2004

# Age-Related Differences in Memory Accuracy and Memory Monitoring: Relationship to Executive Processes

Matthew Gerard Rhodes

THE FLORIDA STATE UNIVERSITY

COLLEGE OF ARTS AND SCIENCES

# AGE-RELATED DIFFERENCES IN MEMORY ACCURACY AND MEMORY MONITORING: RELATIONSHIP TO EXECUTIVE PROCESSES

By

MATTHEW GERARD RHODES

A Dissertation submitted to the
Department of Psychology
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Degree Awarded:
Fall Semester, 2004

The members of the Committee approve the dissertation of Matthew G. Rhodes defended on 13 July 2004.

Colleen M. Kelley
Professor Directing Dissertation

Michelle Bourgeois
Outside Committee Member

Neil Charness
Committee Member

Katinka Dijkstra
Committee Member

E. Ashby Plant
Committee Member

The Office of Graduate Studies has verified and approved the above named committee members.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The current study examined the neuropsychological correlates of memory accuracy in older and younger adults. Participants were tested in a memory monitoring paradigm developed by Koriat and Goldsmith (1996), which permits separate assessments of the accuracy of responses generated during retrieval and the accuracy of monitoring those responses. Participants were also administered a battery of tests designed to measure executive functioning and speed of processing. Results indicated that both age and executive measures were predictive of accuracy, while speed of processing measures accounted for little of the variability in accuracy. In addition, executive measures explained a moderate amount of the variability in monitoring. Mediational analyses demonstrated that a large portion of the effect of executive function measures on accuracy was mediated by memory monitoring. These data suggest that individual differences in executive function are important in memory accuracy.

# INTRODUCTION

Aging is associated with a general decline in memory that may be chiefly evident for older adults who forget names, appointments, or neglect to take medication. The prevalence of such errors of omission is well documented and most acute on tests of recall (e.g., Craik & McDowd, 1987). However, equally troubling are errors of commission, characterized by elevated levels of false memories reported by older adults (e.g., Norman & Schacter, 1997). The current study will attempt to examine the locus of these errors of commission and, specifically, whether they may be related to deficits in executive functions (West, 1996), such as monitoring the contents of memory, that decline with age.

The last decade has witnessed a considerable increase in research investigating the accuracy of memory (see Koriat, Goldsmith, & Pansky, 2001; Roediger, 1996, for reviews) with a particular focus on errors of commission. For example, a number of researchers have investigated false memories in the Deese-Roediger-McDermott (Deese, 1959; Roediger & McDermott, 1995), or DRM paradigm. In the DRM paradigm, participants study a list of words that are semantically related to a central theme word (referred to as the *critical lure*) that is not presented. For example, participants might study *bed*, *rest*, *awake*, *tired*, *dream*, *wake*, *snooze*, *blanket*, *doze*, *slumber*, *snore*, and *nap*, all of which are semantically related to the critical lure *sleep*. The typical finding is that participants recall or recognize the critical lure (i.e., *sleep*) at levels comparable to or even exceeding that for presented list items (e.g., Anastasi, Rhodes, & Burns, 2000; Rhodes & Anastasi, 2000; Roediger, Watson, McDermott, & Gallo, 2001). Several investigators have shown that older adults exhibit more false memories in the DRM paradigm than younger adults (Koustaal & Schacter, 1997; Norman & Schacter, 1997; Tun, Wingfield, Rosen, & Blanchard, 1998) on both recall and recognition measures. Evidence from a number of other paradigms also indicates that older adults, in comparison to younger adults, are particularly susceptible to false memories (e.g., Bartlett, Halpern, & Dowling, 1995; Castel & Craik, 2003;

Dywan & Jacoby, 1990; Hess & Slaughter, 1990; Jacoby, 1999a; 1999b; Kelley & Sahakyan, 2003).

Why do older adults exhibit elevated levels of false memories in comparison to younger adults? Older adults certainly demonstrate impairments in memory performance (see Balota, Dolan, & Duchek, 2000; Zacks, Hasher, & Li, 2000, for reviews) that may be indicative of a more general decline in cognitive functioning. Several researchers have suggested that a decline in higher level cognitive processes, termed "executive functions," are a primary factor in the cognitive deficits present in aging populations (e.g., Moscovitch & Winocur, 1995; West, 1996; Whelihan & Lesher, 1985). Executive processes have been specified in a number of ways and may include the monitoring and control of behavior, suppression of irrelevant information, reasoning, updating information in working memory, inhibition of prepotent behavior, planning, shifting, and control of attention, among others (e.g., Baddeley, 1996; Kane & Engle, 2002; Koechlin, Ody, & Kouneiher, 2003; Miyake, Friedman, Emerson, Witzki, & Howerter, 2000; Shimamura, 2000a; 2000b; Waltz et al., 1999). These executive functions are presumably localized in the prefrontal cortex (PFC) (e.g., Duncan, 1993, 1995; Goldman-Rakic, 1987; Kane & Engle, 2002; Kimberg & Farah, 1993; Shallice & Burgess, 1991; Waltz et al., 1999).

Neurologically, there is compelling evidence that the neural substrates underlying the PFC, the putative "seat" of executive function, deteriorate more rapidly with age than other cortical regions (see Raz, 2000, for a review). There is also extensive behavioral evidence indicating that older adults perform more poorly than young controls on a number of tasks thought to tap executive functioning that are likewise sensitive to frontal lobe lesions. For example, older adults make more perseverative errors on the Wisconsin Card Sorting Test (WCST; see Rhodes, in press, for a review), show greater interference on incongruent trials of the Stroop task (Houx, Jolles, & Vreeling, 1993), produce fewer words on tests of verbal fluency (e.g., Howard, 1980) and demonstrate impairments on a variety of source memory tasks (e.g., Glisky, Polster, & Routhieaux, 1995; McIntyre & Craik, 1987; Schacter, Harbluk, & McLachlan, 1984). Moreover, several studies have demonstrated correlations between performance on such executive tasks and volumetric measures of the PFC (Head, Raz, Gunning-Dixon, Williamson, & Acker, 2002; Raz, Gunning-Dixon, Head, Dupuis, & Acker, 1998; Schretlen et al., 2000).

In regard to the focus of the current study, evidence also exists suggesting a relationship between older adults' scores on tests of executive functioning such as the WCST and memory accuracy. For example, Meade and Roediger (2002, November) recently demonstrated that older adults scoring poorly on tests of executive function were highly likely to make errors of commission during recall in the DRM paradigm. In contrast, older adults with relatively high scores on such executive tests exhibited levels of accuracy comparable to young adults. This is consistent with data from other memory tasks showing significantly better memory performance by older adults scoring high on measures of executive functioning in comparison to groups of older adults with relatively poor performance on such tasks (e.g., Davidson & Glisky, 2002; Glisky et al., 1995; Glisky, Rubin, & Davidson, 2001; McDaniel, Glisky, Rubin, Guynn, & Routhieaux, 1999). Other investigators have shown that measures of executive function are predictive of recall accuracy (Bryan, Luszcz, & Pointer, 1999; Crawford, Bryan, Luszcz, Obonsawin, & Stewart, 2000) and may largely mediate age-related deficits in episodic memory (Troyer, Graves, & Cullum, 1994).

Thus, it appears that a relationship exists between performance on neuropsychological tests of executive function and memory accuracy. This begs the question of why such measures are predictive of older adults' memory accuracy. To preview, the current study will examine this relationship from two perspectives. First, executive functions may be crucial to the retrieval of accurate information. For example, executive functions may be important in binding critical details during encoding (e.g., Henkel, Johnson, De Leonardis, 1998) or generating cues during retrieval that will elicit information diagnostic of prior occurrence (cf., Burgess & Shallice, 1996). Second, executive functions may be important for monitoring candidate responses for accuracy (Shimamura, 2000a, 2000b). The current study will use the memory monitoring framework developed by Koriat and Goldsmith (1994; 1996) to examine the relationship between measures of executive function and memory accuracy. Koriat and Goldsmith's framework is ideal for investigating such issues because it permits a separate assessment of the accuracy of retrieved responses and the effectiveness of monitoring processes. In turn, the relationship between each of these components (i.e., memory accuracy and monitoring) and measures of executive function can be examined. However, before describing the details of the current study, each proposed executive component of memory accuracy will first be considered in turn.

## Executive Processes and Episodic Representations

Executive processes may play a critical role in determining whether veridical information comes to mind during retrieval. For example, executive processes may be important for binding information during encoding (cf., Chalfonte & Johnson, 1996; Glisky et al., 2001; Henkel, et al., 1998) or maintaining an appropriate retrieval set that allows one to generate cues that are likely to lead to the retrieval of accurate information (Burgess & Shallice, 1996; Moscovitch & Melo, 1997; Norman & Schacter, 1996). For example, Norman and Schacter (1996) have proposed that high levels of false memories evident in frontal patients occur either because they are totally unable to recapitulate the study context or because they produce "unfocused retrieval descriptions" that provide a vague representation of the study context.

False memories in older adults may also be a consequence of relying on vague or unfocused representations of study episodes. Jacoby and others (e.g., Jacoby, 1999a; Jacoby, 1999b; Jennings & Jacoby, 1997; Kelley & Sahakyan, 2003; see also Parkin & Walter, 1992) have suggested that memory deficits in aging populations are indicative of a general decline in recollective bases for memory and an increased reliance on general familiarity. For example, Jacoby (1999b) had young and old adults read a list of words in which some items were presented once, and others were repeated two or three times. Participants next heard a list of words presented a single time. They were then administered a recognition test with instructions to endorse only to those items that were heard. Presumably, repetition should increase the chance that participants recollect those words as only read and thus correctly reject them at test. Results showed that this was the case for younger participants, as they were less likely to falsely recognize words read three times than those read only once. In contrast, older adults were significantly more likely to falsely recognize words read three times in comparison to those words read only once. Jacoby suggested that older adults were unable to counter an increase in familiarity due to repetition with recollection of those items that were heard.

Jennings and Jacoby (1997) have reported similar effects of repetition by repeating lures during a recognition test. Their results indicated that older participants were more likely than younger adults to endorse lures repeated at test, even at short lags (see also Kensinger & Schacter, 1999; Koriat, Ben-Zur, & Sheffer, 1988). There is also evidence indicating that older adults are less likely to engage in encoding operations that facilitate recollection (e.g., Perfect & Dasgupta, 1997) and may respond to memory queries based only on the plausibility of a cue

(Reder, Wible, & Martin, 1986). In addition, encoding manipulations that increase the amount of detail available to participants during retrieval, such as providing pictorial information (e.g., Dodson & Schacter, 2003), sharply reduce the level of false memories exhibited by older adults. Thus, older adults' increased susceptibility to false memories may in part stem from a deficit in retrieving detailed memories of a study episode, such that they may rely on information that is merely familiar and unlikely to be diagnostic of prior occurrence (cf., Kelley & Sahakyan, 2003). Given that older adults may be left to rely on a paucity of detail from the study episode, achieving source monitoring accuracy (i.e., Johnson, Hashtroudi, & Lindsay, 1993) will be difficult. Typically, source monitoring is most accurate when candidate memories are accompanied by specific details, such as information about the perceptual qualities of the memory in question (e.g., Johnson, Foley, Suengas, & Raye, 1988). As older adults are less likely to generate such details during retrieval, accurate source monitoring may be problematic.

There is extensive evidence that source monitoring accuracy is impaired in older adults (e.g., Craik, Morris, Morris, & Loewen, 1990; Dywan & Jacoby, 1990; Glisky et al., 1995; Henkel et al., 1998; McIntyre & Craik, 1987; Spencer & Raz, 1995). Glisky et al. (1995; 2001), among others (e.g., Craik et al., 1990; Janowsky, Shimamura, & Squire, 1989a; Johnson, Kounios, & Nolde, 1996; Spencer & Raz, 1995; but see Dywan, Segalowitz, & Williamson, 1994; Johnson, De Leonardis, Hashtroudi, & Ferguson, 1995) have linked such source monitoring deficits in older adults to a general decline in frontal lobe function. Source monitoring accuracy is clearly disrupted in frontal patients (Janowsky, et al., 1989a; Shimamura, Janowsky, & Squire, 1990; Schacter, et al., 1984). Behavioral evidence from older adults also suggests that the frontal lobe plays a critical role. For example, Glisky et al. (1995) preexperimentally divided older adults into groups based on a composite measure of frontal and medial temporal lobe function, with the medial temporal measure presumed to control for differences in memory. This permitted groups of older adults to be classified as either high or low on each composite measure. Source memory was assessed by having participants judge whether previously presented sentences were read by a male or female speaker. Results indicated that the high frontal group exhibited significantly better source memory than the low frontal group, who performed at chance levels. In contrast, recognition memory for the sentences did not differ between the two frontal groups.

Therefore, taken together, it is apparent that older adults are less likely to generate details during retrieval that will be diagnostic of prior study. In turn, given the lack of accurate information available during retrieval, monitoring will be hampered, as there may be little detail available to distinguish between studied and nonstudied information. Executive functions may play a key role in retrieving accurate information by binding important details during study (e.g., Chalfonte & Johnson, 1996) or generating cues that will lead to retrieval of accurate and sufficiently detailed information (i.e., Burgess & Shallice, 1996). As well, executive functions may also be important for accurate source monitoring (e.g., Glisky et al., 1995).

**Executive Processes and Memory Monitoring**

Executive functions may also be crucial to memory accuracy via metacognitive process such as memory monitoring. Monitoring and controlling the contents of cognitive processes likely comprises an important executive function (cf. Fernandez-Duque, Baird, & Posner, 2000) that has implications for the attainment of accuracy (e.g.; Shimamura, 2000a, 2000b). Evidence from frontal patients indicates that they have considerable difficulty with complex metacognitive tasks such as memory monitoring (Janowsky, Shimamura & Squire, 1989b; Vilkki, Servo, & Surma-aho, 1998; see Shimamura, 1996, for a review).

Older adults demonstrate performance comparable to younger adults on a number of metacognitive tasks (see Hertzog & Hultsch, 2000, for a review). However, age differences do exist and depend primarily on the type of judgment elicited. For example, Connor, Dunlosky, and Hertzog (1997; Experiment 2) presented older and younger participants with unrelated word pairs and had participants make predictions about the probability of later recalling the second item of the pair (i.e., judgments of learning or JOLs). Results showed that both younger and older adults demonstrated strong concordance between their JOLs and subsequent recall performance. In contrast Souchay, Isingrini, and Espagnet (2000) have reported age differences in metamemory performance using feeling-of-knowing (FOK) judgments. Younger and older participants were presented with unrelated word pairs and administered a cued-recall test. For items that could not be recalled, participants were asked to rate how likely they would be to recognize that item if it was presented on a recognition test (i.e., FOK judgments). Older adults' FOK judgments were generally much less accurate than those of younger adults. Interestingly, the accuracy of younger and older adults' FOK judgments was positively correlated with measures of executive function, such as the WCST and verbal fluency. Souchay, Isingrini, Clarys, Taconnat, and Eustache (2004) have recently replicated this finding and further demonstrated that performance on executive tasks was not correlated with JOLs. Thus, age differences are apparent in certain metamemory tasks (e.g., Souchay et al., 2000; 2004). Combined with evidence that frontal patients show severe impairments on metamemory tasks (Janowsky et al., 1989b) and measures of executive function are predictive of metamemory performance (e.g., Souchay et al., 2000; 2004) it appears that memory monitoring is an executive function that is impaired on some tasks in older adults.

## The Current Study

The current study examined the role of executive functions in memory accuracy using the memory monitoring framework developed by Koriat and Goldsmith (1994; 1996). Koriat and Goldsmith suggest that memory performance is not simply a function of retrieving a latent memory trace, but also depends on the degree to which the rememberer can monitor the correctness of responses, uses this information to select a correct response, and sets an appropriate criterion based on incentives for accuracy. An important component of this perspective is that it distinguishes between quantity and accuracy-based measures of performance. Specifically, quantity is measured as the number of correct responses elicited during an initial, forced recall stage. However, control processes may then operate during a subsequent, free recall stage, permitting participants to control the accuracy of their output. Accuracy can thus be measured as the proportion of correct responses volunteered out of the number of responses offered. In particular, Koriat and Goldsmith's framework suggests that participants monitor the correctness of a candidate response by assigning it a probability of being correct ($P_a$). They then control their responding by setting a response criterion, $P_{rc}$. If $P_a$ exceeds or is equal to $P_{rc}$ an answer is volunteered; otherwise, the response is withheld. The criterion ($P_{rc}$) can be adjusted based on payoffs specifying the relative costs and benefits for correct and incorrect responses. Consequently, performance at free report is dependent on three factors. The first is *monitoring effectiveness*, which is the extent to which assessed probabilities of being correct (i.e., $P_a$ ) successfully distinguish between correct and incorrect responses. A second factor is *control sensitivity*, which captures the degree to which participants base their decision to respond on the assessed probability of being correct. Finally, participants may adjust their response criterion depending on incentives for accuracy (*response criterion setting*). For example, stringent penalties for incorrect responses should lead the response criterion to be adjusted upward.

Kelley and Sahakyan (2003) have used Koriat and Goldsmith's (1994; 1996) framework to investigate memory accuracy and monitoring performance in older and younger adults. Specifically, they tested older and younger adults in a cued recall paradigm (Kato, 1985) that results in high levels of false recall of easily retrieved but incorrect candidate responses. In the paradigm, participants study related and unrelated pairs of words. At test, recall of the second word in the pair is cued by pairing the first word with three letters of the target (e.g., the study pair CLOCK DOLLAR is tested with the cue CLOCK DO _ _ _ R). The cues for some of the unrelated pairs result in easily produced but incorrect responses. For example, for the study pair NURSE-DOLLAR, the test cue is NURSE DO _ _ _ R, with the associatively related but incorrect item "doctor" a strong competitor for response.

Participants in Kelley and Sahakyan's (2003) experiment first studied a list of word pairs and were then administered a cued-recall test in several stages. During an initial, forced recall stage, participants generated a studied item in response to a cue (e.g., NURSE DO _ _ _ R) or guessed if necessary to provide a response. Next, participants assessed the probability that the response was correct, permitting an evaluation of monitoring effectiveness. In a second, free report phase, participants decided whether to volunteer a candidate response, providing a measure of control sensitivity. Results showed that older adults were significantly less likely to produce a correct response during the forced report stage (i.e., quantity) and, as well, were less accurate than younger adults for those responses that they chose to volunteer during the free report phase (i.e., accuracy). These differences were particularly acute for deceptive items. In addition, older adults were less effective at monitoring the correctness of responses, as indicated by their lower $\gamma$ correlations (a nonparametric measure of the relationship between assessed confidence and correctness) and greater calibration error.

The current study used the same approach as Kelley and Sahakyan (2003) in order to examine the role of executive processes in memory accuracy. In particular, participants were tested in the memory monitoring task employed by Kelley and Sahakyan and were also administered a battery of neuropsychological tests designed to measure executive functioning. As noted, executive functions may be important in retrieving accurate information and may also be crucial in monitoring the accuracy of candidate responses (Shimamura, 2000a, 2000b). The paradigm to be used in the current study is well suited for examining these issues as it permits separate measurements of the quantity of correct items available at forced report and the accuracy of responses at free report, which will be a function of memory monitoring.  One can then determine whether executive functions are predictive of quantity, accuracy, and/or memory monitoring.  Figure 1 illustrates this notion conceptually.  The top panel of Figure 1 (a) proposes that accuracy in free recall is the product of a direct relationship with executive function.  This conceptualization underlies, at least implicitly, much of the work that has examined individual differences in memory accuracy, particularly in cases where individual differences have been described using measures of executive function (e.g., McDaniel et al, 1999; Meade & Roediger, 2002, November; but see Troyer et al., 1994).  However, the current study also permits an assessment of memory monitoring, illustrated in the bottom panel (b) of Figure 1.  In this case, any effect of executive functions on memory accuracy is mediated by memory monitoring accuracy.  That is, the figure suggests that the efficacy of memory monitoring accounts (fully or partially) for the relationship between executive functions and memory accuracy.  The current study will thus test whether memory monitoring mediates (cf. Baron & Kenny, 1986) the executive-accuracy relationship.

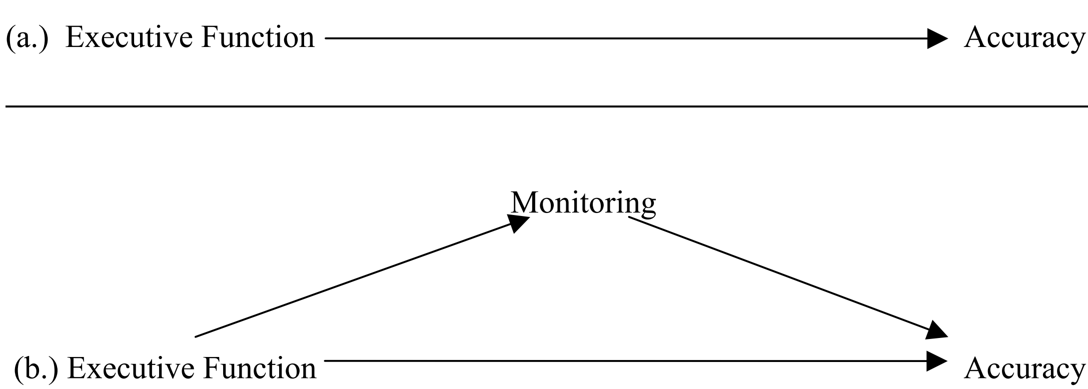



**Figure 1.  Proposed Relationships Between Accuracy and Executive Function**
*Note*.  The top panel (a) illustrates a direct relationship between accuracy and executive function while the lower panel (b) illustrates the relationship with monitoring as a mediator.

Two other types of individual difference measures were taken. First, measures of speed of processing were collected (Salthouse, 1996). Salthouse (see also Cerella, 1985) has suggested that age-related differences in cognition are largely the result of a more basic deficit in processing efficacy. Slower processing speed may act by reducing the number of operations that can be executed simultaneously and by making inaccessible the contents of earlier operations. Data from speed tasks have revealed substantial negative correlations between age and performance on a variety of measures (e.g., Salthouse, 1994; Salthouse, Fristoe, Lineweaver, & Coon, 1995). Several researchers have shown that speed of processing measures may account for the majority of variability reflected in measures of executive function (Crawford et al., 2000; Parkin & Java, 1999). Thus, analyses in the current study controlled for speed of processing in addition to examining executive function correlates of memory accuracy and memory monitoring.

Second, working memory was measured using the operation span (OSPAN) task (Turner & Engle, 1989). OSPAN was originally included as an additional measure of executive functioning. Kane and Engle (2002) have suggested that working memory span scores are indicative of a general ability to control attention and maintain information in the face of interference, and therefore tap into a basic component of executive function. However, to preview, results of a factor analysis showed that OSPAN did not load on a factor comprised of executive functioning measures. Thus, OSPAN was included as a separate measure of working memory span to determine the contribution of executive functions beyond that explained by working memory span.

# EXPERIMENT

## Method

### Participants

Participants were 50 undergraduates at Florida State University who received course credit for their participation, and 50 older adults aged 65 and over who were recruited from lists of university alumni and through notices published in local newspapers (characteristics of the sample are presented in Table 1). Older adults were paid $10 an hour for their participation. All participants reported good or excellent health and normal or corrected vision. In addition, participants were screened for dementia using the Mini-Mental Status Exam (Folstein, Folstein, & McHugh, 1975). All participants achieved scores of 27 or above (i.e., normal cognitive function) and were included in the current study.

**Table 1. Means (and Standard Deviations) of Demographic Characteristics of the Participants.**

| Variable | Younger Adults | Older Adults |
|---|---|---|
| *N* | 50 | 50 |
| Age | 19.64 (1.19) | 72.00 (5.40) |
| % Females | 55 | 59 |
| Years of Education | 14.21 (1.20) | 16.70 (2.04) |
| MMSE[a] | 29.46 (.86) | 28.70 (1.07) |
| NAART-35[b] | 13.90 (5.16) | 22.79 (7.14) |

[a] Mini-Mental State Examination (MMSE) score represent the number of points earned out of a possible total of 30. [b] North American Adult Reading Test-35 scores represent the total number of correctly pronounced irregular words out of a maximum total of 35.

## Materials

The materials used in the current experiment consisted of 75 word pairs, one third of which were related filler items (e.g., "morning-evening") and two thirds of which were unrelated word pairs. Half of the unrelated word pairs (*deceptive items*) had potentially interfering competitors that were not only semantically related to the cue word but shared the same first two letters and last letter as the target word (e.g., the "nurse-dollar" pair had the interfering competitor, "doctor"). The other half of the word pairs consisted of unrelated *control items* that have no such interfering competitors (e.g., "clock-dollar").

All study items were taken from Kato (1985) and Kelley and Sahakyan (2003). The study list consisted of 60 pairs of words, distributed equally across control items, unrelated deceptive items, and related fillers. As well, a study buffer consisting of one related pair ("month-year") and one unrelated pair ("turkey-opera") were included in the study list to serve as practice items at test. The test list consisted of all 60 test items and 5 new items of each type (i.e., control, unrelated deceptive, and related filler items), for a total of 75 items. Test items were counterbalanced for old-new status and for presentation as control or deceptive items. A 3-item practice buffer was also included to familiarize the participant with the test procedure. The practice buffer consisted of two pairs of items presented as the study buffer and an additional related pair ("bride-gr_ _ m").

## Measures

Participants were administered several measures designed to assess vocabulary, executive function, working memory, and speed of processing. These measures are as follows.

### Measure of Vocabulary.

*North American Adult Reading Test-35 (NAART-35)*. The North American Adult Reading Test (Blair & Spreen, 1989) requires participants to read from a list of irregularly spelled words (e.g., *debris*, *hors d'oeuvre*, *façade*). Scores were calculated based on the total number of correct pronunciations. The test is highly correlated with measures of vocabulary and verbal IQ (see Spreen & Strauss, 1998, for a review). A typical administration requires the participant to read 50 words but Uttl (2002) has demonstrated that the test can be shortened to 35 words with no loss of validity or reliability. Thus, the 35-item version of the NAART (i.e., the NAART-35) was administered. Data from this task was collected as a descriptor of the sample tested and was not included as a predictor in any of the analyses reported.

**Speed of Processing Measures.**

*Digit-Symbol Substitution Task*. In the Digit Symbol Substitution task, participants are presented with a code table consisting of a row of numbers (1-9) above a row of symbols. The participant is given a set of problems—a digit accompanied by a blank—and instructed to write the correct symbol for each digit as fast as possible. The number of correct symbols written in 90 s was recorded as the dependent measure.

*Number Comparison*. In the Number Comparison task (Salthouse & Babcock, 1991) participants are presented with columns consisting of pairs of numbers. Participants are given 90 s to determine which pairs of numbers are the same and which are different. The number of correct responses minus the number of incorrect responses given in 90 s is recorded as the dependent measure. Participants were tested in two trials of 90 s.

**Executive Function Measures.**

The tests of executive function described below were chosen because they are common measures of executive function and/or because they have been proven to be sensitive to lesions to the frontal lobe.

*Wisconsin Card Sorting Test (WCST)*. In the WCST (Grant & Berg, 1948; Heaton, Chelune, Talley, Kay, & Curtiss, 1993) the participant is instructed to sort 2 sets of 64 response cards based on several dimensions (e.g., color, form, number), which change periodically during the course of sorting. Participants are not instructed as to how to sort the cards but must infer the correct sorting principles through limited feedback from the experimenter, who only tells the participant whether the sort is correct or incorrect. After 10 consecutive successful sorts the sorting principle changes and the participant must adjust accordingly. In the current study the two most common measures, the number of categories achieved and the number of perseverative errors committed, were recorded as dependent variables. A participant is scored as having achieved a category when they complete 10 correct, consecutive sorts for a particular category. Perseverative errors occur when the participant sorts response cards according to a category that was formerly correct, but is no longer in effect. The test was administered via computer.

*Controlled Oral Word Association Test (COWA; or FAS-Test).* Participants in the *COWA* (Benton & Hamsher, 1976) are instructed to orally produce in separate, 1-minute trials, as many words as possible beginning with the letters *F*, *A*, and *S*. (Given that the most common letters used are *F*, *A*, and *S* and the test is often referred to as the *FAS-Test*.) Participants are instructed to avoid using proper nouns, numbers, and the same words ending in different suffixes (e.g., *eat*, *eating*). The total number of unique words produced across all three trials was recorded as the dependent measure.

*Trail Making Test* (*TMT*). The TMT (Reitan, 1958) is divided into two parts. Part A of the test requires participants to connect encircled numbers that are randomly arranged on a page in ascending order. Part of the B the test has the same general format, with the exception that it requires participants to alternate between encircled numbers and encircled letters. Performance was measured in terms of the time in seconds to complete Parts A and B of the test and also as the difference (in seconds) between the time to complete Parts A and B.

**Working Memory Span Measure.**

*Operation-Span (OSPAN)*. Working memory capacity was measured using the OSPAN task developed by Turner and Engle (1989). In the task, participants are simultaneously presented with simple math problems and unrelated words. Participants were instructed to read the math problem aloud, confirm whether the answer given was correct or incorrect, and then immediately read the word aloud. For example, a participant might see the following: Is (9 x 2) + 2 = 16 ARM. The participant would say, "Is nine times two plus two equal to sixteen, no, arm". Immediately after this, the experimenter would press a key and another math problem and unrelated word appeared. After a series of two to five math/word presentations (the order is random with the condition that there is an equal number of series that are two, three, four, or five entries in length), the participant was signaled to write down the words they studied in the order that they were presented. Participants were given credit only if they recalled the entire sequence of words in the order that it was presented. For example, if a participant were presented with:

Is (8 x 4) – 2 = 34, BEAN

Is (9 x 3) – 3 = 24, ARM

Is (4 ÷ 1) + 1 = 4, GROUND

he or she was only given credit for recalling BEAN, ARM, and GROUND, in that order. Participants were not given credit if any words were missing in their recall (e.g., BEAN) or if the serial order was incorrect (e.g., GROUND, ARM, BEAN). The participant's score was recorded as the total number of words recalled from those trials in which all words were recalled in the correct order.

**Procedure**

After providing informed consent participants filled out a survey asking them to list their age, years of education, and rate their current health status. They were then administered the MMSE. Following the MMSE, participants completed Parts A and B of the Trail Making Test and the NAART-35. Next, participants were tested in the memory monitoring task that is the primary focus of the current study. All study and test items were presented via computer using Micro Experimental Laboratory (MEL) software (Schneider, 1990). Study pairs were presented at an 8 s rate and participants were instructed to remember the pairs for a later test. Words pairs were presented in the center of a computer screen in a large font. The study list was followed by a brief filler activity requiring participants to write down as many names of boys, girls, and last names of United States presidents as they could think of in 5 minutes.

Following the filler task, participants moved on to the test phase. The cued recall test took place in two stages. In the first stage, participants were instructed to use the cue (e.g., NURSE DO _ _ _ R) to either retrieve an item that was studied or, failing that, to generate an item that fit the cue. Once the participant either recalled or guessed a target item they were asked to assess the likelihood that the target item was studied using a 0 – 100% confidence scale. They were informed that a confidence rating of 100% would correspond to a 100% likelihood that the target was the correct answer whereas a confidence rating of 0% would correspond to a 0% likelihood that the target was the correct answer. Participants were encouraged to use the full range of the confidence scale. After making the confidence judgment, participants moved on to the free-recall stage of the test. During this stage, participants were instructed to either report an item that was studied, say "new" if the test cue was new, or "pass" if they were unsure whether the item was old or new. Prior to the test, participants were informed that the test would contain some new cues and that they should avoid guessing during the free report phase. Participants were also given instructions to be accurate. This was accomplished by telling participants that they would earn 10 points for each correct answer and would be penalized 10 points for each incorrect answer. They were also instructed that passing would result in no award or penalty. There was no time limit for any of the stages of the memory monitoring test.

Upon completion of the memory monitoring test, participants were given a rest break of 10 to 20 minutes. Following the break, participants were administered, in order, the WCST, FAS-Test, DSST, Number Comparison task, and the OSPAN task. Upon completion of these tasks, participants were debriefed and thanked for their participation.  The experiment took approximately one and a half to three hours to complete.

# RESULTS

Analyses for the current study can be broadly categorized into analyses of performance on the memory monitoring task (memory performance and monitoring performance), performance on the various tests administered, and the relationship between performance on these tests and the memory monitoring task. Each of these components of the data analysis will be considered in turn. Unless otherwise noted the alpha level for all statistical tests reported was set to $p < .05$.

**Memory Performance.**

Memory performance on the monitoring task (see Table 2) was analyzed in several different ways based on responses given during the forced-report and free-report stages.

**Table 2. Means (and Standard Deviations) of Quantity and Accuracy Scores for the Free and Forced Report Condition by Age and Item Type.**

| | | Report Option | | |
|---|---|---|---|---|
| | | Forced | Free | |
| Age Group | Item Type | Quantity and Accuracy | Quantity | Accuracy |
| Younger | Control | .59 (.18) | .54 (.21) | .80 (.17) |
| | Deceptive | .42 (.24) | .42 (.24) | .58 (.27) |
| Older | Control | .51 (.19) | .44 (.22) | .76 (.24) |
| | Deceptive | .31 (.18) | .31 (.18) | .43 (.26) |

*Forced report*. The quantity of correct responses produced was first assessed by examining retrieval during the forced report stage. These data were analyzed in a 2 (Age: younger, older) x 2 (Item Type: control, deceptive) mixed-factor analysis of variance (ANOVA).

Results showed that older adults retrieved a significantly lower quantity of correct responses (*M* = .41) in comparison to younger adults (*M* = .51), *F* (1, 98) = 7.15, *MSe* = .45. Thus, younger adults began the monitoring process with a greater number of correct responses. Participants were far more likely to retrieve correct responses for control items (*M* = .55) than for deceptive items (*M* = .36), *F* (1, 98) = 117.49, *MSe* = 1.85. Finally, age did not interact with item type, *F* (1, 98) < 1.

*Free report*. Accuracy of responses was assessed based on data from the free report stage. Specifically, accuracy was measured as the number of correct responses given divided by the total number of responses volunteered. If participants are sensitive to the correctness of candidate responses, they should demonstrate gains in accuracy from the forced to the free report stage. Accuracy scores were analyzed in a 2 (Age: younger, older) x 2 (Item Type: control, deceptive) x 2 (Report Option: forced, free) mixed-factor ANOVA. These data indicated that participants were considerably more accurate at free report (*M* = .64) than at forced report (*M* = .43), *F* (1, 98) = 251.87, *MSe* = 3.42. In addition, accuracy was much higher for control items (*M* = .67) than for deceptive items (*M* = .43), *F* (1, 98) = 154.34, *MSe* = 5.39, and younger adults (*M* = .60) were more accurate than older adults (*M* = .50), *F* (1, 98) = 6.63, *MSe* = .91. A triple interaction of age, item type, and report option was also present, *F* (1, 98) = 4.63, *MSe* = .03. This reflects the fact that younger adults made largely equivalent gains in accuracy for control and deceptive items from forced to free report. In contrast, while older adults made gains in accuracy for control items that were somewhat larger than those made by younger adults, older adults made a less substantial gain in accuracy for deceptive items.

The accuracy data clearly shows that participants made considerable gains in accuracy from forced to free report. However, gains in accuracy may come at the cost of lower quantity correct, as participants may be unable to differentiate between correct and incorrect candidate responses produced at forced report. To address this issue, quantity correct was examined in a 2 (Age: younger, older) x 2 (Item Type: control, deceptive) x 2 (Report Option: forced, free) mixed-factor ANOVA. Results confirmed that the quantity of correct responses was diminished at free report (*M* = .42) in comparison to forced report (*M* = .46), *F* (1, 98) = 65.42, *MSe* = .15. This finding is qualified by a significant item type x report option interaction, *F* (1, 98) = 37.14, *MSe* = .06. Specifically, participants demonstrated greater losses in quantity from forced to free report for control items than for deceptive items. A trend was also evident for older adults to exhibit greater losses in quantity in comparison to younger adults from forced to free report. However, the interaction of age and report option was only marginally significant, $F$ (1, 98) = 3.19, $MSe$ = .007, $p$ = .08.

As an additional measure of gains from forced to free report, gain scores were calculated for each participant. Gain scores were based on the relative rather than the absolute magnitude of the gain from forced to free report. For example, consider a participant whose proportion of correct responding during forced report was .80, and who attained an accuracy level of .90 at free report. Based on absolute magnitude, that participant made a gain in accuracy of .10 (i.e., .90 - .80 = .10). However, after achieving a proportion correct of .80 at forced report, the largest gain possible (i.e., that gain which would result in perfect accuracy, 1.0) was .20. Given that the largest possible gain was .20, a gain in accuracy at free report of .10 was half of what was possible (i.e., .10/.20 = .50). In contrast, a participant reporting .40 correct at forced report and an accuracy level of .50 at free report would have made the same absolute gain but, in relative terms, one that was significantly smaller given that the largest possible gain was .60. Thus, in relative terms, their gain score would be .17 (i.e., .10/.60 = .17). Therefore, gains were quantified for control and deceptive items, based on the formula below:

(1) Gain = (Free Report Accuracy – Forced Report Quantity)/(1 – Forced Report Quantity)

where *free report accuracy* is the proportion of correct responses out of the total number volunteered at free report and *forced report quantity* is the proportion of correct responses given at forced report. Free report accuracy was entered first in the numerator to ensure that all scores would be positive.

Examination of gain scores indicates that older and younger adults made largely equivalent gains from forced to free report for control items (.57 and .56 for older and younger adults, respectively). However, older adults (.21) made smaller gains than younger adults (.34) for deceptive items. These data were subjected to a 2 (Item Type: control, deceptive) x 2 (Age: young, old) mixed ANOVA. Results showed that participants made significantly larger gains for control items (.56) than for deceptive items (.28), $F$ (1, 98) = 84.15, $MSe$ = 4.04. Gains did not differ between age groups [$F$ (1, 98) = 1.51, $MSe$ = .19, $p$ = .22] but a significant interaction of age and item type was present, $F$ (1, 98) = 5.59, $MSe$ = .27. This reflects the fact while younger and older individuals made equivalent gains for control items, older adults made smaller gains than younger adults for deceptive items.

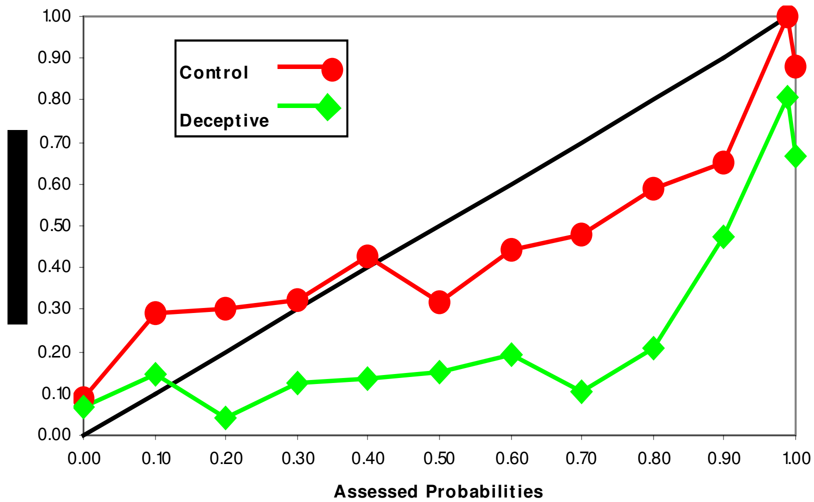
Overall, older adults produced fewer correct responses during forced report than did younger adults. In turn, younger adults exhibited greater accuracy than older adults at free report. Older adults made nearly equivalent gains in accuracy as younger adults from forced to free report for control items, but made less substantial gains in accuracy for deceptive items. Differences in gains in accuracy from forced to free report may reflect the efficacy of monitoring processes, which will be examined next.

**Memory Monitoring.**

Participants provided an immediate confidence rating for each item given during forced report[1]. Confidence ratings were then grouped into 12 levels (e.g., 0, .01-.10, .11-.20, .21-.30,… .91-.99, 1.0). Using these confidence ratings, memory monitoring effectiveness was quantified in two ways. First, *calibration* refers to the absolute correspondence between assessed levels of confidence and the actual probability that a participant's response is correct. Perfect calibration would suggest that items given an assessed confidence of 50% are correct 50% of the time, items given an assessed confidence of 60% would be correct 60% of the time, and so forth. Second, *monitoring resolution* was calculated using the Kruskal-Goodman $\gamma$ correlation (Nelson, 1984) and the Adjusted Normalized Discrimination Index (ANDI; Yaniv, Yates, & Smith, 1991). These measures of monitoring resolution permit an assessment of the degree to which assessed confidence distinguished between correct and incorrect responses.

Figures 2 and 3 display calibration curves for older and younger participants for control and deceptive items. The proportion correct is plotted against the mean assessed probability across participants. A diagonal line with an intercept of 0 and a slope of 1.0 represents perfect calibration. Inspection of the figures shows that both younger and older adults were well calibrated for control items. For example, younger adults reported an average assessed confidence of .64 while their actual proportion correct was .59. Older adults reported an average assessed confidence of .54 while their actual proportion correct was .51. However, calibration on deceptive items was significantly worse for both younger and older participants. Younger adults reported an average assessed confidence of .70 while their actual proportion correct was .42. Older adults likewise demonstrated considerable overconfidence on deceptive items, with an average assessed confidence of .66 whereas their actual proportion correct was .31.[2]

**a.**




**b.**

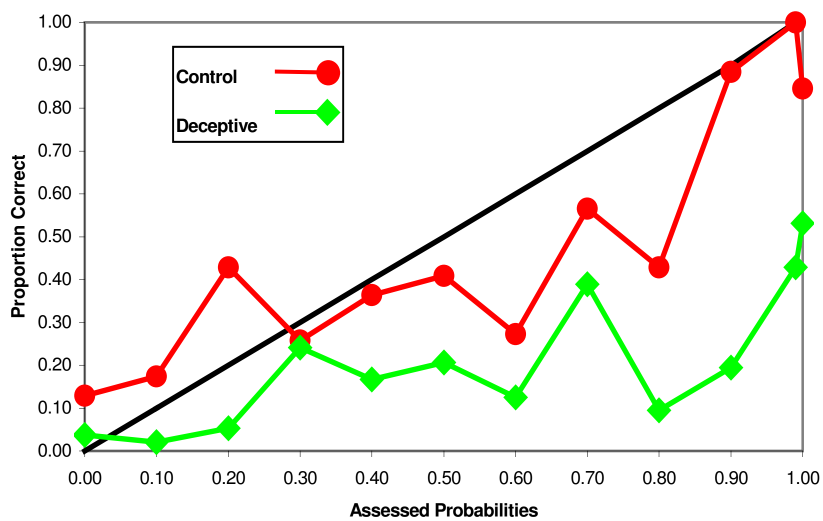



**Figure 2. Calibration Curves**
*Note:* Calibration curves for deceptive and control items for younger adults (a) and older adults (b). The diagonal line represents perfect calibration.

A further index of calibration is captured using calibration error scores (cf. Oskamp, 1962). Such scores are computed as the weighted mean of the absolute difference between the assessed probability and the actual proportion correct for each category. Older and younger adults exhibited identical calibration error scores (.19) for control items. In contrast, for deceptive items, older adults (.39) demonstrated slightly larger calibration error scores than younger adults (.33). A 2 (Age: younger, older) x 2 (Item Type: control, deceptive) mixed-factor ANOVA on calibration error scores confirmed that calibration error was greater for deceptive items ($M$ = .36) than for control items ($M$ = .19), $F$ (1, 98) = 63.91, $MSe$ = 1.49. However, neither the main effect of age, $F$ (1, 98) = 1.14, $MSe$ = .04, $p$ = .29, nor the the interaction of age and item type was significant, $F$ (1, 98) = 1.47, $MSe$ = .03, $p$ = .23.

Monitoring resolution was next examined based on γ correlations between assessed confidence and quantity correct at forced report. Occasionally, γ correlations could not be calculated, as in the case when responses were either completely correct or completely incorrect. This occurred for deceptive items for one younger participant and one older participant and these data were excluded from the present analysis. On the whole, γ correlations did not differ for control items, with young and old adults exhibiting largely equivalent mean γ scores (.81 and .78, for younger and older adults, respectively). Monitoring resolution was worse for deceptive items, with younger and older adults achieving γ correlations of .75 and .67, respectively. A 2 (Age: young, old) x 2 (Item Type: control, deceptive) mixed ANOVA confirmed that resolution was marginally better for control items ($M$ = .80) in comparison to deceptive items ($M$ = .71), $F$ (1, 96) = 3.50, $MSe$ = .45, $p$ = .07. These data are consistent with mean confidence ratings and calibration error scores showing that participants in both age groups were consistently overconfident for deceptive items. Further analysis of the data showed that γ correlations did not differ between young and old adults[3], $F$ (1, 96) < 1, nor did age interact with item type, $F$ (1, 96) < 1. Thus, while monitoring resolution was poorest for deceptive items, age differences were not apparent. This may be contrasted with the age differences reported by Kelley and Sahakyan (2003). Inspection of their data indicates that much of the inconsistency stems from the fact that older adults' γ correlations for deceptive items in that study (.46) were considerably lower than the figure (.67) reported in the current study .

One possible problem with the γ correlations reported is that they were often subject to ceiling effects, such that participants frequently attained the highest score possible (i.e., 1.0). For example, older adults attained γ correlations of 1.0 32% of the time for control items and 43% of the time for deceptive items. Younger adults show a similar pattern with γ correlations of 1.0 30% of the time for control items and 38% of the time for deceptive items. Therefore, ANDI scores were calculated as a second measure of monitoring resolution (Yaniv et al., 1991). ANDI scores may be interpreted as the proportion of variability in the correctness of responses that is accounted for by the participant's confidence judgments. Older and younger adults exhibited largely similar ANDI scores[4]. For example, for control items, younger adults had a mean ANDI score of .55 while older adults had a mean ANDI score of .49. Younger adults exhibited slightly higher ANDI scores for deceptive items (.38 and .32 for younger and older adults, respectively). ANDI scores were analyzed in a 2 (Item Type: control, deceptive) x 2 (Age: old, young) mixed-factor ANOVA. Results showed that participants had significantly higher ANDI scores for control items ($M = .52$) in comparison to deceptive items ($M = .35$), $F(1, 98) = 15.72$, $MSe = 1.42$. ANDI scores did not differ on the basis of age [$F(1, 98) = 1.32$, $MSe = .19$, $p = .25$], nor did age interact with item type, $F(1, 98) < 1$. Thus, data from ANDI scores concurred with the pattern of data evident for $\gamma$ correlations. That is, both measures indicated that monitoring resolution did not differ between age groups and that resolution was significantly poorer for deceptive items in comparison to control items.

**Control Processes**

An additional method of investigating metacognitive processes in the achievement of memory accuracy is to examine control processes (i.e., those processes that determine whether a response is volunteered or withheld). Therefore, $\gamma$ correlations were calculated to assess the degree to which assessed confidence was related to the decision to volunteer or withhold a response[5]. Results showed that there was a strong relationship between the decision to volunteer a response and confidence ratings, particularly for control items. Specifically, mean $\gamma$ correlations for control items were .95 and .91 for older and younger adults, respectively. The relationship between assessed confidence and the decision to respond was somewhat weaker by comparison for deceptive items, with older adults exhibiting a mean $\gamma$ correlation of .89 and younger adults exhibiting a mean $\gamma$ correlation of .83. Analysis of these data indicated that assessed confidence was indeed more strongly linked to the decision to respond for control items (mean $\gamma = .93$) than for deceptive items (mean $\gamma = .86$), though this difference was only marginally significant, $F(1, 88) = 2.97$, $MSe = .22$, $p = .09$. Volunteering mean $\gamma$s did not differ by age, $F(1, 88) = 1.69$, $MSe = .12$, $p = .20$, and the interaction of item type and age was not significant, $F(1, 88) < 1$.

One may also examine the number of responses volunteered and response criterion as a measure of control over memory accuracy. Table 3 displays descriptive statistics for the number of responses volunteered for each type of item and the corresponding response criterion. The number of responses volunteered at free report was analyzed in a 2 (Item Type: control, deceptive) x 2 (Age: young, old) mixed-factor ANOVA. Overall, younger adults (*M* = .67) volunteered a slightly greater proportion of items than older adults (*M* = .62), though this difference was not significant, $F(1, 98) = 2.25$, $MSe = .14$, $p = .14$. Participants also volunteered significantly more responses for control items (*M* = .68) than for deceptive items (*M* = .62), $F(1, 98) = 8.23$, $MSe = .18$. However, the main effect of item type is qualified by a significant interaction of item type and age, $F(1, 98) = 6.55$, $MSe = .15$. Specifically, whereas younger adults volunteered roughly the same proportion of responses for control and deceptive items, older adults volunteered significantly more responses for deceptive items (*M* = .67) than for control items (*M* = .56), $t(49) = -3.65$.

**Table 3. Means of Response Criteria, Fit Ratios, and Number of Responses Volunteered and Withheld At Confidence Levels Above and Below the Response Criteria, by Age and Item Type.**

| Variable | Item Type | Younger Adults | | Older Adults | |
|---|---|---|---|---|---|
| | | $P_A \geq P_{RC}$ | $P_A < P_{RC}$ | $P_A \geq P_{RC}$ | $P_A < P_{RC}$ |
| Mean Number of Responses Volunteered | Control | 12.76 | 0.82 | 10.38 | 0.74 |
| | Deceptive | 13.10 | 0.65 | 13.00 | 0.48 |
| Mean Number of Responses Withheld | Control | 0.48 | 6.08 | 0.52 | 8.08 |
| | Deceptive | 1.18 | 4.92 | 0.78 | 5.58 |
| $P_{RC}$ | Control | 0.55 | | 0.62 | |
| | Deceptive | 0.61 | | 0.57 | |
| Fit Ratio | Control | 0.94 | | 0.94 | |
| | Deceptive | 0.91 | | 0.94 | |

Per the recommendation of Koriat and Goldsmith (1996), response criterion ($P_{RC}$) was calculated by first assessing those values on the confidence scale, from 0 to 100%, at which participants chose to volunteer a response. Each value can be considered as a possible response criterion, such that participants would volunteer responses equal to or exceeding the criterion (defined as hits in this framework) and withhold those values below the criterion (defined as correct rejections in this framework). For each possible response criterion value, one can then calculate the total number of hits and correct rejections out of the total number of items possible, defined as the *fit ratio*. That value which results in the largest possible fit ratio is considered to be the response criterion. More succinctly, the response criterion is that value which maximizes the number of hits and correct rejections. Response criterion estimates were examined in a 2 (Item Type: control, deceptive) x 2 (Age: old, young) mixed-factor ANOVA[6]. Results showed that response criterion did not differ on the basis of age, $F$ (1, 97) < 1, or item type, $F$ (1, 97) < 1. However, an age x item type interaction was present, $F$ (1, 97) = 4.88, $MSe$ = .17. Specifically, older adults exhibited a more stringent criterion for volunteering responses for control items in comparison to deceptive items. Younger adults, in contrast, demonstrated the opposite pattern, as they maintained a stricter criterion for deceptive items than for control items.

***Discussion***

Overall, younger adults demonstrated greater accuracy than older adults, though age differences were most apparent for deceptive items. Older adults began with a lower quantity of correct items and did not achieve the level of accuracy evident for younger individuals. Both groups of participants showed striking differences in performance for control versus deceptive items. In particular, older and younger adults reported a lower quantity of correct responses on deceptive items and did not achieve comparable levels of accuracy at free report. A lower quantity of correct responses available at forced report does not necessarily dictate that accuracy will also be diminished. That is, monitoring processes may operate to achieve accuracy, such that participants may only volunteer those responses that are correct. However, while age differences in monitoring resolution and calibration were largely absent, both groups of participants were persistently overconfident for deceptive items and demonstrated significantly poorer monitoring resolution on such items.

**Psychometric Tests: Performance, Correlations, and Factor Loadings.**

Performance on each of the executive, speed, and memory tests administered is summarized in Table 4. Inspection of the table indicates that younger adults performed better on each of the tasks administered than older adults with the exception of the verbal fluency task (FAS). This is not inconsistent with the extant literature, as age differences are not always apparent on the task (for a review, see Spreen & Strauss, 1998). Correlations between performance on the various tests administered, accuracy, and monitoring measures are presented in Appendix A.

**Table 4. Means (and Standard Deviations) of Scores From Executive Functioning and Speed Measures for Older and Younger Adults.**

| Task | Younger Adults | Older Adults | *F (1, 98)* |
|---|---|---|---|
| Trails Making Test | | | |
| Version A time (*s*) | 23.08 (9.74) | 36.32 (11.50) | 38.58** |
| Version B time (*s*) | 53.60 (25.84) | 79.72 (27.34) | 24.11** |
| Difference | 30.52 (25.75) | 43.40 (21.06) | 7.50* |
| Wisconsin Card Sorting Test | | | |
| Categories Achieved | 5.98 (.14) | 4.37 (2.03) | 31.47** |
| Perseverative Errors | 7.44 (3.67) | 16.96 (12.62) | 26.18** |
| Verbal Fluency[a] | 42.55 (11.38) | 43.74 (12.53) | .24 |
| Operation Span[b] | 16.55 (7.30) | 12.40 (6.58) | 8.85* |
| Digit Symbol Substitution[c] | 69.47 (9.24) | 49.28 (11.74) | 90.13** |
| Number Comparison[d] | 54.18 (11.42) | 41.02 (11.13) | 33.73** |

* $p < .05$
** $p < .01$
[a] Scores represent the total number of unique words produced in 1 min for each of the three letters tested (*F*, *A*, and *S*). [b] Scores represent the total number of words produced in their entirety and in the correct sequence out of a total of 42 words presented. [c] Scores represent the number of correct responses made in 90 s. [d] Scores represent the total number of correct responses made in two sessions of 90 s.

Scores on the Trails A-B difference measure, and each of the other executive and speed measures (see Table 4) were subjected to factor analysis. Because the oblique solution resulted in only a moderate correlation ($r = .28$) the orthogonal solution was accepted. Two factors with eigenvalues exceeding 1.0 were obtained (see Table 5 for a summary). The two speed tests and the difference in Trails A and Trails B performance loaded on the first factor, accounting for 32% of the total variance. The number of categories achieved on the WCST, the number of perseverative errors committed on the WCST, and the number of unique words produced on the FAS test loaded on the second factor, accounting for an additional 26% of the total variance. The first factor was labeled the *speed composite*. While this label is certainly applicable in the context of the two speed measures it is not conventionally applied to the Trail Making Test. However, the Trail Making Test is a timed a task, in common with the speed measures, and optimal performance on the task is predicated on completing it as fast as possible for each portion. This, in conjunction its strong loading on factor 1, provides sufficient justification to include the difference in times on Parts A and B of the Trail Making Test in the speed composite. The second factor, composed of scores from the WCST and the FAS test, was labeled the *executive composite*. Weightings based on each factor were used to derive an individual speed composite measure and executive composite measure for each participant. These scores were used as predictors of individual differences in memory accuracy and memory monitoring, described below.

**Table 5. Loading Patterns From the Orthogonal Rotation of the Factor Analysis of Executive Scores.**

| Test | Factor 1 | Factor 2 |
|---|---|---|
| Trail Making Test: A-B Difference | **-.70** | -.09 |
| Digit Symbol Substitution | **.80** | .27 |
| Number Comparison | **.85** | .10 |
| WCST-Categories | .36 | **.76** |
| WCST-Perseverative Errors | -.36 | **-.78** |
| Verbal Fluency (FAS Test) | -.19 | **.63** |
| Operation Span | .30 | .42 |
| Eigenvalue | 3.00 | 1.09 |
| Variance Proportion | .32 | .26 |

| | Correlation | |
|---|---|---|
| | Factor 1 | Factor 2 |
| 1. Age | -.60 | -.30 |
| 2. Factor 1 | -- | .00 |

**Memory Quantity and Accuracy Models**

Individual differences in memory quantity and memory accuracy for both control and deceptive items were examined by first entering age as a predictor. Age was followed in the model by composite factor scores for speed of processing and executive function. If executive processes make a contribution to the attainment of quantity and accuracy, then scores from the executive composite measure should contribute unique explained variability beyond that explained by age and speed of processing. Data from regression analyses of memory quantity and memory accuracy are presented in Table 6. (Similar analyses confined only to older participants are presented in Appendix B.)  Inspection of these data indicates that, with the exception of free report accuracy for control items, age alone explained a significant proportion of the variability in performance.  The speed composite measure contributed little, if any, unique explained variability in performance.  However, in all cases, the executive composite measure contributed a significant portion of unique explained variability to quantity and accuracy measures for both control and deceptive items.

**Table 6.  Proportion of Variability and Regression Coefficients for Quantity and Accuracy Measures Accounted for by Age, Before Control of Speed and Executive Composites.**

| | Age Alone | | After speed composite | | After executive composite | |
|---|---|---|---|---|---|---|
| Measure | $R^2$ | $\beta$ | Change in $R^2$ | $\beta$ | Change in $R^2$ | $\beta$ |
| Forced Quantity | | | | | | |
| Control | .04* | -.20 | .00 | -.02 | .11** | .36 |
| Deceptive | .08** | -.28 | .00 | -.01 | .09** | .33 |
| Free Quantity | | | | | | |
| Control | .05* | -.22 | .00 | .05 | .08** | .31 |
| Deceptive | .10** | -.31 | .00 | .10 | .08** | .31 |
| Free Accuracy | | | | | | |
| Control | .01 | -.11 | .00 | -.01 | .17** | .45 |
| Deceptive | .09** | -.31 | .00 | .05 | .10** | .34 |

* $p < .05$
** $p < .01$
*Note*: All regression estimates reported are standardized.

These data suggest that executive processes explain variability in individual differences in memory quantity and memory accuracy beyond that explained by age and speed of processing alone.  However, one can make the case that the executive tasks employed depend on working memory (e.g., Hartman, Bolton, & Fehnel, 2001), such that their contributions to explaining quantity and accuracy solely operate through individual differences in working memory. Inspection of the pattern of correlations (Table 5) indicates that OSPAN scores were positively correlated with all quantity and accuracy measures.  Thus, to investigate the contribution of working memory to performance, additional regression models were examined that included OSPAN scores as a predictor.  These models were identical to those previously reported with two exceptions. First, given that the speed composite measure contributed virtually no explained variability to previous models, it was excluded from the current regression analyses. Second, scores on the OSPAN task were entered as a predictor prior to entering the executive composite. These data are summarized in Appendix C. Results showed that OSPAN scores contributed significant unique explained variability in performance on most measures, except for free and forced report quantity measures for deceptive items. In addition, even after controlling for OSPAN scores, the executive composite measure still accounted for a significant amount of explained variability in performance. In fact, the executive composite measure accounted for more variability in performance than OSPAN on all measures of memory with the exception of free report quantity for control items.

In conjunction with previous analyses that controlled for the contribution of the speed of processing, these data suggest that executive processes do predict individual differences in quantity and accuracy. An alternative manner in which to illustrate this point is to compare those participants with high and low scores on the executive composite measure. While splitting participants into categories based on a continuous variable is statistically questionable (Maxwell & Delaney, 1993), it does illustrate the relationship between executive measures and quantity and accuracy. Table 7 shows quantity and accuracy data for older participants divided on the basis of the executive composite scores into roughly upper and lower quartiles ($n$ = 13 per group). These data indicate that older participants classified as high on the executive composite measure exhibited performance comparable to that of younger adults (see Table 2). Quantity correct at forced report was subjected to a 2 (Executive Score: high, low) x 2 (Item Type: control, deceptive) mixed-factor ANOVA. Results showed that participants classified as high on the executive composite score produced a greater quantity of correct responses ($M$ = .50) than participants classified as low on the executive score (.33), $F$ (1, 24) = 7.27, $MSe$ = .01. A similar pattern of data was evident for quantity correct at free report $F$ (1, 24) = 5.30, $MSe$ = .29, and accuracy at free report, $F$ (1, 24) = 8.16, $MSe$ = .70. That is, participants classified as high on the executive composite score produced more correct responses at forced report and demonstrated greater accuracy at free report than participants classified as low on the executive composite score[7]. Thus, it appears that executive processes made a contribution to quantity correct and memory accuracy. However, executive processes may also be crucial to the monitoring of responses for correctness. The next analyses will therefore examine whether executive processes also make a contribution to memory monitoring.

**Table 7. Means (and Standard Deviations) of Quantity and Accuracy Scores for the Free and Forced Report Condition by Item Type and Executive Composite Score for Older Participants (N = 26).**

| | | Report Option | | |
|---|---|---|---|---|
| Executive | | Forced | Free | |
| Composite | Item Type | Quantity and Accuracy | Quantity | Accuracy |
| High | Control | .63 (.17) | .55 (.21) | .88 (.18) |
| | Deceptive | .36 (.20) | .33 (.18) | .55 (.28) |
| Low | Control | .45 (.21) | .38 (.23) | .63 (.21) |
| | Deceptive | .22 (.09) | .20 (.10) | .34 (.28) |

**Memory Monitoring Models**

Before examining predictors of memory monitoring, a composite score was first derived by subjecting monitoring data for control and deceptive items (ANDI, calibration error, $\gamma$ correlations) to principle component factor analysis. The oblique solution resulted in a modest correlation ($r$ = .18) and thus the orthogonal solution was accepted. A summary of the data is presented in Table 8. Two factors with eigenvalues exceeding 1.0 were obtained. All measures of monitoring for deceptive items loaded on the first factor, accounting for 34% of the total variance. Likewise, all measures of monitoring for control items loaded on the second factor, accounting for 33% of the total variance. Thus, the first factor was labeled the *deceptive monitoring composite* and the second factor the *control monitoring composite*. Weightings based on each factor were used to derive monitoring scores for each type of item.

**Table 8. Loading Patterns From the Orthogonal Rotation of the Factor Analysis of Monitoring Scores.**

| Test | Factor 1 | Factor 2 |
|---|---|---|
| γ Correlation - Deceptive | **.72** | -.07 |
| ANDI - Deceptive | **.84** | .14 |
| Calibration Error - Deceptive | **-.86** | -.23 |
| γ Correlation - Control | -.04 | **.73** |
| ANDI - Control | .09 | **.93** |
| Calibration Error - Control | -.20 | **-.73** |
| Eigenvalue | 2.47 | 1.53 |
| Variance Proportion | .34 | .33 |
| | Correlation | |
| | Factor 1 | Factor 2 |
| 1. Age | -.16 | -.05 |
| 2. Factor 1 | -- | .00 |

Monitoring composite scores for control and deceptive items were subjected to the same regression analyses employed for memory quantity and accuracy. That is, age was entered first in the model, followed by the speed composite score and the executive composite score. These data are presented in Table 9. In general, the models explained a modest amount of the variability in monitoring performance, with executive scores only accounting for a significant amount of the variability in monitoring performance for deceptive items ($R^2 = .04$)[8]. Thus, on the whole, executive processes made a modest contribution to explaining variability in memory monitoring performance, and did so primarily for deceptive items.

**Table 9. Proportion of Variability and Regression Coefficients for Memory Monitoring Measures Accounted for by Age, Before Control of Speed and Executive Composites.**

| Monitoring Composite | Age Alone | | After speed composite | | After executive composite | |
|---|---|---|---|---|---|---|
| | $R^2$ | $\beta$ | Change in $R^2$ | $\beta$ | Change in $R^2$ | $\beta$ |
| Control Items | .00 | -.05 | .00 | .06 | .03 | .18 |
| Deceptive Items | .03 | -.16 | .02 | .15 | .04* | .22 |

* $p < .05$
*Note*: All regression estimates reported are standardized.

**Mediational Models of Executive Function: The Role of Monitoring**

Thus far, the data indicate that a relationship exists between scores on the executive composite measure and free report accuracy. This relationship may simply be the result of a direct relationship between accuracy and executive functions (see Figure 1, part a). However, the achievement of accuracy likely depends in part on accurate memory monitoring. In this sense, any effect of executive functions on accuracy is accounted for by monitoring accuracy (see Figure 1, part b). As noted, monitoring has been largely neglected in investigations of individual differences in memory performance (e.g., McDaniel et al., 1999; Mead & Roediger, 2002, November), with only the direct relationship between executive processes and accuracy used to account for individual differences. Scores on the executive composite measure were at least moderately related to memory monitoring, particularly for deceptive items, and may in part mediate the relationship between executive scores and accuracy measures. Therefore, the current analysis examined whether the relationship between executive scores and accuracy was mediated completely or partially by the accuracy of memory monitoring[9].

The procedure for mediational analyses followed that prescribed by Baron and Kenny (1986). In the first step, the independent variable (the executive composite score) was used to predict the dependent variable (accuracy at free report). In the second step, the independent variable was used to predict the mediating variable (the monitoring composite score). Third, the mediating variable (the monitoring composite score) and the independent variable (the executive composite score) was used to predict free report accuracy. Finally, mediation was examined by assessing whether the mediator reduced the predictive value of the independent variable. In order to achieve mediation, the regression of the independent variable to the dependent variable and the mediator (steps 1 and 2) must be significant and the regression of the mediator to the dependent variable (step 3) must be significant. When these steps are satisfied, partial or complete mediation of the independent variable will typically be present (step 4).

*Mediation of deceptive items*. Data for deceptive items was first examined and indicated that monitoring partially mediated accuracy at free report. Specifically, results showed that the first step was met, as the executive composite measure predicted free report accuracy for deceptive items ($\beta = .37, p < .01$). The second step was also met, as the executive composite measure was a significant predictor of the mediator, the deceptive monitoring composite ($\beta = .27, p < .01$). When the executive composite score and the monitoring composite score for deceptive items (i.e., the mediator) were included as predictors, the resulting regression estimate was significant (β = .60, $p < .01$; Step 3 met). The mediator in turn reduced the executive composite predictor of free report accuracy for deceptive items (β = .21, $p < .05$). A follow-up Sobel test confirmed that the predictive value of the executive composite score decreased significantly when monitoring was introduced as a mediator (Sobel = 2.67, $p < .05$; Step 4 met). Because all steps were met, Baron and Kenny's (1986) requirements for establishing mediation were satisfied.

Given that the executive composite measure continued to be a significant predictor of accuracy after introducing the monitoring composite (i.e., the mediator), the current data indicate that the monitoring composite partially mediated accuracy at free report. Shrout and Bolger (2002) suggest that partial mediation is best described in terms of the proportion of the original effect mediated. Thus, proportion mediated ($P_M$) was calculated by dividing the products of the coefficients produced in steps 2 and 3 by the original coefficient of prediction evident in step 1 (Shrout & Bolger, 2002). Results showed that 43% of the effect of the executive composite score on accuracy at free report for deceptive items was mediated by the monitoring composite. Thus, while monitoring did not completely mediate free report accuracy for deceptive items, it did account for nearly half of the effect of the executive composite score on accuracy[10].

*Mediation of control items*. The role of monitoring as a mediator of free report accuracy for control items was examined in a manner identical to that for deceptive items. The only exceptions were that free report accuracy for control items was included as the dependent variable and the monitoring composite for control items was entered as the mediator. Results showed that the first step was met, as the executive composite score was a significant predictor of accuracy at free report, (β = .42, $p <.01$). However, the second step was not met, as the executive composite was not a significant predictor of the monitoring composite measure, (β = .16, $p = .16$)[11]. Thus, these data suggest that monitoring mediated the relationship between executive scores and accuracy for deceptive items, but not for control items.

# GENERAL DISCUSSION

The current study examined individual differences in memory accuracy in older and younger adults. Consistent with Kelley and Sahakyan (2003), younger adults demonstrated greater quantity correct and greater accuracy than older adults, particularly for deceptive items. In contrast to Kelley and Sahakyan, differences on monitoring measures were far smaller, with younger adults only showing better calibration for deceptive items than older adults. Analyses of the relationship between performance on measures of executive function (reduced to a composite score) and measures of memory quantity, accuracy, and monitoring produced several notable findings. First, for quantity and accuracy measures, the executive composite explained unique variability in performance not explained by age and speed of processing. This relationship held even when controlling for scores on the OSPAN task. Second, whereas age, speed, and the executive composite measure did not explain unique variability in the monitoring composite score for control items, it did explain a significant amount of the variability in the monitoring composite measure for deceptive items. Further, the data indicated that the relationship between the executive composite measure and free report accuracy for deceptive items was partially mediated by memory monitoring, accounting for nearly half of the relationship. Monitoring did not mediate the relationship between the executive composite and accuracy for control items.

These data provide support for the major assumptions tested in the current study, namely that a) executive functions contribute to the quantity of accurate information available at retrieval and that b) executive functions operate in part by memory monitoring. Executive functions clearly explained unique variability in the quantity of correct responses available at forced report. Specifically, results showed that from 9 to 17% of the variability in quantity was explained by executive measures after accounting for age and speed of processing. Thus, there was a relationship between the level of memory available at forced report and executive functions.

Results from mediational analyses also provided support for the second assumption, largely ignored in previous work, that executive processes operate in part through monitoring of candidate responses. However, mediation was only apparent for deceptive items. One possible explanation for this is that achieving accuracy for deceptive items was a far more demanding task than that required for control items. For example, any gains in accuracy at free report for deceptive items will hinge on discerning those items that are correct versus those that are plausible, but incorrect responses (e.g., "doctor" in the case of the cue, NURSE DO_ _ _ R). In this case, monitoring will be a far more difficult task and any possible gains in accuracy will demand effective memory monitoring. The greater difficulty in monitoring deceptive items is evident in the fact that participants demonstrated high levels of overconfidence for such items. In contrast, control items (e.g., CLOCK DO _ _ _ R) were far less demanding of monitoring processes, as participants were less likely to retrieve plausible but incorrect competitors. That is, control items were far less likely to elicit plausible alternatives, making monitoring an easier task, as evidenced by the strong relationship between assessed confidence and the probability that a response was correct.

Mather, Johnson, and De Leonardis (1999) have offered a similar explanation for the finding that frontal (executive) measures are inconsistent predictors of source monitoring accuracy. Specifically, several studies have shown that frontal measures are predictive of source accuracy (e.g., Glisky et al. 1995; 2001; Janowsky et al., 1989a) while others have failed to detect a relationship (e.g., Dywan et al., 1994; Henkel et al., 1998; Johnson et al., 1995). Mather et al. (1999) suggest that source monitoring tasks differ in the degree to which they depend on binding of features of the initial event and later evaluation of those features. Given that the frontal lobes (and executive processes) are construed as playing a major role in generating appropriate cues during retrieval (Burgess & Shallice, 1996; Norman & Schacter, 1996) and evaluating the contents of cognitive processes (Shimmamura, 1996; 2000a; 2000b) they will be important in any demanding source monitoring task. Specifically, Mather et al. suggest that the extent of frontal lobe involvement will depend on the degree to which the task requires "reflective processing", with greater demands for reflective processing leading to greater dependence on frontal processes. Such an explanation is consistent with the finding from the current study that the monitoring composite score was a substantial mediator of executive function only for deceptive items. In this case, retrieval of deceptive items likely required greater reliance on monitoring to achieve accuracy. This in turn suggests that any relationship between executive processes and performance in paradigms that facilitate false recall of easily available, nonstudied items (such as the DRM paradigm; e.g., Meade & Roediger, 2002) may be mediated in part by monitoring efficacy.

Kelley and Sahakyan (2003) contend that monitoring will depend largely on the quality of information available to the monitoring process. That is, monitoring will be more difficult if participants retrieve vague or undifferentiated information. Results from the current study showed that monitoring resolution, as reflected by $\gamma$ correlations and ANDIs, was significantly better for control items than for deceptive items. This presumably occurred because participants were less likely to retrieve accurate information for deceptive items, and may instead have only had available fluently generated but incorrect items. As noted, monitoring scores only partially mediated the effect of executive processes on accuracy for deceptive items. The remaining variability may thus be a function of the quality of information available to participants. The issue is further compounded when one considers that items given at forced report may have already undergone at least a rudimentary form of monitoring for accuracy. For example, an older participant remarked during forced recall of a deceptive item that she was sure that a potential target was not semantically related to the cue and instead reported a different item that was not related. Thus, monitoring at free report may capture only a portion of the full monitoring process during retrieval, rendering quantity correct at forced report in part dependent on an initial monitoring process. This at least suggests that role of monitoring in the current study may have been underestimated and certainly demands future investigation.

Although the current study has focused primarily on the role of the frontal lobe in memory vis a vis executive measures, processes localized in the medial temporal region undoubtedly contribute to memory accuracy. For example, Davidson and Glisky (2002) have demonstrated that measures of recollection are sensitive to both frontal and medial temporal functioning. Medial-temporal lobe scores may also be predictive of source memory (Henkel et al., 1998; Mather et al., 1999) and it has long been established that medial temporal lesions are associated with amnesia (e.g., Scoville & Milner, 1957; Cohen & Squire, 1980). The current study did not include a medial temporal battery of tests but controlling for variability associated with working memory span did reduce the proportion of variability explained by executive measures. Does this mean that a battery medial temporal measures would account for all variability in accuracy associated with executive measures? Evidence from neuroimaging studies suggests that along with medial temporal regions, frontal regions are important for both encoding (see Buckner, Logan, Donaldson, & Wheeler, 2000, for a review) and retrieval (e.g., Rugg & Henson, 2001) operations[12]. For example, Fletcher, Shallice, and Dolan (1998) demonstrated that dorsolateral frontal activity was greatest when participants were required to use an organizational strategy during encoding. Others have shown increases in frontal activity when retrieval tasks demand recollection of details (see Yonelinas, 2002, for a review). Medial temporal and frontal regions may in fact work in complementary fashion, such that frontal regions play a large part in determining encoding strategies, with encoded information stored in medial temporal regions. Frontal regions may then provide support during retrieval, particularly when recollection of details is required. Thus, there is reason to suggest that frontal processes make a contribution to memory in addition to that made by medial temporal regions.

One anomalous finding from the current study concerns data from speed of processing tasks. Specifically, data from these tasks in general added little or no predictive power to regression analyses of accuracy and monitoring performance. Similar results have been reported in the memory literature when speed has been used a predictor of encoding strategies (Bryan et al., 1999; Dunlosky & Hertzog, 1998). However, this may be contrasted with findings from a number of other studies indicating substantial correlations between speed of processing and performance on a variety of cognitive tasks (see Salthouse, 1996, for a review). Part of this may be a function of the older adults tested in the current study. Specifically, a number of the older adults tested in the current study were very high functioning (see Table 1) and had, on average, some post-graduate education. Thus, many of the older adults tested, while scoring well below younger adults on measures of speed of processing, exhibited scores comparable to younger adults on accuracy measures. Certainly, younger adults demonstrated greater accuracy than older adults on all memory measures but the existence of a number of older adults performing well on accuracy measures may in part explain the null finding for speed. This explanation may be further extended to monitoring data, where age differences were not apparent, in contrast to previous work (i.e., Kelley & Sahakyan, 2003). As a consequence, speed of processing, typically sensitive to age differences, had little predictive power. One may thus contend that a more rigorous test of speed of processing as a predictor of memory accuracy should include older adults from a more diverse population. At the very least, the nature of the sample suggests that caution should be exercised when interpreting data from speed tasks.

While the current study has attempted to clarify the role of executive functions in memory accuracy any attempt at clarifying the role of executive functions is certainly formidable. At the most fundamental level, executive functions are typically construed as serving a supervisory role in cognition (e.g., Baddeley, 1996; Norman & Shallice, 1986) that is often linked to the frontal lobe. Much of this link to a specific and highly complex neural substrate such as the frontal lobe has come from evidence showing that patients with frontal lesions are exceptionally poor at tasks that demand planning (e.g., Penfield & Evans, 1935; Shallice & Burgess, 1991) or inhibition of automatic responses (Lhermitte, 1983). The issue becomes more muddled when one considers the diversity of functions that have been deemed executive functions (e.g., Miyake et al., 2000) and the task of accurately measuring those functions (cf., Salthouse, Atkinson, & Berish, 2003). The current study attempts to pinpoint two possible components of the relationship between executive function and memory, but greater clarification of the relationship is necessary.

In conclusion, results from the current study indicate that measures of executive functioning are predictive of memory accuracy. While previous studies have examined this relationship based on a single measure of accuracy (e.g., McDaniel et al., 1999; Meade & Roediger, 2002) the current study permitted an assessment of both the accuracy of the items produced during retrieval and monitoring processes. Results showed that executive processes consistently explained unique variability in both quantity and accuracy. Examination of memory monitoring scores indicated that executive processes had their effect on accuracy in part through memory monitoring, particularly for deceptive items that were quite demanding of monitoring processes. Thus, both the accuracy of retrieved responses and memory monitoring must be taken into consideration when individual differences in memory accuracy are examined.

## FOOTNOTES

[1]For 1.3% of the items, participants changed their response between forced and free report. These items were excluded from the monitoring data reported.

[2]Further analysis of confidence scores showed that, overall, younger adults had a mean confidence level of .87 for volunteered items and .18 for items that were withheld. Older adults exhibited a mean confidence level of .85 for volunteered items and .25 for withheld items.

[3]Younger and older adults exhibited equal levels of polarization (i.e., instances of using extreme confidence categories of 0 and 100%). Specifically, younger and older adults used the categories of 0 and 100% for control items, 66% and 67% of the time, respectively. The pattern is similar for deceptive items, as younger adults showed polarized responding 61% of the time compared to 62% of the time for older adults.

[4]Calculation of ANDI scores occasionally resulted in cases in which the ANDI score was negative. Yaniv et al. (1991) suggest that such scores be converted to 0 primarily because, as a measure of the proportion of variability in correctness accounted for by confidence judgments, negative values are largely uninterpretable. Therefore, all negative values were converted to 0. Analyses which included negative values resulted in no discernable difference in the data reported.

[5]In some cases, $\gamma$ correlations could not be calculated, as in instances when participants either did not withhold or volunteer any responses. This occurred in 6 cases for older participants for deceptive items and once for older participants for control items. In addition, this occurred in 2 cases for younger participants for deceptive items and once for control items.

[6]One participant did not volunteer any responses for deceptive items and was excluded from the analysis.

[7]It must be noted that these data were also examined with OSPAN held as a covariate, as regression analyses had indicated that OSPAN did account for unique variability in accuracy and quantity correct (see Appendix B). Results showed that the magnitude of the main effect of the executive composite was diminished when OSPAN was introduced as a covariate. For example, the classification based on the executive composite score was only marginally significant when covaried with OSPAN for forced quantity correct, $F$ (1, 23) = 3.50, $MSe$ = .17, $p$ = .07. The same pattern of a diminished effect of the executive composite score was likewise evident for free quantity correct, $F$ (1, 23) = 2.16, $MSe$ = .12, $p$ = .16, and free accuracy, $F$ (1, 23) = 2.96, $MSe$ = .23, $p$ = .10. However, such null findings may in part be due to the relative lack of power that resulted from confining analyses to small subgroups of participants. Examination of effect sizes (Cohen's $d$) also indicates that relatively largely effect sizes are still present when OSPAN is held as a covariate. For example, both forced report quantity correct ($d$ = .78) and free report accuracy ($d$ = .72) are characterized by relatively large effect sizes when OSPAN is entered as a covariate.

[8]The monitoring composite scores derived included two measures of monitoring resolution (i.e., γ correlations and ANDIs). As noted, γ correlations were somewhat skewed and were often subject to ceiling effects, such that participants frequently attained the maximum score (1.0). An additional analysis was conducted in which factor scores were derived for monitoring without including γ correlations. The specific data are not presented, but factors with eigenvalues exceeding 1.0 were once again attained that resulted in two different factors: One reflecting monitoring of deceptive items and one reflecting monitoring of control items. Results showed that after controlling for age and speed of processing, executive composite scores accounted for approximately 7% of the variability in monitoring for deceptive items. Excluding γ correlations did not improve the proportion of explained variability in monitoring for control items.

[9]Models were also tested that examined whether monitoring functioned as a moderater (rather than a mediator) of the relationship between executive processes and accuracy. In the familiar parlance of ANOVA, moderators are apparent when an interaction exists between variables of interest (see Baron and Kenny, 1986, for a more detailed discussion). Thus, an interaction term (the cross products of raw scores) was derived for the analysis of the deceptive monitoring composite and the executive composite. This term was regressed on accuracy for

deceptive items, after controlling for the individual effects of the deceptive monitoring composite and the executive composite. Results showed that the interaction term was not significant, $t$ (95) < 1. An identical analysis for control items likewise revealed that an interaction was not present, $t$ (95) = -1.62, $p$ = .11.

[10]Monitoring composites that excluded γ correlations (see Footnote 8) were also examined in a mediational analysis of deceptive items, identical to that just described. All steps were met in this analysis and results showed partial mediation was once again present and that $P_M$ = .56. Thus, exclusion of γ correlations made monitoring a stronger mediator of free report accuracy for deceptive items.

[11]This was the case even when γ correlations were excluded from the monitoring composite score (see Footnote 8).

[12]Interestingly, patients with frontal lobe lesions are also less likely to use mnemonic strategies that may facilitate later recollection (e.g., Hirst, 1982). For example, examinations of frontal patients' free recall reports indicates that they are less likely to encode and retrieve words in an organized fashion (Eslinger & Grattan, 1994; Gershberg & Shimmamura, 1995).

APPENDIX A.

CORRELATION MATRIX

.

| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. | Age | 1.0 | | | | | | | | | |
| 2. | Trails A-B Difference | .29** | 1.0 | | | | | | | | |
| 3. | WCST-Categories | -.52** | -.30** | 1.0 | | | | | | | |
| 4. | WCST-Persev Errors | .45** | -.34** | -.80** | 1.0 | | | | | | |
| 5. | Verbal Fluency (FAS) | .03 | -.03 | .14 | -.17 | 1.0 | | | | | |
| 6. | Operation Span | -.32** | -.18 | .26* | -.26** | .17 | 1.0 | | | | |
| 7. | DSST | -.70** | -.39** | .41** | -.40** | .15 | .36** | 1.0 | | | |
| 8. | Number Comparison | -.51** | -.43** | .34** | -.34** | .07 | .24* | .70** | 1.0 | | |
| 9. | Forced Quan-Control | -.20* | -.13 | .19 | -.29** | .32** | .31** | .21* | .18 | 1.0 | |
| 10. | Forced Quan-Decep | -.26** | -.12 | .31** | -.26** | .29** | .22* | .28** | .21* | .63** | 1.0 |
| 11. | Free Quan-Control | -.21* | -.09 | .20* | -.23* | .27** | .31** | .20* | .17 | .93** | .66** |
| 12. | Free Quantity-Decep | -.29** | -.10 | .32** | -.27** | .28** | .22* | .31** | .25* | .62** | .98** |
| 13. | Free Accuracy-Control | -.12 | -.17 | .28** | -.30** | .28** | .32** | .15 | .08 | .71** | .46** |
| 14. | Free Accuracy-Decep | -.29** | -.13 | .31** | -.25* | .30** | .30** | .34** | .26** | .59** | .87** |
| 15. | γ Correlation-Control | -.05 | -.11 | .14 | -.01 | .01 | .10 | .09 | .06 | .15 | .21* |
| 16. | γ Correlation-Deceptive | -.10 | .13 | .08 | -.04 | -.04 | .15 | .17 | .00 | .12 | .23* |
| 17. | Calibration Error-Control | .03 | -.03 | .15 | .09 | -.02 | -.18 | -.04 | .04 | -.26** | .22* |
| 18. | Calibration Error-Decep | .17 | .04 | -.26** | .23* | -.31** | -.25* | -.27** | -.13 | -.45** | -.67** |
| 19. | ANDI - Control | -.10 | .01 | .15 | -.07 | .11 | .21* | .11 | .03 | .25** | .29** |
| 20. | ANDI - Deceptive | -.12 | .03 | .12 | -.08 | .20* | .21* | .05 | -.11 | .34** | .40** |

*Note*. WCST = Wisconsin Card Sorting Test; DSST = Digit-Symbol Substitution Task; Quan = Quantity; Decep = Deceptive; ANDI = Adjusted Normalized Discrimination Index
* $p < .05$ ** $p < .01$

| | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
|---|---|---|---|---|---|---|---|---|---|
| 1. Age | | | | | | | | | |
| 2. Trails A-B Difference | | | | | | | | | |
| 3. WCST-Categories | | | | | | | | | |
| 4. WCST-Persev Errors | | | | | | | | | |
| 5. Verbal Fluency (FAS) | | | | | | | | | |
| 6. Operation Span | | | | | | | | | |
| 7. DSST | | | | | | | | | |
| 8. Number Comparison | | | | | | | | | |
| 9. Forced Quan-Control | | | | | | | | | |
| 10. Forced Quan-Decep | | | | | | | | | |
| 11. Free Quan-Control | 1.0 | | | | | | | | |
| 12. Free Quantity-Decep | .68** | 1.0 | | | | | | | |
| 13. Free Accuracy-Control | .69** | .46** | 1.0 | | | | | | |
| 14. Free Accuracy-Decep | .60** | .87** | .55** | 1.0 | | | | | |
| 15. γ Correlation-Control | .27** | .23* | .38** | .26* | 1.0 | | | | |
| 16. γ Correlation-Deceptive | .19 | .28** | .02 | .33** | .10 | 1.0 | | | |
| 17. Calibration Error-Control | -.32** | -.22* | -.29** | -.28** | .22* | -.04 | 1.0 | | |
| 18. Calibration Error-Decep | -.41** | -.66** | -.41** | -.79** | -.16 | -.46** | .27** | 1.0 | |
| 19. ANDI - Control | .38** | .29** | .42** | .31** | .59** | .01 | -.63** | -.27** | 1.0 |
| 20. ANDI - Deceptive | .30** | .41** | .35** | .55** | .02 | .35** | -.29** | -.68** | .23* |

*Note*. WCST = Wisconsin Card Sorting Test; DSST = Digit-Symbol Substitution Task; Quan = Quantity; Decep = Deceptive; ANDI = Adjusted Normalized Discrimination Index
* $p < .05$ ** $p < .01$

# APPENDIX B.

# PROPORTION OF VARIABILITY AND REGRESSION COEFFICIENTS FOR QUANTITY AND ACCURACY MEASURES—OLDER PARTICIPANTS.

| Measure | Speed Alone $R^2$ | Speed Alone $\beta$ | After executive composite Change in $R^2$ | After executive composite $\beta$ |
|---|---|---|---|---|
| Forced Quantity | | | | |
| Control | .02 | .15 | .08* | .28 |
| Deceptive | .03 | .19 | .08* | .29 |
| Free Quantity | | | | |
| Control | .01 | .12 | .06 | .24 |
| Deceptive | .05 | .22 | .07† | .26 |
| Free Accuracy | | | | |
| Control | .04 | .19 | .17** | -.41 |
| Deceptive | .04 | .20 | .09* | .30 |

* $p < .05$
** $p < .01$
† $p = .07$
*Note*: All regression estimates reported are standardized.

# APPENDIX C.

# PROPORTION OF VARIABILITY AND REGRESSION COEFFICIENTS FOR QUANTITY AND ACCURACY MEASURES AFTER CONTROLLING FOR AGE, OPERATION SPAN (OSPAN) SCORE, AND EXECUTIVE FUNCTION

| Measure | Age Alone $R^2$ | Age Alone $\beta$ | After OSPAN score Change in $R^2$ | After OSPAN score $\beta$ | After executive composite Change in $R^2$ | After executive composite $\beta$ |
|---|---|---|---|---|---|---|
| Forced Quantity | | | | | | |
| Control | .04* | -.20 | .06* | .26 | .07** | .29 |
| Deceptive | .08** | -.28 | .02 | .14 | .07** | .30 |
| Free Quantity | | | | | | |
| Control | .05* | -.22 | .06* | .26 | .04* | .24 |
| Deceptive | .10** | -.31 | .02 | .14 | .06* | .28 |
| Free Accuracy | | | | | | |
| Control | .01 | -.11 | .09** | .31 | .10** | .36 |
| Deceptive | .09** | -.31 | .04* | .21 | .05* | .26 |

* $p < .05$
** $p < .01$
*Note*: All regression estimates reported are standardized.

APPENDIX D.

IRB APPROVAL

Florida State
UNIVERSITY

Office of the Vice President
For Research
Tallahassee, Florida 32306-2763
(850) 644-8673 · FAX (850) 644-4392

# APPROVAL MEMORANDUM

Human Subjects Committee

Date: 10/7/2003

**Matthew Rodes/ Colleen Kelley**
**Mc 1270**

Dept.: **Psychology**

From: **David Quadagno, Chair** DQ/ph

Re: **Use of Human Subjects in Research**
**Memory Monitoring and Aging**

The forms that you submitted to this office in regard to the use of human subjects in the proposal referenced above have been reviewed by the Secretary, the Chair, and two members of the Human Subjects Committee. Your project is determined to be exempt per 45 CFR § 46.101(b) 2 and has been approved by an accelerated review process.

**The Human Subjects Committee has not evaluated your proposal for scientific merit, except to weigh the risk to the human participants and the aspects of the proposal related to potential risk and benefit. This approval does not replace any departmental or other approvals, which may be required.**

If the project has not been completed by **10/6/2004** you must request renewed approval for continuation of the project.

You are advised that any change in protocol in this project must be approved by resubmission of the project to the Committee for approval. Also, the principal investigator must promptly report, in writing, any unexpected problems causing risks to research subjects or others.

By copy of this memorandum, the chairman of your department and/or your major professor is reminded that he/she is responsible for being informed concerning research projects involving human subjects in the department, and should review protocols of such investigations as often as needed to insure that the project is being conducted in compliance with our institution and with DHHS regulations.

This institution has an Assurance on file with the Office for Protection from Research Risks. The Assurance Number is IRB00000446.

Cc: Colleen Kelley
HSC No. 2003.511

APPENDIX E.

INFORMED CONSENT FOR UNDERGRADUATE PARTICIPANTS

# INFORMED CONSENT FORM
For Undergraduate Participants

I freely and voluntarily consent to be a participant in the research project entitled "Memory Monitoring and Aging".

This project is being conducted by Matthew Rhodes who is a graduate student in the Department of Psychology at Florida State University.

The experiment does not in any way constitute a risk to me. I will receive course credit for this experiment.

I will be studying and attempting to remember lists of words and/or nonwords. My responses will be recorded on a computer in a particular data file, but no record will be kept of whose data corresponds to which data file. In another task, I will try to think of words that begin with certain letters. In other tasks, I may be asked to arrange discs in an alignment specified by the experimenter, to name the color of words and letters presented to me, or to trace a pattern specified by the experimenter. In another short task, I will see pictures of geometric figures and try to guess which one follows a rule the experimenter has in mind. I will be told when I am correct and incorrect on this task. I will also answer a demographic questionnaire (age, education, and health status), and fill out a vocabulary test. These records will be kept locked and accessible only to Dr. Kelley. I understand that no individual results of this experiment will be reported and that my results will be kept confidential to the extent allowed by law.

I understand that these tasks will involve pencil and paper as well as computer-based testing and will take approximately 3 hours to complete.

I may contact Mr. Rhodes (644-9873) if I have questions about this project.

My consent may be withdrawn at any time without prejudice, penalty, or loss of benefits to which I am otherwise entitled. That is, my grade in the course will not be affected if I choose to withdraw from the experiment, nor will I receive an experiment credit penalty. However, I will still be obliged to fulfill my experiment participation obligation for the General Psychology course.

I have read and understand this consent form, and I am 18 years or older.

_______________________________ (Participant)

_______________________________ (date)

FLORIDA STATE UNIVERSITY
Approved: 10.7.03
HS No. 03.511
Void After 10/6/04
INSTITUTIONAL REVIEW BOARD

APPENDIX F.

INFORMED CONSENT FOR COMMUNITY PARTICIPANTS

Rhodes -- IRB

INFORMED CONSENT FORM
For Community Participants

I freely and voluntarily consent to be a participant in the research project entitled "Memory Monitoring and Aging".

This project is being conducted by Matthew Rhodes who is a graduate student in the Department of Psychology at Florida State University.

The experiment does not in any way constitute a risk to me. I also understand that I will receive $30 contingent on completing the experiment.

I will be studying and attempting to remember lists of words and/or nonwords. My responses will be recorded on a computer in a particular data file, but no record will be kept of whose data corresponds to which data file. In another task, I will try to think of words that begin with certain letters. In other tasks, I may be asked to arrange discs in an alignment specified by the experimenter, to name the color of words and letters presented to me, or to trace a pattern specified by the experimenter. In another short task, I will see pictures of geometric figures and try to guess which one follows a rule the experimenter has in mind. I will be told when I am correct and incorrect on this task. I will also answer a demographic questionnaire (age, education, and health status), and fill out a vocabulary test. These records will be kept locked and accessible only to Dr. Kelley. I understand that no individual results of this experiment will be reported and that my results will be kept confidential to the extent allowed by law.

I understand that these tasks will involve pencil and paper as well as computer-based testing and will take approximately 3 hours to complete.

I may contact Mr. Rhodes (644-9873) if I have questions about this project.

My consent may be withdrawn at any time without prejudice, penalty, or loss of benefits to which I am otherwise entitled

I have read and understand this consent form, and I am 18 years or older.

______________________________ (Participant)

__________________ (date)

FLORIDA STATE UNIVERSITY
Approved: 10.7.03
HS No. 03-511
Void After 10/6/04
INSTITUTIONAL REVIEW BOARD

REFERENCES

Anastasi, J. S., Rhodes, M. G., & Burns, M. C. (2000). Distinguishing between memory illusions and actual memories utilizing phenomenological measurements and explicit warnings. *American Journal of Psychology, 113*, 1-26.

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

Balota, D. A., Dolan, P. O., & Duchek, J. M. (2000). Memory changes in healthy older adults. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford Handbook of Memory* (pp. 395-409). Oxford: Oxford University Press.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173-1182.

Bartlett, J. C., Halpern, A. R., & Dowling, W. J. (1995). Recognition of familiar and unfamiliar melodies in normal aging and Alzheimer's disease. *Memory & Cognition, 23*, 531-546.

Benton, K. L., & Hamsher, K. (1976). *Multilingual Aphasia Examination manual*. Iowa City: University of Iowa.

Blair, J. R., & Spreen, O. (1989). Predicting premorbid IQ: A revision of the National Adult Reading Test. *The Clinical Neuropsychologist*, *3*, 129-136.

Bryan, J., Luszcz, M. A., & Pointer, S. (1999). Executive function and processing resources as predictors of adult age differences in the implementation of encoding strategies. *Aging, Neuropsychology*, *and Cognition*, *6*, 273-287.

Buckner, R. L., Logan, J., Donaldson, D. I., & Wheeler, M. E. (2000). Cognitive neuroscience of episodic memory encoding. *Acta Psychologia*, *105*, 127-139.

Burgess, P. W., & Shallice, T. (1996). Confabulation and the control of recollection. *Memory*, *4*, 359-411.

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

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

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

Cohen, N. J., & Squire, L. R. (1980). Preserved learning and retention of pattern-analyzing skill in amnesia: Dissociation of knowing how and knowing that. *Science*, *210*, 207-210.

Connor, L. T., Dunlosky, J., & Hertzog, C. (1997). Age-related differences in absolute but not relative metamemory accuracy. *Psychology and Aging*, *12*, 50-71.

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, L. W., Morris, R. G., Loewen, E. R. (1990). Relations between source amnesia and frontal lobe functioning in older adults. *Psychology and Aging, 5*, 148-151.

Crawford, J. R., Bryan, J., Luszcz, M. A., Obonsawin, M. C., & Stewart, L. (2000). The executive decline hypothesis of cognitive aging: Do executive deficits qualify as differential deficits and do they mediate age-related memory decline? *Aging, Neuropsychology, and Cognition*, *7*, 9-31.

Davidson, P. S. R., & Glisky, E. (2002). Neuropsychological correlates of recollection and familiarity in normal aging. *Cognitive, Affective, & Behavioral Neuroscience*, *2*, 174-186.

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

Dodson, C. S., & Schacter, D. L. (2003) Aging and strategic retrieval processes: Reducing false memories with a distinctiveness heuristic. *Psychology and Aging*, *17*, 405-415.

Duncan, J. (1993). Selection of input and goal in the control of behavior. In A. Baddeley & L. Weiskrantz (Eds.), *Attention: Selection, awareness, and control. A tribute to Donald Broadbent* (pp. 53-71). Oxford: Oxford University Press, Clarendon Press.

Duncan, J. (1995). Attention, intelligence, and the frontal lobes. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 721-733). Cambridge, MA: MIT Press.

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

Dywan, J., & Jacoby, L. L. (1990). Effects of aging on source monitoring: Differences in susceptibility to false fame. *Psychology and Aging, 5*, 379-387.

Dywan, J., Segalowitz, S. J., & Williamson, L. (1994). Source monitoring during name recognition in older adults: Psychometric and electrophysiological correlates. *Psychology and Aging*, *9*, 568-577.

Eslinger, P. J., & Grattan, L. M. (1994). Altered serial position learning after frontal lobe lesion. *Neuropsychologia*, *32*, 729-739.

Fernandez-Duque, D., Baird, J. A., & Posner, M. I. (2000). Executive attention and metacognitive regulation. *Consciousness and Cognition*, *9*, 288-307.

Fletcher, P. C., Shallice, T., & Dolan, R. J. (1998). The functional roles of prefrontal cortex in episodic memory. I. Encoding. *Brain*, *121*, 1239-1248.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). 'Min-Mental State': A practical method for grading the cognitive states of patients for the clinician. *Journal of Psychiatric Research*, *12*, 189-198.

Gershberg, F. B., & Shimmamura, A. P. (1995). The role of the frontal lobes in the use of organizational strategies in free recall. *Neuropsychologia*, *13*, 1305-1333.

Glisky, E. L., Polster, M. R., & Routhieaux, B. C. (1995). Double dissociation between item and source memory. *Neuropsychology*, *9*, 229-235.

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

Goldman-Rakic, P. S. (1987). Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In F. Plum (Ed.), *Handbook of physiology*: *The nervous system* (Vol. 5, pp. 373-417). Bethesda, MD: American Physiological Society.

Grant, D. A., & Berg, E. A. (1948). A behavioral analysis of reinforcement and ease of shifting to new responses in a Weigel-type card-sorting problem. *Journal of Experimental Psychology*, *38*, 404-411.

Hartman, M., Bolton, E., & Fehnel, S. E. (2001). Accounting for age differences on the Wisconsin Card Sorting Test: Decreased working memory, not inflexibility. *Psychology and Aging*, *16*, 385-399.

Head, D., Raz, N., Gunning-Dixon, F., Williamson, A., & Acker, J. D. (2002). Age-related differences in the course of cognitive skill acquisition: The role of regional cortical shrinkage and cognitive resources. *Psychology and Aging*, *17*, 72-84.

Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G., & Curtiss, G. (1993). *Wisconsin Card Sorting Test manual: Revised and expanded*. Odessa, FL: Psychological Assessment Resources.

Henkel, L. A., Johnson, M. K., & De Leonardis, D. M. (1998). Aging and source monitoring: Cognitive processes and neuropsychological correlates. *Journal of Experimental Psychology: General*, *127*, 251-268.

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

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

Hirst, W. (1982). The amnesic syndrome: Descriptions and explanations. *Psychological Bulletin*, *91*, 435-460.

Houx, P. J., Jolles, J., & Vreeling, F. W. (1993). Stroop interference: Aging effects assessed with the Stroop Color-Word Test. *Experimental Aging Research, 19*, 209-224.

Howard, D. V. (1980). Category norms: A comparison of the Battig and Montague (1960) with the response of adults between the ages of 20 and 80. *Journal of Gerontology, 35*, 225-231.

Jacoby, L. L. (1999a). Deceiving the elderly: Effects of accessibility bias in cued-recall performance. *Cognitive Neuropsychology*, *16*, 417-436.

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

Janowsky, J. S., Shimamura, A. P., & Squire, L. R. (1989a). Source memory impairment in patients with frontal lobe lesions. *Neuropsychologia, 27*, 1043-1056.

Janowsky, J. S., Shimamura, A. P., & Squire, L. R. (1989b). Memory and metamemory: Comparisons between patients with frontal lobe lesions and amnesic patients. *Psychobiology*, *17*, 3-11.

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.

Johnson, M. K., De Leonardis, D. M., Hashtroudi, S., & Ferguson, S. A. (1995). Aging and single versus multiple cues in source monitoring. *Psychology and Aging*, *10*, 507-517.

Johnson, M. K., Foley, M. A., Suengas, A. G., & Raye, C. L. (1988). Phenomenal characteristics of memories for perceived and imagined autobiographical events. *Journal of Experimental Psychology: General*, *117*, 371-376.

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

Johnson, M. K., Kounios, J., & Nolde, S. F. (1996). Electrophysiological brain activity and memory source monitoring. *NeuroReport*, *7*, 2929-2932.

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review*, *9*, 637-671.

Kato, T. (1985). Semantic-memory sources of episodic retrieval failure. *Memory & Cognition*, *13*, 442-452.

Kelley, C. M., & Sahakyan, L. (2003). Memory, monitoring, and control in the attainment of memory accuracy. *Journal of Memory and Language*, *48*, 704-721.

Kensinger, E. A., & Schacter, D. L. (1999). When true recognition suppresses false recognition: Effects of aging. *Cognitive Neuropsychology*, *16*, 399-415.

Kimberg, D. Y., & Farah, M. J. (1993). A unified account of cognitive impairments following frontal lobe damage: The role of working memory in complex, organized behavior. *Journal of Experimental Psychology: General*, *4*, 411-428.

Koechlin, E., Ody, C., & Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. *Science*, *302*, 1181-1185.

Koriat, A., Ben-Zur, H., & Sheffer, D. (1988). Telling the same story twice: Output monitoring and age. *Journal of Memory and Language*, *27*, 23-39.

Koriat, A., & Goldsmith, M. (1994). Memory in naturalistic and laboratory contexts: Distinguishing the accuracy-oriented and quantity oriented approaches to memory assessment. *Journal of Experimental Psychology*, *General*, *123*, 297-316.

Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review, 103*, 490-517.

Koriat, A., Goldsmith, M., & Pansky, A. (2001). Toward a psychology of memory accuracy. *Annual Review of Psychology*, *51*, 481-537.

Koustaal, W., & Schacter, D. L. (1997). Gist-based false recognition of pictures in older and younger adults. *Journal of Memory and Language*, *37*, 555-583.

Lhermitte, F. (1983). "Utilization behavior" and its relation to lesions of the frontal lobes. *Brain, 106*, 237-255.

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

Maxwell, S. E., & Delaney, H. D. (1993). Bivariate median splits and spurious statistical significance. *Psychological Bulletin*, *113*, 181-190.

McDaniel, M. A., Glisky, E. L., Rubin, S. R., Guynn, M. J., & Routhieaux, B. C. (1999). Prospective memory: A neuropsychological study. *Neuropsychology*, *13*, 103-110.

McIntyre, J. S., & Craik, F. I. M. (1987). Age differences in memory for item and source information. *Canadian Journal of Psychology*, *41*, 175-192.

Meade, M. L., & Roediger, H. L. III. (2002, November). *Age differences in illusory memories produced by forced recall*. Paper presented at the 43rd Annual Meeting of the Psychonomics Society, Kansas City, MI.

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

Moscovitch, M., & Melo, B. (1997). Strategic retrieval and the frontal lobes: Evidence from confabulation and amnesia. *Neuropsychologia*, *35*, 1017-1034.

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

Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin*, *95*, 109-113.

Norman, K. A., & Schacter, D. L. (1996). Implicit memory, explicit memory, and false recollection: A cognitive neuroscience perspective. In L. Reder (Ed.), *Implicit memory and metacognition* (pp. 229-257). Erlbaum: Mahwah, New Jersey.

Norman, K. A. & Schacter, D. L. (1997). False recognition in younger and older adults: Exploring the characteristics of illusory memories. *Memory & Cognition, 25*, 838-848.

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwarts, & D. Shapiro (Eds.), *Consciousness and self-regulation. Advances in research and theory* (Vol. 4, pp. 1-18). New York: Plenum.

Oskamp, S. (1962). The relationship of clinical experience and training methods to several criteria of clinical prediction. *Psychological Monographs*, *76*, 1-27.

Parkin, A. J., & Java, R. I. (1999). Deterioration of frontal lobe function in normal aging: Influences of fluid intelligence versus perceptual speed. *Neuropsychology*, *13*, 539-545.

Parkin, A. J., Walter, B. M. (1992). Recollective experience, normal aging, and frontal dysfunction. *Psychology and Aging*, *7*, 290-298.

Penfield, W., & Evans, J. (1935). The frontal lobe in man: A clinical study of maximum removals. *Brain, 58*, 115-133.

Perfect, T. J., & Dasgupta, Z. R. R. (1997). What underlies the deficit in reported recollective experience in old age? *Memory & Cognition*, *25*, 849-858.

Raz, N. (2000). Aging of the brain and its impact on cognitive performance: Integration of structural and functional findings. In F. I. M. Craik & T. Salthouse (Eds.), *The Handbook of Aging and Cognition*. Lawerence Erlbaum: Mahwah, NJ.

Raz, N., Gunning-Dixon, F. M., Head, D., Dupuis, J. H., & Acker, J. D. (1998). Neuroanatomical correlates of cognitive aging: Evidence from structural magnetic resonance imaging. *Neuropsychology*, *12*, 95-114.

Reder, L. M., Wilbe, C., & Martin, J. (1986). Differential memory changes with age: Exact retrieval versus plausible inference. *Journal of Experimental Psychology: Learning, Memory*, *and Cognition*, *12*, 72-81.

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

Rhodes, M. G. (in press). Age-related differences in performance on the Wisconsin Card Sorting Test: A meta-analytic review. *Psychology and Aging*.

Rhodes, M. G. & Anastasi, J. S. (2000). The effects of a levels of processing manipulation on false recall. *Psychonomic Bulletin & Review, 7*, 158-162.

Roediger, H. L. (1996). Memory illusions. *Journal of Memory and Language*, *35*, 76-100.

Roediger, H. L. & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 21*, 803-814.

Roediger, H. L., Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin & Review*, *8*, 385-407.

Rugg, M. D., & Henson, N. A. (2001). Episodic Memory Retrieval and (event related) functional neuroimaging perspective. In A. E. Parker, E. L. Wilding, & E. L. Bussey (Eds.), *The cognitive neuroscience of memory encoding and retrieval*. Hove, UK: Psychology Press.

Salthouse, T. A. (1994). Aging associations: Influence of speed on adult age differences in associative learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 1486-1503.

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

Salthouse, T. A., Atkinson, T. M., & Berish, D. E. (2003). Executive functioning as a potential mediator of age-related cognitive decline in normal adults. *Journal of Experiment Psychology: General*, *132*, 566-594.

Salthouse, T. A., & Babcock, R. L. (1991). Decomposing adult age differences in working memory. *Developmental Psychology*, *27*, 763-776.

Salthouse, T. A., Fristoe, N. M., Lineweaver, T. T., & Coon, V. E. (1995). Aging of attention: Does the ability to divide decline? *Memory & Cognition*, *23*, 59-71.

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, and Psychiatry*, *20*, 11-21.

Schacter, D. L., Harbluk, J. L., & McLachlan, D. R. (1984). Retrieval without recollection: An experimental analysis of source amnesia. *Journal of Verbal Learning and Verbal Behavior, 23*, 593-611.

Schneider, W. (1990). *MEL user's guide: Computer techniques for real time experimentation*. Pittsburgh: Psychology Software Tools.

Schretlen, D., Pearlson, G. D., Anthony, J. C., Aylward, E. H., Augustine, A. M., Davis, A., et al. (2000). Elucidating the contributions of processing speed, executive ability, and frontal lobe volume to normal age-related differences in fluid intelligence. *Journal of the International Neuropsychological Society*, *6*, 52-61.

Shallice, T., & Burgess, P. W. (1991). Deficits in strategy application following frontal lobe damage in man. *Brain, 114*, 727-741.

Shimamura, A. P. (1996). The role of the prefrontal cortex in controlling and monitoring memory processes. In L. M. Reder (Ed.), *Implicit memory and metagcognition* (pp. 259-274). Erlbaum: Mahwah, NJ.

Shimamura, A. P. (2000a). The role of the prefrontal cortex in dynamic filtering. *Psychobiology*, *28*, 207-218.

Shimamura, A. P. (2000b). Toward a cognitive neuroscience of metacognition. *Consciousness and Cognition*, *9*, 313-323.

Shimamura, A. P., Janowsky, J. S., & Squire, L. R. (1990). Memory for the temporal order of events in patients with frontal lobe lesions. *Neuropsychologia*, *28*, 801-813.

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, *7*, 422-445.

Souchay, C., Isingrini, M., & Espagnet, L. (2000). Aging, episodic memory feeling-of-Knowing, and frontal functioning. *Neuropsychology*, *14*, 299-309.

Souchay, C., Isingrini, M., Clarys, D., Taconnat, L., & Eustache, F. (2004). Executive functioning and judgment-of-learning versus feeling-of-knowing in older adults. *Experimental Aging Research*, *30*, 47-62.

Spencer, W. D., & Raz, N. (1995). Memory for facts, source, and context: Can frontal lobe dysfunction explain age-related differences? *Psychology and Aging*, *9*, 149-159.

Spreen, O., & Strauss, E. (1998). *A compendium of neuropsychological tests: Administration, norms, and commentary*. New York: Oxford University Press.

Troyer, A. K., Graves, R. E., & Cullum, C. M. (1994). Executive functioning as a mediator of the relationship between age and episodic memory in healthy aging. *Aging and Cognition*, *1*, 45-53.

Tun, P. A., Wingfield, A., Rosen, M. J., & Blanchard, L. (1998). Response latencies for false memories: Gist-based processes in normal aging. *Psychology and Aging*, *13*, 230-241.

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

Uttl, B. (2002). North American Adult Reading Test: Age norms, reliability, and validity. *Journal of Clinical and Experimental Neuropsychology*, *24*, 1123-1137.

Vilkki, J., Servo, A., & Surma-aho, O. (1998). Word list learning and prediction of recall after frontal lobe lesions. *Neuropsychology*, *12*, 268-277.

Waltz, J. A., Knowlton, B. J., Holyoak, K. J., Boone, K. B., Mishkin, F. S., de Menezes Santos, M., et al. (1999). A system for relational reasoning in human prefrontal cortex. *Psychological Science*, *10*, 119-125.

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

Whelihan, W. M., Lesher, E. L. (1985). Neuropsychological changes in frontal function with aging. *Developmental Neuropsychology*, *1*, 371-380.

Yaniv, I., Yates, J. F., & Smith, J. E. K. (1991). Measures of discrimination skill in probabilistic judgment. *Psychological Bulletin*, *110*, 611-17.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*, 441-517.

Zacks, R. T., Hasher, L., & Li, K. Z. H. (2000). Human memory. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 293-357). Erlbaum, Mahwah, NJ.

# BIOGRAPHICAL SKETCH

Matthew Gerard Rhodes received his Bachelor of Science degree in psychology from Francis Marion University in Florence, South Carolina in 1999 and his Masters of Science degree in psychology from Florida State University in 2002. He has authored or co-authored several peer-reviewed publications and has taught courses in cognition and research methods. His current research interest is in memory.