).
54. The baby was born in the (summer / fall).
55. The bar ran out of (chips / nuts).
56. The bicycling family is going to visit their (grandmother / aunt).
57. The bike had broken (pedals / handlebars).
58. The boat is smuggling (opium / hashish).
59. The book on the patio table has (recipes / maps).
60. The buildings are made out of (steel / iron).
61. The canoeists first met at a (contest / competition).
62. The (captain / general) led his troops into battle.
63. The (cathedral / castle) is over one thousand years old.
64. The cloud was shaped like a (dolphin / seal).
65. The coastal house is home to a (bachelor / couple).
66. The computer genius can do (statistics / programming).
67. The computer's (keyboard / printer) is not working properly.
68. The couple is using a map to get to the (theatre / museum).
69. The couple is watching their children play (soccer / hockey).
70. The dancing young people are at a (club / disco).
71. The daredevil is going to jump off the (tower / cliff).
72. The deadly lightning was followed by some (sleet / rain).
73. The desert area has some (lizards / snakes).
74. The destructive fire was started by a (cigarette / grease fire).
75. The driver of the Cadillac studies (chemistry / physics).
76. The electricity was generated by (solar / wind) power.
77. The estate is owned by a (governor / senator).
78. The exercising cricket player knows how to ride a (helicopter / boat).
79. The family room is in the home of a (dentist / lawyer).
80. The father was fixing the (radio / clock).
81. The fields have some (violets / gardenias).
82. The fighter plane is armed with (rockets / bombs).
83. The flower's (stem / leaf) got damaged by the wind.
84. The forest has deposits of (rubies / emeralds).
85. The friends keep in touch with (postcards / letters).
86. The girl is getting dressed up to go to (synagogue / church).
87. The goalie is particularly good at (shooting / blocking).
88. The golfer has already lost (2 / 3) balls.
89. The hair stylist enjoys dancing (waltzes / tangos).
90. The horse's owner lives in a (cottage / cabin).
91. The hotel's swimming pool has a (slide / volleyball net).
92. The insects in a flower garden include (wasps / ants).
93. The kids like to eat (plums / apricots) in the play room.
94. The kitchen was full of (fumes / odors).
95. The koala lives in a (park / reserve).

Appendix 2.1 (continued)

96. The lady with the back pains had surgery on her (kidney / liver).
97. The lady with the bad headache has trouble (studying / working).
98. The little girl is moving and has packed all of her (puzzles / crayons).
99. The man caught littering was once charged with (assault / robbery).
100. The man holding the rifle often drinks (scotch / brandy).
101. The man in the hotel room is working on a (budget / report).
102. The man purchasing the CD was once bitten by a (viper / rattlesnake).
103. The man who shaved his head just got (the measles / chicken pox).
104. The men in the boat at sea rarely (eat / sleep).
105. The mother was cleaning the (windows / mirror).
106. The muffins are low in (fat / cholesterol).
107. The music was too (soft / slow) for Anthony's taste.
108. The musicians are touring across the (province / country).
109. The new model car is being tested for its (emissions / engine).
110. The novel did not have a good (setting / plot).
111. The old bridge was made out of (rope / boards).
112. The old coat has a lot of (buttons / zippers).
113. The old men are playing (chess / checkers).
114. The patient suffered from (diabetes / heart disease).
115. The people are dining on (chicken / lobster).
116. The people in the car are going (fishing / camping).
117. The prisoner is praying for (forgiveness / compassion).
118. The prize in the snowshoe race is a (trophy / snowmobile).
119. The running boy is playing (hide-and-seek / tag).
120. The science student just got a new (microscope / telescope).
121. The space station is a long-term (project / goal).
122. The spelling bee was held in the (auditorium / gymnasium).
123. The strawberry farmer also grows some (soy beans / corn).
124. The streetcar is headed (north / west).
125. The tourists on the boat wanted to see the (mountains / island).
126. The tourists on the train are going next to the (roller coaster / Ferris wheel).
127. The truck in the parade usually transports (furniture / food).
128. The two kids blowing bubbles are (friends / sisters).
129. The used car is covered in (rust / dirt).
130. The waterfall flows into a (lake / river).
131. The wealthy man with the umbrella invests in (stocks / bonds).
132. The weight lifter is expected to win a (silver / gold) medal.
133. The woman at the shooting range does not wear (lipstick / rouge).
134. The woman in the revolving door likes to wear (slippers / sandals).
135. The woman is reaching into the fridge for (milk / juice).
136. The woman with the sunflower bakes good (pies / cakes).
137. The wood in the fireplace is mostly (spruce / pine).
138. These people are grieving because of a terrible (earthquake / tidal wave).
139. These skiers are friends from (college / university).
140. Tim wants to add a (sofa / coffee table) to his family room.
141. Tony paid for his video camera in (yen / lira).
142. Vivian likes to write and wants to be a (journalist / poet).
143. Whales can jump several (feet / yards).
144. When it is cold outside, Sharon wears a (hat / scarf).

Appendix 2.2
<u>Experiment 1 Instructions</u>

<u>Practice Training Phase:</u>
"You will be hearing a series of sentences via audio tape. These same sentences will also be presented visually, *except* that one word will be missing in the visual version. That is, a 'blank' or underline will appear on the page and will be the part of the sentence that you should try to memorize or remember for an upcoming memory test. In other words, you should be able to provide the completion or missing word when you receive the sentence cue at testing.

There is a further 'catch' to this experiment – namely, each sentence will actually have two *possible* completions across the experimental trials. You should try your hardest to remember *both* possible completions for the memory test a few minutes later. This may sound somewhat difficult but we are going to do some practice trials. Just do the best you can."

<u>Practice Test Phase:</u>
"You will now hear and see sentence cues and you should try to come up with *both* completions for each sentence cue. You will have a total of 10 seconds (until the next "beep" on the tape recorder) to come up with a response."

<u>Training Phase:</u>
"This next part is a lengthier version of what you did before. You will be hearing a number of sentences that will appear more than once. Each sentence will have two possible completions across the experimental trials. You should try your hardest to remember *both* completions for the memory test. This should take about 20 minutes in total."
[Following the training phase, the participant completed a demographic questionnaire.]

<u>Study Phase:</u>
"I know I told you that you would be getting a memory test on the sentences I just presented, but I wasn't exactly telling the truth. What I want you to do now is to try and put out of your mind the information that I just presented and *now* remember the sentences that I am now going to present. Each sentence will be presented once only and I will want you to remember the 'real' completion that goes with each sentence. This is somewhat difficult, so just do the best you can."

<u>Interpolated Tasks:</u>
"Before we get to the memory test for the last list of sentences, I have a couple of short tasks I would like you to perform."

<u>Task A: ***Digit Symbol Task***</u>
[The experimenter used the standardized instructions from the WAIS Digit Symbol Substitution Task.]

<u>Task B: ***Mill Hill Vocabulary Test***</u>
"I want you to select the lower-case word that is closest in meaning to the upper-case word in each case. I would like you to come up with a response for each item, *even if you have to guess*. There is no time limit."

Appendix 2.2 (continued)

Task C: ***Stroop Task***

"The next task involves your ability to name colors. There are two columns of rectangles on the sheet I am about to show you. What I want you to do is to name, as quickly and accurately as possible, the color of each rectangle. I will want you to start at the top of the first column and work your way to the bottom of the column. Then I'll want you to start naming colors at the top of the second column and work your way to the bottom...."

"For this next part of the task, you will see a number of words, presented in black ink. What I want you to do this time around is to *read the words* as quickly and accurately as possible...."

"I am now going to be presenting you with two columns of color words, as before. However, this time around, the words are going to be presented in different colors. What I want you to do this time is to ignore the meaning of the words and simply name the ink color as quickly and accurately as possible. So, in other words, I do not want you to read the words anymore. For example, if a word is presented in brown ink, I want you to say 'brown,' no matter what the word itself says. Once again, I will want you start at the beginning of the first column and end at the bottom of the second column. This is a difficult task, so just do the best you can."

Task D : ***Verbal Fluency Test***

"In the first part of this task, I want you to come up with as many English words as you can think of in 1 minute that start with a certain letter of the alphabet. The only restrictions on your responses are as follows: No proper nouns (i.e., names of an actual person, place, or thing, such as Cher, Toronto, or Volkswagen); no foreign words; no numbers (for example, 'nine,' 'nineteen,' and 'ninety'); and no variants on the same word. In other words, it is not permissible to use the same root word over again, but with different suffixes – for example, you can say 'run' if you're coming up with 'R' words, but you can't then say, 'runs,' 'running,' 'ran,' 'runner,' and 'runners'."

I now want you to come up with as many words as you can think of in 1 minute that begin with 'F'....'A'....'S'....Now I want you to come up with as many *animals* as you can think of in 1 minute."

Test Phase:

"Now it's time for the actual test phase. Once again you are going to be hearing sentence cues read aloud over the tape recorder. However, this time, you will not see anything written down in front of you. Instead, you will hear a sentence with a missing word, as indicated by the word 'blank'. The word 'blank' is your signal to fill in the missing word. ***In each case, I will want you to come up with the word that was presented during the last list of sentences that you saw and heard. However, if you can't remember anything from the last study list, then I want you to guess.*** In all cases, I will want you to try your hardest to come up with a response. I want you to say all your responses out loud and I will write them down.

For each sentence, you will have a total of 10 seconds to come up with a response. So, when the tape 'beeps,' I will want you to move on to the next sentence, even if you've not come up with a response. However, I want you to try your hardest to come up with a completion from the last study list, ***even if you have to guess.*** This is somewhat difficult, so just do the best you can."

Appendix 2.2 (continued)

Debriefing:
"In this study, we are interested in the effects of habit and context on people's memory for sentences. Habit was manipulated by having a training phase in which each sentence was presented four times each: Three times with one 'frequent' completion, and once with an 'infrequent' completion. We expect memory to be better when the subsequent study list provides the frequent completion (presented three times), rather than the less frequent completion (presented once only). Context was manipulated by presenting sentences with either the same voice, or different voices, at training and test. We expect memory to be better in the same-voice condition than in the different-voice condition. Both habitual responding and supportive contexts may serve to reduce age differences in memory."

Appendix 2.3
Experiment 2 Stimuli

85 Sentences from Bloom and Fischler (1980)

Sentence frame	Medium-cloze completion	Base rate	Low-cloze completion	Base rate
1. Barry wisely chose to pay the _____.	FINE	.21	TICKET	.04
2. Her dress was made of very fine _____.	MATERIAL	.21	CLOTH	.05
3. It's easy to get lost without a _____.	COMPASS	.21	GUIDE	.06
4. Coming in he took off his _____.	HAT	.21	SHIRT	.09
5. Wally wanted to buy a beer, but he was too _____.	BROKE	.21	DRUNK	.09
6. The sandwich wasn't very good without a slice of _____.	BREAD	.21	MEAT	.10
7. There's something grand about the _____.	OPERA	.22	CIRCUS	.04
8. They went to the rear of the long _____.	LINE	.22	HALL	.12
9. The difficult concept was beyond his _____.	COMPREHENSION	.22	IMAGINATION	.13
10. The earth is shaped like a _____.	SPHERE	.23	GLOBE	.03
11. The rider walked his beautiful _____.	BIKE	.23	MARE	.03
12. During class Jack had to borrow some _____.	MONEY	.23	PENCILS	.05
13. He was soothed by the gentle _____.	MUSIC	.23	BREEZE	.11
14. His ring fell into a hole in the _____.	SINK	.23	GROUND	.18
15. Ira turned on the radio and listened to the _____.	NEWS	.24	SONGS	.05
16. The airplane went into a _____.	DIVE	.24	TAILSPIN	.07
17. Diane slowly sank into the hot _____.	BATH	.24	WATER	.10
18. The actor was praised for being very _____.	GOOD	.24	TALENTED	.12
19. Don found that he had no spare _____.	CHANGE	.24	TIME	.13
20. Rushing out he forgot to take his _____.	COAT	.24	WALLET	.15
21. The truck that Bill drove crashed into the _____.	WALL	.24	POST	.15
22. I thought the sermon was very _____.	BORING	.24	LONG	.17
23. Mary couldn't leave the parlor until her hair was _____.	DONE	.25	FINISHED	.07
24. Few nations are now ruled by a _____.	DICTATOR	.25	MONARCH	.08
25. The Smiths had never visited that _____.	PLACE	.25	COUNTRY	.09
26. His view was blocked by the music _____.	STAND	.25	BOX	.10
27. The storm made the air damp and _____.	COLD	.25	HUMID	.14
28. John poured himself a glass of _____.	WINE	.26	BEER	.03
29. The pain she felt was all in her _____.	MIND	.26	STOMACH	.03
30. The old house was built entirely of _____.	BRICK	.26	STONE	.09
31. The surface of the water was nice and _____.	SMOOTH	.26	CLEAR	.09
32. The kind old man asked us to _____.	STAY	.26	LEAVE	.10
33. The long test left the class _____.	EXHAUSTED	.26	DRAINED	.12
34. Some of the ashes dropped on the _____.	RUG	.26	CARPET	.15
35. Sometimes success is simply a matter of _____.	LUCK	.26	WORK	.15
36. Jim hit his horse with a _____.	STICK	.27	CROP	.04
37. There are times when life seems _____.	DULL	.27	HOPELESS	.07
38. Getting the shot really didn't _____.	HELP	.27	MATTER	.09
39. My sister bought tickets for the opening _____.	GAME	.27	PLAY	.13
40. He disliked having to commute to the _____.	CITY	.27	OFFICE	.16
41. The cup of tea felt very _____.	HOT	.28	SOOTHING	.04
42. They liked to sleep out under the _____.	TREES	.28	MOON	.04
43. The child went ever higher on the _____.	SWING	.28	MOUNTAIN	.08
44. You could count on Dale on being late for _____.	CLASS	.28	DINNER	.13
45. Being stood up made Paul _____.	ANGRY	.29	UPSET	.05
46. Stan slowed down going around the _____.	CURVE	.29	TURN	.12
47. No one wanted to accuse him of _____.	STEALING	.29	MURDER	.18
48. The death of his dog was a great _____.	SHOCK	.29	LOSS	.18

Appendix 2.3 (continued)

Sentence frame	Medium-cloze completion	Base rate	Low-cloze completion	Base rate
49. They rested under a tree in the _____.	PARK	.30	YARD	.09
50. The young boy was granted a small _____.	ALLOWANCE	.30	FAVOR	.10
51. He drove the nail into the _____.	WOOD	.32	PLANK	.03
52. Dan caught the ball with his _____.	GLOVE	.32	MITT	.08
53. Every month Rick had to clean his _____.	ROOM	.32	CLOSET	.09
54. He wondered if the storm would be _____.	OVER	.32	BAD	.19
55. The car stalled because the engine failed to _____.	START	.32	RUN	.19
56. Jean had learned the passage by _____.	HEART	.33	ROTE	.03
57. He had to fill his truck with _____.	GAS	.33	CEMENT	.04
58. You can't buy anything for a _____.	DIME	.33	NICKEL	.10
59. She cleaned the dirt from her _____.	SHOE	33	DRESS	.11
60. The ruby was so big, it looked like a _____.	ROCK	.33	CHERRY	.13
61. The surgeon tried vainly to save his _____.	LIFE	.34	SON	.02
62. The fire was small, and there was no reason to _____.	WORRY	.34	FEAR	.03
63. Their money was divided by the _____.	BANK	.34	LAWYER	.03
64. Harriet sang while my brother played the _____.	GUITAR	.34	FLUTE	.04
65. Larry chose not to join the _____.	CLUB	.34	TEAM	.08
66. The person who caught the thief deserves our _____.	THANKS	.34	REWARD	.09
67. Without food a man would die in several _____.	WEEKS	.37	MONTHS	.03
68. We used to get company every _____.	NIGHT	.37	WEEKEND	.05
69. The final score of the game was _____.	TIED	.37	CLOSE	.09
70. A direct attack failed, so they changed the _____.	STRATEGY	.37	TACTICS	.13
71. Even their friends were left in the _____.	DARK	.37	RAIN	.13
72. Helen reached up to dust the _____.	SHELF	.37	CABINET	.18
73. Rita walked down the shaky _____.	LADDER	.39	STAIRCASE	.04
74. She dropped a glass and woke up the _____.	BABY	.39	NEIGHBOR	.13
75. Joan showed her friend a new card _____.	TRICK	.40	TABLE	.06
76. To find the body, they had to drain the _____.	LAKE	.40	POND	.13
77. Jill looked back through the open _____.	WINDOW	.41	GATE	.02
78. Bob would often sleep during his lunch _____.	BREAK	.41	PERIOD	.05
79. Suzy liked to play with her toy _____.	DOLLS	.41	ANIMALS	.05
80. The wooded lake made a pretty _____.	SCENE	.41	SITE	.15
81. Plants will not grow in dry _____.	SOIL	.42	SAND	.03
82. The birds in the yard ate every last _____.	CRUMB	.42	BIT	.11
83. The senator was startled by the sudden pain in his _____.	CHEST	.42	SIDE	.13
84. The hunter shot and killed a large _____.	DEER	.43	MOOSE	.05
85. The sail got loose, so they tightened the _____.	ROPES	.44	MAST	.10

59 Sentences from Brown's (Unpublished) Norms

Sentence frame	Medium-cloze completion	Base rate	Low-cloze completion	Base rate
1. The man in the hotel room is working on a _____.	NOVEL	.22	BOOK	.11
2. Jenny wants to paint her bedroom _____.	BLUE	.23	GREEN	.07
3. Rescuers at the lake are trying to find a lost _____.	CHILD	.23	PUPPY	.07
4. The book on the patio table has _____.	PICTURES	.23	PHOTOGRAPHS	.07
5. The electricity was generated by _____ power.	NUCLEAR	.23	STEAM	.07
6. The forest has deposits of _____.	COAL	.23	CLAY	.07
7. The mother was cleaning the _____.	KITCHEN	.23	BATHROOM	.07
8. The woman in the revolving door likes to wear _____.	RED	.23	FUR	.07
9. Julie served her cake with some _____.	TEA	.23	COFFEE	.10
10. Michelle likes her bedsheets to be made out of _____.	SILK	.23	FLANNEL	.10

Appendix 2.3 (continued)

Sentence frame	Medium-cloze completion	Base rate	Low-cloze completion	Base rate
11. The dancing young people are at a _____.	PARTY	.23	BAR	.10
12 The daredevil is going to jump off the _____.	BRIDGE	.23	RAMP	.10
13. The lady with the bad headache has trouble _____.	CONCENTRATING	.23	THINKING	.10
14. The bicycling family is going to see their _____.	FRIENDS	.23	RELATIVES	.13
15. The goalie is particularly good at _____.	SAVING	.23	CATCHING	.13
16. The astronomer just discovered a new _____.	PLANET	.27	GALAXY	.07
17. The golfer has already lost _____ balls.	THREE	.27	FOUR	.07
18. Tim wants to add a _____ to his living room.	TELEVISION	.27	STEREO	.09
19. The bike had broken _____.	WHEELS	.27	TIRES	.10
20. The computer genius can do _____.	PROGRAMMING	.27	MATH	.10
21. The new model car is being tested for its _____.	SAFETY	.27	PERFORMANCE	.10
22. The tourists on the boat wanted to see the _____.	WHALES	.27	STATUE	.10
23. In the businessman's suitcase, there are many _____.	PAPERS	.27	FILES	.13
24. The hotel's swimming pool has a _____.	DIVING BOARD	.27	HEATER	.13
25. The kids like to eat _____ in the playroom.	COOKIES	.27	CANDY	.13
26. The little girl is moving and has packed all of her _____.	TOYS	.27	CLOTHES	.13
27. The running boy is playing _____.	TAG	.27	BALL	.13
28. The used car was covered in _____.	MUD	.27	DUST	.13
29. These people are grieving because of the terrible _____.	ACCIDENT	.27	TRAGEDY	.13
30. The canoeists first met at a _____.	CAMP	.28	RACE	.07
31. The desert area has some _____.	CACTI	.30	VEGETATION	.07
32. The old men are playing _____.	CHESS	.30	POKER	.07
33. The truck in the parade usually transports _____.	FLOATS	.30	CLOWNS	.07
34. A man cut his _____ while shaving.	CHIN	.30	CHEEK	.10
35. The computer's _____ is not working properly.	HARD DRIVE	.30	MEMORY	.10
36. The lady with the back pains had surgery on her _____.	SPINE	.30	DISK	.10
37. The streetcar is headed _____.	SOUTH	.30	DOWNTOWN	.10
38. Angie uses the stamp on her desk to stamp her _____.	LETTERS	.30	MAIL	.13
39. Cellists often complain about stiff _____.	NECKS	.30	FINGERS	.13
40. The fields have some _____.	FLOWERS	.30	GRASS	.13
41. The people in the car are going _____.	HOME	.30	AWAY	.13
42. The boat is smuggling _____.	DRUGS	.33	DOPE	.07
43. The father was fixing the _____.	CAR	.33	STOVE	.07
44. The man purchasing the CD was once bitten by a _____.	DOG	.33	SPIDER	.07
45. After the game, the coach bought the kids _____.	ICE CREAM	.33	PIZZA	.10
46. The fighter plane is armed with _____.	MISSILES	.33	WEAPONS	.10
47. The waterfall flows into a _____.	RIVER	.33	STREAM	.10
48. The horse's owner lives in a _____.	HOUSE	.33	MANSION	.13
49. The woman at the shooting range does not wear _____.	GLASSES	.33	GOGGLES	.13
50. The estate is owned by a _____.	MILLIONAIRE	.36	BILLIONAIRE	.07
51. Jane puts pictures in her scrapbook with _____.	TAPE	.37	CARE	.07
52. The baby was born in the _____.	HOSPITAL	.37	BEDROOM	.07
53. One tool needed to make the cabinets is a _____.	HAMMER	.37	CHISEL	.13
54. David wanted some _____ for breakfast.	CEREAL	.40	EGGS	.13
55. The old coat has a lot of _____.	HOLES	.43	STAINS	.07
56. Next, the chess player will move his _____.	PAWN	.43	QUEEN	.10
57. On the stove, Jeff is heating up some _____.	SOUP	.43	SPAGHETTI	.10
58. The insects in a flower garden include _____.	BEES	.43	ANTS	.10
59. Kids like to pitch _____ in the water fountain.	PENNIES	.43	COINS	.17

Appendix 2.4
<u>Experiment 2 Instructions</u>

<u>Practice Phase, Training Phase, Study Phase, and Interpolated Tasks A-D:</u>
[The instructions for these experimental phases are provided in Appendix 2.2.]

<u>Interpolated Task E:</u> *<u>Trail Making Task</u>*
[The experimenter used the standardized instructions for Reitan's (1955, 1958) Trail Making Task, Parts A and B.]

<u>Test Phase:</u>
"Now it's time for the actual test phase. Once again you are going to be hearing sentence frames read aloud over the tape recorder. In addition, you'll be seeing the same sentence frames in the binder. The appearance of the underline on the page, and the word 'blank' in the recording, should prompt you to come up with the missing word. ***In each case, I'll want you to come up with the new word that was presented during the <u>last</u> list of sentences that you saw and heard. However, if you can't remember anything from the last study list, then I want you to guess.*** In all cases, I will want you to try your hardest to come up with a response *from the last study list.* I want you to say all your responses out loud and I will write them down.

There is an additional 'catch' here: Although most of the sentences are from the *last* list that you saw and heard, there will occasionally be a few sentences that are brand new. Don't let this throw you: Even if you don't remember the sentence, I'll want you to provide your best guess as to how the sentence would be completed.

For each sentence, you will have a total of 10 seconds to come up with a response. In each case, you should always try to come up with the completion for the sentence from the *last* study list. However, after 8 seconds, you will hear a 'beep' on the tape recording. This 'beep' will be your signal that your time limit is almost up and you should *immediately* provide your best guess as to how the sentence would be completed. In other words, you should *always* come up with a response, no matter what.

Then, when the full 10 seconds is up, you will hear a 'double beep,' which will be your signal that the particular test trial is over. I will want you to move on to the next sentence, even if you've not come up with a response. However, I will want you to *always* try and come up with a completion from the last study list, ***even if you have to guess.*** It is important that you try to guess, even if the sentence is not familiar to you. This is somewhat difficult, so just do the best you can."

<u>Debriefing:</u>
"In this study, we are interested in the effects of habit and prior knowledge on people's memory for sentences. Habit was manipulated by having a training phase in which each sentence was presented four times each: Three times with one completion, and one time with another completion. We expect memory to be better when the subsequent study list provides the thrice-presented completion, rather than the once-presented completion. In addition, we varied the extent to which prior knowledge can assist in memory for the sentences. For example, if I gave you the sentence frame, 'He had to fill his truck with ___,' you would be more likely to think of 'gas' than 'cement.' This is the case, because in real life, you are more likely to fill a truck with gas than cement. We expect memory to be better when information is repeated and when people can rely on prior knowledge. Both habitual responding and real-world knowledge may serve to reduce age differences in memory."

Appendix 3.1
Stimuli for Experiments 3, 4, and 5

48 Cues Used in Experiments 3, 4, and 5 *(from Reingold's [1995] Experiment 3)*

The two possible versions of each cue are presented with their respective solutions. The letter to be deleted is in bold. In the four columns on the left, cues are presented in which the leftmost underlined letter must be deleted to form a word. In the four columns on the right, cues are presented in which the rightmost underlined letter must be deleted to form a word.

Left Letter Deletions				Right Letter Deletions			
Cue	**Word**	**Cue**	**Word**	**Cue**	**Word**	**Cue**	**Word**
A**P**RSON	ARSON	APR**S**ON	APRON	BR**E**AID	BRAID	BREA**I**D	BREAD
A**U**DMIT	ADMIT	AUD**M**IT	AUDIT	CH**E**AMP	CHAMP	CHEA**M**P	CHEAP
B**E**ATCH	BATCH	BEA**T**CH	BEACH	FA**B**LSE	FALSE	FABL**S**E	FABLE
B**I**ROTH	BROTH	BIR**O**TH	BIRTH	GR**O**AIN	GRAIN	GROA**I**N	GROAN
B**R**ASIN	BASIN	BRA**S**IN	BRAIN	NO**R**TCH	NOTCH	NORT**C**H	NORTH
B**R**OSOM	BOSOM	BRO**S**OM	BROOM	PE**D**ARL	PEARL	PEDA**R**L	PEDAL
C**A**URSE	CURSE	CAU**R**SE	CAUSE	PL**E**AID	PLAID	PLEA**I**D	PLEAD
C**H**ABIN	CABIN	CHA**B**IN	CHAIN	QU**I**EST	QUEST	QUIE**S**T	QUIET
C**O**ATCH	CATCH	COA**T**CH	COACH	RE**P**ADY	READY	REPA**D**Y	REPAY
D**R**IVER	DIVER	DRI**V**ER	DRIER	RO**B**OST	ROOST	ROBO**S**T	ROBOT
L**E**APSE	LAPSE	LEA**P**SE	LEASE	SA**D**LTY	SALTY	SADL**T**Y	SADLY
P**A**URSE	PURSE	PAU**R**SE	PAUSE	SC**R**ALP	SCALP	SCRA**L**P	SCRAP
P**L**AINT	PAINT	PLA**I**NT	PLANT	SH**O**UNT	SHUNT	SHOU**N**T	SHOUT
P**L**ASTE	PASTE	PLA**S**TE	PLATE	SP**A**REE	SPREE	SPAR**E**E	SPARE
P**O**UNCH	PUNCH	POU**N**CH	POUCH	SP**E**ARK	SPARK	SPEA**R**K	SPEAK
P**R**IECE	PIECE	PRI**E**CE	PRICE	SP**O**ILL	SPILL	SPOI**L**L	SPOIL
P**R**OISE	POISE	PRO**I**SE	PROSE	ST**E**ALL	STALL	STEA**L**L	STEAL
R**A**INSE	RINSE	RAI**N**SE	RAISE	ST**O**RAY	STRAY	STOR**A**Y	STORY
R**E**ANCH	RANCH	REA**N**CH	REACH	ST**R**AMP	STAMP	STRA**M**P	STRAP
S**H**ANDY	SANDY	SHA**N**DY	SHADY	SU**R**LKY	SULKY	SURL**K**Y	SURLY
S**H**EWER	SEWER	SHE**W**ER	SHEER	TH**R**EME	THEME	THRE**M**E	THREE
S**L**AINT	SAINT	SLA**I**NT	SLANT	TR**O**UST	TRUST	TROU**S**T	TROUT
S**P**AUCE	SAUCE	SPA**U**CE	SPACE	VA**U**LET	VALET	VAUL**E**T	VAULT
S**T**OLVE	SOLVE	STO**L**VE	STOVE	WI**D**TCH	WITCH	WIDT**C**H	WIDTH

24 Additional Cues Generated for Use in Experiments 4 and 5

Left Letter Deletions				Right Letter Deletions			
Cue	**Word**	**Cue**	**Word**	**Cue**	**Word**	**Cue**	**Word**
B**L**OUND	BOUND	BLOU**N**D	BLOND	BA**S**THE	BATHE	BAST**H**E	BASTE
C**L**EDAR	CEDAR	CLE**D**AR	CLEAR	BL**E**ANK	BLANK	BLEA**N**K	BLEAK
C**U**RAVE	CRAVE	CUR**A**VE	CURVE	BU**D**GLE	BUGLE	BUDG**L**E	BUDGE
F**A**ILTH	FILTH	FAI**L**TH	FAITH	CH**E**ART	CHART	CHEA**R**T	CHEAT
F**I**ROST	FROST	FIR**O**ST	FIRST	DI**A**RTY	DIRTY	DIAR**T**Y	DIARY
G**L**IANT	GIANT	GLI**A**NT	GLINT	DR**I**EAD	DREAD	DRIE**A**D	DRIED
O**U**THER	OTHER	OUT**H**ER	OUTER	GR**E**ANT	GRANT	GREA**N**T	GREAT
P**E**ATCH	PATCH	PEA**T**CH	PEACH	HA**I**RDY	HARDY	HAIR**D**Y	HAIRY
S**C**ABLE	SABLE	SCA**B**LE	SCALE	QU**A**ILL	QUILL	QUAI**L**L	QUAIL
S**L**OUTH	SOUTH	SLO**U**TH	SLOTH	SN**E**ACK	SNACK	SNEA**C**K	SNEAK
W**O**RATH	WRATH	WOR**A**TH	WORTH	ST**E**ALK	STALK	STEA**L**K	STEAK
W**O**RIST	WRIST	WOR**I**ST	WORST	ST**O**UNT	STUNT	STOU**N**T	STOUT

Appendix 3.1 (continued)

24 Additional Cues Generated for Use in Experiment 5

Left Letter Deletions				Right Letter Deletions			
Cue	**Word**	**Cue**	**Word**	**Cue**	**Word**	**Cue**	**Word**
ANGILE	AGILE	ANGILE	ANGLE	CREACK	CRACK	CREACK	CREAK
ENTHER	ETHER	ENTHER	ENTER	DONUBT	DOUBT	DONUBT	DONUT
FAIERY	FIERY	FAIERY	FAIRY	FIRELD	FIELD	FIRELD	FIRED
FREWER	FEWER	FREWER	FREER	FORAMY	FOAMY	FORAMY	FORAY
GOUARD	GUARD	GOUARD	GOURD	FRAILL	FRILL	FRAILL	FRAIL
HEANDY	HANDY	HEANDY	HEADY	GROUNT	GRUNT	GROUNT	GROUT
MEADLY	MADLY	MEADLY	MEALY	HORNEY	HONEY	HORNEY	HORNY
SHEVEN	SEVEN	SHEVEN	SHEEN	INEPUT	INPUT	INEPUT	INEPT
SLINCE	SINCE	SLINCE	SLICE	PIANTO	PINTO	PIANTO	PIANO
SNEVER	SEVER	SNEVER	SNEER	STARUT	STRUT	STARUT	START
THEASE	TEASE	THEASE	THESE	STORUM	STRUM	STORUM	STORM
WEAIVE	WAIVE	WEAIVE	WEAVE	TRAILL	TRILL	TRAILL	TRAIL

Appendix 3.2

Instructions for Experiments 3 and 4

Letter Deletion Instructions:

"We're now going to do some word puzzles on the computer. There will be several trials. In each trial, a string of 6 letters will be presented, which is a nonword. Your task is to decide which letter can be removed from the nonword to make a meaningful word from the remaining letters.

More specifically, in each trial, two letters of the nonword will be surrounded by boxes. If one of the two 'boxed' letters is removed, it will become a 5-letter word. You are asked to decide which of the two boxed letters can be removed to leave a meaningful word. If the boxed letter on the left is to be removed, type the key marked 'L' for 'left.' If the boxed letter on the right is to be removed, type 'R' for 'right'."

[Two stickers labeled 'L' and 'R' were affixed to the '1' and '2' keys of the numeric keypad of the computer keyboard, respectively.]

"I want you to position your dominant hand *(i.e., right, if you're right-handed)* over the 'L' and 'R' keys *(i.e., '1' and '2').* The base of your palm should be resting comfortably on the table between trials. I will want you to press the 'L' and 'R' keys with the index and middle fingers of your dominant hand.

Make sure your other hand is positioned near the 'ENTER' key *(on the numeric keypad if you're right-handed, and on the alphanumeric lay-out if you're left-handed).* This hand should also be resting on the table.

You will be pressing the 'ENTER' key with your non-dominant hand to start each trial. A fixation cross will be your signal to start each trial, whenever you're ready to do so.

One more thing, and this is very important: When pressing the 'L' and 'R' keys, ***try to respond as quickly as you can without making errors***. Everyone makes mistakes, so don't worry about it if you do. But I will ask you to be as fast and as accurate as possible in hitting the 'L' and 'R' keys. So both speed and accuracy are equally important.

On the other hand, you can take your time when pressing the 'ENTER' key to start each trial. But please make your selection of the 'L' or 'R' key as quickly and as accurately as you can."

[Participants were then shown two paper examples, one at a time, and were asked for each example, "What would you do if this came up on the computer screen?" *Both examples were in a large, 128-point Times New Roman font. Experiment 3 participants received the following cues as examples:* "c [h] [e] a r t" *and* "w o r [i] [s] t". *However, these exemplars became critical stimuli in Experiment 4. Therefore, Experiment 4 participants received two new cues as examples:* "a n g [i] [l] e" *and* "c [r] [e] a c k"*.]*

"Now we're going to try this on the computer. We'll start off with some practice trials – the first dozen trials or so won't really count for anything, except to give you some practice on the task. So just do the best you can. But after a dozen trials, the computer will switch automatically between the practice and actual trials, and there won't be a distinct break between the two types of trials."

Appendix 3.2 (continued)

Debriefing for Experiment 3:

"We are interested in age differences in the ability to solve word puzzles. More specifically, we're interested in age differences in susceptibility to interference. Interference was manipulated by having some word puzzles come up twice, but with different letters surrounded by boxes from time one to time two. *[Participants were then shown two debriefing examples on sheets of paper.]*

For example, if 'c h e a r t' comes up the first time, with the 'h' and 'e' surrounded by boxes, the person deletes the 'e' to produce the word 'chart.' However, if this nonword comes up later, with different letters surrounded by boxes *(e.g.,* 'c h e a r t,' *with boxes around the 'a' and 'r')* then the initially appropriate solution becomes inappropriate. In other words, the 'e' is no longer a candidate for solution, and instead the 'r' must be deleted to produce the word, 'cheat.' Therefore, the person's performance is interfered with, and he or she becomes slower and/or less accurate the second time around.

We expect that older adults might show greater interference in this task than younger adults. In other words, older adults might show greater slowing or make more errors than younger adults, when different letters are surrounded from the first to the second presentation of a stimulus. Thank you for your participation."

Debriefing for Experiment 4:

"We are interested in age differences in the ability to solve word puzzles. In other words, we are comparing older and younger adults on this task.

You may have noticed that some of these puzzles came up more than once. In most cases, the puzzle appeared in the same exact form from one repetition to the next. This type of repetition causes people to become faster and/or more accurate across repetitions.

This finding is called repetition priming, or just plain 'priming.' It's basically been shown for the last ten years or so that there tend to be no age differences in this type of priming, so this is nothing new.

However, sometimes the nonwords came up with different letters surrounded by boxes from one time to the next. In this case, it becomes harder to solve the puzzles across repetitions. *[Participants were then shown two debriefing examples on sheets of paper.]*

For example, if 'c h e a r t' comes up the first time, with the 'h' and 'e' surrounded by boxes, the person deletes the 'e' to produce the word 'chart.' However, if this nonword comes up later, with different letters surrounded by boxes *(e.g.,* 'c h e a r t,' *with boxes around the 'a' and 'r')* then the initially appropriate solution becomes inappropriate later on.

In other words, the 'e' is no longer a candidate for solution, and instead the 'r' must be deleted to produce the word 'cheat.' Therefore, the person's performance is interfered with, and he or she becomes slower and/or less accurate the second time around. This finding is, not surprisingly, called interference priming.

We expect that older adults might show greater interference in this task than younger adults. In other words, older adults might show greater slowing or make more errors than younger adults, when different letters are surrounded from the first to the second presentations of a stimulus.

We are making this prediction because some memory researchers believe that aging makes people more susceptible to interference. Thank you for your participation."

Appendix 3.3
Experiment 5 Instructions

Instructions for Participants Receiving the Boxed Task First:

"In this experiment, you will be getting a number of trials. In each trial you will be solving a word puzzle. More specifically, in each trial, a string of letters will be presented which is a nonword.

Instructions for First (Boxed) Task
In the first set of trials, a string of letters will appear on each trial. Two letters in the string will be surrounded by boxes. You are asked to decide which boxed letter can be removed to leave a meaningful word.

However, you won't actually be removing any letters from the nonword. Instead, what you'll be doing is telling the computer whether the first or second boxed letter can be removed from the nonword to form a word.

If the first or left boxed letter is to be removed, type '1' for first. If the second or right boxed letter is to be removed, type '2' for second.

Then, after you've pressed the '1' or '2' key to indicate whether the first or second boxed letter should be removed, the computer screen will be erased and a fixation cross will come up in the center of the screen. The cross will be your signal to press the 'ENTER' key to start the next trial as soon as you're ready.

I will want you to position your dominant hand *(i.e., right, if you're right-handed)* over the '1' and '2' keys. The base of your palm should be resting comfortably on the table between trials. I will want you to press the '1' and '2' keys with the index and middle fingers of your dominant hand.

Make sure that your other hand is positioned near the 'ENTER' key *(on the numeric keypad if using your right hand, and on the alphanumeric lay-out if using your left hand)*. This hand should also be resting on the table.

So, once again, as a reminder, each trial will start off with a fixation cross in the center of the computer screen. The cross will be your signal to press the "ENTER" key as soon as you're ready to bring up the next nonword onto the computer screen. Then I will want you to indicate on the keyboard whether the first or the second boxed letter should be removed in order to produce a word. Then the fixation cross will appear again and it will be time for the next trial.

One more thing – ***try to respond as quickly as you can without making errors***. Everyone makes mistakes, so don't worry about it if you do. But you should try your hardest to press the '1' and '2' keys as quickly and and as accurately as possible.

So, the first several trials will be practice trials, so you can get a better feel for what this is like. Have fun and good luck. *[Participants were then shown two paper examples of boxed stimuli:* 't r e a s h' *and* 's m i r l e.']...."

Appendix 3.3 (continued)

Instructions for Second* (Working Memory) *Task
"....Now that you've completed the first set of trials, we're going to try a slightly different task. Once again, you'll decide which letter can be removed from a nonword to produce a word.

However, this time around, you won't be selecting one of two boxed letters for removal. In fact, there will be no letters surrounded by boxes at all.

Instead, at the start of each trial, I will be saying aloud the two letters that I will want you to try and delete for that trial. So, in other words, before we see the nonword, I will always say two letters aloud (for example, 'T,L') that I will want you to repeat aloud and commit to memory.

Then it will be your job to call up the nonword on the computer screen. You'll do this by hitting the 'ENTER' key. Then you'll see the nonword and have to decide which of the two letters I gave you can be removed to form a word.

For example, if I said the letters 'T,L' and you think the 'T' can be removed to form a word, then you would type '1' for 'first,' since this was the first letter I presented for that trial. However, if you think the 'L' can be removed to form a word, then you would type '2' for 'second,' since this was the second letter I presented for that trial.

Then, after you've pressed the '1' or '2' key to indicate whether the first or second letter should be removed, the computer screen will be erased and a fixation cross will come up in the center of the screen.

The cross will be your signal to listen for the next two letters that I will present in the upcoming trial. You will then repeat back to me the letters that I just presented, in the same order that I gave them to you, and then press the 'ENTER' key to bring up the next stimulus as soon as you're ready.

But don't press the 'ENTER' key until after you've memorized the two letters for that trial. Then I will want you to indicate on the keyboard whether the first or the second letter I gave you should be removed in order to produce a word. Then the fixation cross will appear again and it will be time for the next trial.

As before, I'll want you to make your selection of either the '1' or '2' key ***as quickly and as accurately as you can***.

I'll also want you to hit the '1' and '2' keys with the index and middle fingers of your dominant hand. And I'll want you to use your non-dominant hand to hit the 'ENTER' key when you're ready to bring up the nonword on the computer screen. We'll start off with several practice trials, as before. *[Participants were then presented with two paper examples –* 'stlair' *and* 'movuie' *– which were prefaced with 'T,L' and 'U,I,' respectively.]* ...."

Appendix 3.3 (continued)

Instructions for Participants Receiving the Working Memory Task First:

"In this experiment, you will be getting a number of trials. In each trial you will be solving a word puzzle. More specifically, in each trial, a string of letters will be presented which is a nonword.

Instructions for First* (Working Memory) *Task

In the first set of trials, you will be asked to decide which letter can be removed from the letter string to leave a meaningful word.

However, there's a bit of a catch here: At the start of each trial, I will be saying aloud the two letters that I will want you to try and delete for that trial. So, in other words, before we see the nonword, I will always say two letters aloud (for example, 'T,L') that I will want you to repeat aloud and commit to memory.

Then, it will be your job to call up the nonword on the computer screen. You'll do this by hitting the 'ENTER' key. Then you'll see the nonword and have to decide which of the two letters I gave you can be removed to form a word.

However, you won't actually be removing any letters from the nonword. Instead, what you'll be doing is telling the computer whether the first or second letter I mentioned can be removed from the nonword to form a word.

For example, if I said the letters 'T,L' and you think the 'T' can be removed to form a word, then you would type '1' for 'first,' since this was the first letter I presented for that trial. However, if you think the 'L' can be removed to form a word, then you would type '2' for 'second,' since this was the second letter I presented for that trial.

Then, after you've pressed the '1' or '2' key to indicate whether the first or second letter should be removed, the computer screen will be erased and a fixation cross will come up in the center of the screen. The cross will be your signal to listen for the next two letters that I will present in the next trial. You will then repeat back to me the letters that I just presented, in the same order that I gave them to you, and then press the 'ENTER' key to bring up the next stimulus as soon as you're ready.

Another thing is that I will want you to do is to position your dominant hand *(i.e., right, if you're right-handed)* over the '1' and '2' keys. The base of your palm should be resting comfortably on the table between trials. I will want you to press the '1' and '2' keys with the index and middle fingers of your dominant hand.

Make sure that your other hand is positioned near the 'ENTER' key *(on the numeric keypad if using your right hand, and on the alphanumeric lay-out if using your left hand)*. This hand should also be resting on the table.

So, once again, as a reminder, each trial will start off with a fixation cross in the center of the computer screen. The cross will be your signal to hear me read off the two letters for the upcoming trial. I'll want you to repeat the letters back to me in the same order in which I said them and then commit them to memory...

Appendix 3.3 (continued)

...Then, once you've said the letters aloud, I will want you to press the 'ENTER' key to bring up the corresponding nonword onto the computer screen. But don't press the 'ENTER' key until after you've committed the letters to memory for that trial. Then I will want you to indicate on the keyboard whether the first or the second letter I gave you should be removed in order to produce a word. Then the fixation cross will appear again and it will be time for the next trial.

One more thing – try to respond as quickly as you can without making errors. Everyone makes mistakes, so don't worry about it if you do. But you should try your hardest to press the '1' and '2' keys ***as quickly and and as accurately as possible***.

So, the first several trials will be practice trials, so you can get a better feel for what this is like. Have fun and good luck. *[Participants were then presented with two paper examples –* 'stlair' *and* 'movuie' – *which were prefaced with 'T,L' and 'U,I,' respectively.]...."*

Instructions for Second (Boxed) Task

"....Now that you've completed the first set of trials, we're going to try a slightly different task. Once again, you'll decide which letter can be removed from a nonword to produce a word.

However, this time around, you won't be memorizing a pair of letters to try and delete from the nonword. Instead, you'll see boxes around two letters of the nonword and you'll have to decide which of these two boxed letters can be taken out to form a word.

Before each trial begins there will be a fixation cross or small cross on the center of the computer screen. The cross will be your signal to press 'ENTER,' whenever you're ready to start the trial. By pressing 'ENTER,' you'll bring up the nonword with the two boxed letters onto the screen.

Then I'll want you to make your selection of whether the first or second boxed letter can be removed to form a word by hitting either the '1' key (first letter) or the '2' key (second letter).

As before, I'll want you to make your selection of either the '1' or '2' key ***as quickly and as accurately as you can***.

I'll also want you to hit the '1' and '2' keys with the index and middle fingers of your dominant hand. And I'll want you to use your non-dominant hand to hit the 'ENTER' key when you're ready to bring up the nonword on the computer screen. We'll start off with several practice trials, as before. *[Participants were then shown two paper examples of boxed stimuli:* 't r e a s h' *and* 's m i r l e.'*]...."*

Appendix 3.3 (continued)

Debriefing for Experiment 5:
"We are interested in age differences in the ability to solve word puzzles. More specifically, we're interested in age differences in susceptibility to interference. Interference was manipulated by having some word puzzles come up twice, but with different solutions from time one to time two. *[Participants were then shown two debriefing examples on sheets of paper.]*

For example, if 'c h e a r t' comes up the first time, with the 'h' and 'e' surrounded by boxes, the person deletes the 'e' to produce the word 'chart.' However, if this nonword comes up later, with different letters surrounded by boxes *(e.g.,* 'c h e a r t,' *with boxes around the 'a' and 'r')* then the initially appropriate solution becomes inappropriate. In other words, the 'e' is no longer a candidate for solution, and instead the 'r' must be deleted to produce the word, 'cheat.' Therefore, the person's performance is interfered with, and he or she becomes slower and/or less accurate the second time around.

We expect that older adults might show greater interference in this task than younger adults. In other words, older adults might show greater slowing or make more errors than younger adults, when the same puzzle comes up twice, but with different solutions each time.

In particular, we predict that older adults will show the most interference when the word puzzles appear without boxes. In other words, when there are no boxes to show the candidate letters for deletion, the person must hold these letters in their mind while solving the word puzzle. This type of memory, called working memory, involves actively manipulating information held in one's mind and is thought to be especially difficult for older adults.

In summary, we predict that age differences in interference will be larger in the no-boxes condition than in the boxes condition. Thank you for your participation."