Neuropsychological Status in Older Adults Influences Susceptibility to False Memories

MICHELLE L. MEADE
Montana State University

LISA D. GERACI
Texas A&M University

HENRY L. ROEDIGER, III
Washington University
in Saint Louis

In 2 experiments we examined the influence of frontal lobe function on older adults' susceptibility to false memory in a categorized list paradigm. Using a neuropsychological battery of tests developed by Glisky, Polster, and Routhieaux (1995), we designated older adults as having high- or low-frontal function. Young and older adults studied and were tested on categorized lists using free report cued recall and forced report cued recall instructions, with the latter requiring participants to produce responses even if they had to guess. Under free report cued recall instructions, frontal lobe function was a strong predictor of false memories in older adults: Older adults who scored low on tests of frontal functioning demonstrated much higher levels of false recall than younger adults, whereas levels of false recall in high-frontal older adults were more similar to those of young adults. However, after forced report cued recall, high- and low-frontal older adults performed similarly to each other, and both demonstrated higher levels of false recall than young adults. On a final recognition test, high-frontal older adults in both the free report cued recall and forced report cued recall conditions were more successful than low-frontal older adults in using source information to reduce memory errors. The results indicate that older adults show higher levels of false recall than younger adults, but type of test (free report or forced report) and neuropsychological status of older adults mediate these effects. Low-frontal older adults are particularly susceptible to producing false memories on free report tests that entail source monitoring.

Older adults generally perform more poorly on tests of recall than do younger adults (see Balota, Dolan, & Duchek, 2000, for a review), and older adults are also more susceptible to false memories than are younger adults (see Roediger & McDaniel, 2007). However, recent studies suggest that not all older adults are equally susceptible to false memories in standard paradigms. Rather, susceptibility is correlated with older adults' abilities as assessed on tests intended to measure frontal lobe functioning (Butler, McDaniel, Dornburg, Price, & Roediger, 2004; Chan & McDermott, 2007; Roediger & Geraci, 2007). Older

American Journal of Psychology
Winter 2012, Vol. 125, No. 4 pp. 449–467 •

adults with poor frontal functioning are generally more susceptible to false memories across a variety of paradigms, and older adults who score high on tests of frontal functioning often show levels of veridical and false responding equal to that of young adults in paradigms such as the Deese–Roediger–McDermott (DRM) false memory paradigm (Deese, 1959; Roediger & McDermott, 1995) or the Loftus misinformation paradigm (Loftus, 1992), among others (see Butler et al., 2004, and Roediger & Geraci, 2007).

The goal of the current experiments was to explore further the effect of frontal functioning on older adults' susceptibility to false memories by using a categorized list paradigm that includes both free report cued recall and forced report cued recall. Free report cued recall tests are given under an instruction that allows participants to determine their own criterion for reporting items on a memory test ("Recall words from the list without guessing wildly"), whereas forced report cued recall tests require participants to write down a certain number of items, guessing if necessary. Participants are then asked on subsequent recall and recognition tests to remember items presented during study (which may or may not include the items just produced under free report cued recall or forced report cued recall). We chose these tasks because varying demands of free report cued recall and forced report cued recall will affect response criterion and source monitoring, which (according to prior work) should illuminate the role of frontal lobe function on older adults' susceptibility to false memories. In particular, forced report cued recall should harm low-frontal older adults relative to high-frontal older adults and young adults. We clarify these expectations in the next few paragraphs by reviewing pertinent research.

Healthy older adults can be characterized as varying along a continuum from high frontal to low frontal according to a battery of neuropsychological tests developed by Glisky and colleagues (Glisky, Polster, & Routhieaux, 1995; Glisky, Rubin, & Davidson, 2001). The battery includes five tests that were selected originally from a much larger number of putative frontal tests because they converged on the construct of frontal function in the early studies. It includes various subscales of the modified Wisconsin Card Sort Test (Hart, Kwentus, Wade, & Taylor, 1988), the Controlled Oral Word Association Task (Benton & Hamsher, 1976), the Wechsler Adult Intelligence Scale–Revised (Wechsler, 1981), the Wechsler Memory Scale–Revised (Wechsler, 1987), and the California Verbal Learning Test (Delis, Kramer, Kaplan, & Ober, 1987). A composite score based on the combined tests has been used in prior research to designate participants as having high- or low-frontal function (Glisky et al., 1995).

Much evidence has linked older adults' frontal status as determined by this battery of neuropsychological tests to source memory abilities. For example, Glisky et al. (1995) designated older adults as having high or low frontal functioning and then presented them with sentences spoken in a male or female voice. On a subsequent test, participants were asked to recognize the content of the sentence (item memory) and the voice in which the sentence had been spoken (source memory). Participants who scored high on the frontal battery demonstrated better source memory than those who scored low on the frontal battery, although the two groups did not differ in item memory (i.e., recognition of whether sentences were old or new). Several other subsequent studies demonstrated this relationship between frontal status and source memory (Glisky et al., 2001; Glisky & Kong, 2008; Henkel, Johnson, & De Leonardis, 1998; Mather, Johnson, & De Leonardis, 1999). According to Johnson, Hashtroudi, and Lindsay's (1993) theory of source monitoring, participants make source judgments by retrieving any additional characteristics associated with the event in question in order to make a decision about the event's source. Such complex evaluative and decision processes are consistent with the effortful, strategic processing for which frontal regions of the brain are assumed to be responsible (Raz, 2000).

Because low frontal functioning as assessed by neuropsychological tests is associated with reduced source monitoring abilities, and reduced source monitoring abilities in older adults may inflate false memories (Balota et al., 1999), a straightforward prediction is that older adults with low frontal functioning should be especially susceptible to false memories across a variety of paradigms. Several studies have found just that. As noted previously, Butler et al. (2004) demonstrated that performance on Glisky et al.'s (1995) frontal battery predicted older adults' susceptibility to false recall in the DRM paradigm

(Deese, 1959; Roediger & McDermott, 1995). Briefly, participants were presented with lists of semantically related words that each converged on a single nonpresented critical item. Low-frontal older adults demonstrated greater false recall of the critical items than young adults, but high-frontal older adults did not differ from young adults. Frontal functioning has also been shown to predict older adults' susceptibility to false alarms in a standard misinformation paradigm (cf. Loftus, Miller, & Burns, 1978). Roediger and Geraci (2007) found that low-frontal participants were especially susceptible to misinformation, and importantly, low-frontal older adults were not able to reduce errors on a final recognition test that called attention to the source of the misinformation, whereas high-frontal older adults could (cf. Multhaup, 1995).

Finally, Chan and McDermott (2007) demonstrated that frontal functioning is related to false memory performance in a pragmatic inference paradigm (Brewer, 1977; McDermott & Chan, 2006). After studying sentences that implied a given action (e.g., "The hungry python caught the mouse"), low-frontal participants were more likely than high-frontal participants to misremember sentences containing the inferred action (e.g., "The hungry python ate the mouse"). Furthermore, by using a slightly different method to compute the composite score on the frontal battery, Chan and McDermott (2007) were able to classify younger adults and older adults into high- and low-frontal groups. They demonstrated that age and frontal functioning were independent predictors of false recall. Considered together, research on frontal functioning and false memory processes provides converging evidence that measures of frontal functioning are associated with false responding and that older adults' greater susceptibility to false memory effects may be mediated by frontal functioning. By this hypothesis, low-frontal older adults are less able to engage in effortful, recollective processes that would allow one to determine the correct source of the interfering information and therefore to withhold erroneous responses.

The current experiments extend previous research on source monitoring difficulties among high- and low-frontal older adults by examining false recall and recognition in a categorized list paradigm that compares initial free report cued recall (with a warning against guessing) with forced report cued recall (requiring guessing) when these measures are followed by subsequent recall and source monitoring recognition tests (Meade & Roediger, 2006, 2009). In this paradigm, participants are presented with categorized word lists that exclude the most common exemplars of a given category. In the free report cued recall condition, participants are asked to recall exemplars from each list without guessing, and false recall is determined by the proportion of nonpresented, typical exemplars that participants produce. In this case, the source monitoring demands are quite intensive, as the most typical exemplars are familiar yet must be excluded. Consistent with previous work demonstrating that older adults are especially likely to mistake highly accessible items for previously studied items (e.g., Jacoby, 1999), older adults in the categorized list paradigm demonstrated higher rates of false recall than young adults under free report cued recall instructions (Meade & Roediger, 2006, 2009). Meade and Roediger (2006, Experiment 2) showed that this age difference was quite pronounced when young adults were compared with just the older adults who scored low on the frontal battery. Although high-frontal older adults have not yet been tested in the categorized list paradigm, the free report cued recall condition is conceptually similar to previous research conducted on frontal differences in older adults' susceptibility to false memory using the DRM paradigm (e.g., Butler et al., 2004). In the current research, we compared high- and low-frontal older adults' performance with that of young adults in the categorized list paradigm. We predicted that, consistent with previous findings, low-frontal older adults would show higher levels of false recall than both high-frontal older adults and young adult participants in the free report cued recall condition, and high-frontal older adults would not differ from young adults.

The primary novel component of the current experiments is the examination of frontal functioning as a predictor of false recall under a different instructional set than used in previous work, viz., forced report cued recall instructions. In the forced report cued recall condition of the current experiments, participants were asked to recall the categorized lists with the requirement that they must produce a given number of words, even if they must guess. Because forced report cued recall requires participants to generate their own misleading information—their guesses—performance on subsequent source monitoring tasks is especially difficult, because participants must differentiate the items that appeared on the study list from the items generated as guesses to complete the forced report cued recall test (Meade & Roediger, 2006, 2009; cf. Johnson, Foley, & Leach, 1988). Consistent with age-related declines in source monitoring (Hashtroudi, Johnson, & Chrosniak, 1989; McIntyre & Craik, 1987), in the single prior study using forced report cued recall, low-frontal older adults showed larger false memory effects after forced report cued recall in the categorized list paradigm than young adults (Meade & Roediger, 2006, Experiment 2). The effect of forced report cued recall on high-frontal older adults has not yet been examined, although the task demands of forced report cued recall are directly relevant to the source monitoring difficulties associated with frontal function. The current study directly compared high- and low-frontal older adults to examine the influence of frontal functioning on susceptibility to false memories generated using the categorized list paradigm. Based on previous research demonstrating that greater frontal differences are more likely for difficult source monitoring decisions (Henkel et al., 1998), we predicted that differences in levels of false responding between high- and low-frontal older adults would be greater after forced report cued recall than after free report cued recall.

The inclusion of forced report cued recall in the current study is also relevant to another possible explanation for frontal differences: strategic processing. Pansky, Goldsmith, Koriat, and Pearlman-Avnion (2009) demonstrated that lower frontal functioning (determined by the Wisconsin Card Sorting Test and the Controlled Oral Word Association Test only) was associated with less reliance on subjective monitoring. They defined subjective monitoring as the correspondence between confidence in a given item and the likelihood of reporting that item on a memory test. Pansky et al.'s finding that lower frontal functioning was associated with reduced monitoring provides one example of how high- and low-frontal older adults rely on different strategies to respond on a memory test. Another example of strategic differences is that high-frontal older adults are able to use source discrimination to reduce memory errors on a misinformation task. Specifically, high-frontal older adults can improve their memory accuracy when a source monitoring test encourages them to focus on potential sources of memory (Roediger & Geraci, 2007). More generally, Mather et al. (1999) suggested that older adults with high frontal lobe function might be better able to adapt their strategies to task demands. In the current study, we were interested primarily in frontal differences in strategic regulation of memory output, which would include the ability to monitor the accuracy of one's memory performance using multiple sources of information. Forced report cued recall is relevant here because under forced report cued recall instructions, participants no longer determine their own criterion for reporting an item but instead must comply with experimenter instructions to produce a given number of responses. In contrast, on a free report cued recall test, participants can determine how many items they choose to report (cf. Koriat & Goldsmith, 1996). To the extent that high-frontal participants strategically regulate their memory responses differently from low-frontal participants, differences in responding between the groups should occur only on the free report cued recall test where different strategies are allowed. No differences would be predicted on the initial forced report cued recall test.

Furthermore, in both test conditions there are subsequent recall and source monitoring recognition tests, so we can determine the influence of initial free report cued recall and forced report cued recall on participants' ability to regulate memory accuracy over repeated tests (see Henkel, 2007, for a discussion of age differences across repeated testing). Examining performance on subsequent tests is important because the source monitoring requirements on tests following cued and forced report cued recall differ. In both cases, participants must monitor whether items produced on the initial test appeared on the study list. After forced report cued recall, participants must further differentiate items produced on the initial test that appeared on the study list from those generated as plausible guesses to complete the forced report cued recall test. The increased source monitoring demands after forced report cued recall should exaggerate frontal differences (cf. Henkel et al., 1998); we predicted that any difference in false responding between high- and low-frontal older adults would be greater after forced report cued recall than after free report cued recall.

EXPERIMENT 1

Experiment 1 was designed to examine whether tests intended to measure frontal lobe function mediate the effects of aging in a categorized list recall paradigm. Previous research using this paradigm has established that older adults show higher levels of false recall and recognition than do young adults in free report cued recall and that the effect is magnified after forced report cued recall, at least for older adults who score low on the frontal battery (Meade & Roediger, 2006, Experiment 2); high-frontal older adults have not yet been tested using this paradigm. The current experiment examined the effects of age and frontal status on both free report cued recall and forced report cued recall of categorized lists to gain further evidence about whether the putative effects of age on false remembering are mediated by frontal lobe functioning. Based on prior research, we predicted that older adults would show greater levels of false recall than younger adults, low-frontal older adults would show the effect more than high-frontal older adults, and forced report cued recall would magnify the number of errors because it places greater burdens on source monitoring. Finally, we predicted that forced report cued recall would be especially deleterious for older adults with low frontal functioning.

METHOD

Participants

Participants were 30 Washington University undergraduates (age range 18–25 years, mean 20) and 60 older adults recruited from the Washington University Older Adult Participant Pool (age range 66–85 years, mean 74). Young adults received class credit or $10 cash for their participation, and older adults received $15 cash (the rate of pay was $10 per hour; older adults were paid more because they generally took longer to complete the experiment). Older adults were administered a neuropsychological battery containing the Modified Wisconsin Card Sort, controlled oral word association test, mental arithmetic, mental control from the Wechsler Adult Intelligence Scale, and backward digit span, all of which are thought to reflect frontal lobe functioning (Glisky et al., 1995, 2001). Performance on each test was calculated as an age-adjusted *z* score, and a composite measure of frontal function for each participant was created by averaging the *z* scores across all of the tests. Consistent with the guidelines established by Glisky et al. (1995, 2001), 30 older adults with composite scores above 0 were classified as high frontal and 30 older adults with composite scores below 0 were classified as low frontal.

Because of population limitations, the same high- and low-frontal older adults who participated in Experiment 1 also participated in Experiment 2 along with 4 additional participants who participated in Experiment 2 only (2 additional high-frontal older adults and 2 additional low-frontal older adults were added for counterbalancing purposes). Data from all 64 older adult participants revealed that high-frontal participants (M = 73.94 years old) and low-frontal participants (M = 74.47 years old) did not differ in age, $t < 1.0$, but they did differ in both years of education (M = 15.92 years of education for high-frontal participants, M = 14.39 years of education for low-frontal participants), $t(62) = 2.57$, $SEM = .59$, and Shipley vocabulary scores (M = 36.75 for high-frontal participants, M = 34.44 for low-frontal participants), $t(62) = 3.25$, $SEM = .71$. Because of population limitations, the low-frontal older adults in the current studies partially overlapped with the low-frontal older adults whose data were reported in Meade and Roediger (2006; Experiment 2). However, participants' data in the earlier article were combined with additional low-frontal participants and were used for different purposes. In the current analyses, low-frontal older adults' data are compared with high-frontal older adults' data.

Design

The experiment consisted of a 3 × 2 between-subject design. Participant group (young adults, high-frontal older adults, or low-frontal older adults) and retrieval condition on the first recall test (free report cued recall or forced report cued recall) were both between-subject factors. The primary dependent variables were correct and false recall on two recall tests and correct and false recognition on a final source monitoring recognition test.

Materials

The same materials used by Meade and Roediger (2006) were used in the current experiment. Specifically, categorized word lists were constructed from the top 22 exemplars from six categories in the Battig and Montague (1969) norms. The top five exemplars (numbered 1–5) were designated as the critical items and not included in the study list. Exemplars numbered 6–22 were designated as studied items, and these 17 items appeared in the study list. The recognition source monitoring test contained the 5 critical items from each of the six study lists, 5 randomly selected studied items from each study list, and 30 unrelated filler items.

Procedure

The procedure was identical to that used in the no-warning condition of Meade and Roediger's (2006) Experiment 2. Participants were tested in small groups ranging from two to five people of the same age group (but not necessarily the same frontal designation for older adults). Within each group, all participants were tested in the same experimental condition (free report cued recall or forced report cued recall). Participants were told to study each of the six categorized lists in preparation for a memory test. Each list was verbally labeled as it was presented (e.g., "the bird list"), and each word was visually presented for 1.5 s. After presentation of all six lists, participants completed a visual spatial filler task for 4 min in order to eliminate short-term memory effects.

Participants were then presented with individual sheets of paper headed with the category name and asked to recall items from the list under free report cued recall or forced report cued recall instructions. Participants in the free report cued recall condition were asked to recall as many items as possible from the list without guessing. Participants in the forced report cued recall condition were asked to produce 20 items even if they had to generate guesses to fill in all 20 spaces (the actual lists contained only 17 items, so they would have to guess). All participants successfully completed this task. In an effort to roughly equate the amount of time allotted for each item produced, participants in the free report cued recall condition were given 2 min to recall items from each category, and participants in the forced report cued recall condition were given 4 min to produce 20 words from each category. No participants requested additional time. In addition, participants in both the free report cued recall and forced report cued recall conditions were asked to concurrently rate their confidence in each item produced on a scale of 1 to 4 (1 = *sure the item was not on the list,* 2 = *pretty sure the item was not on the list,* 3 = *pretty sure the item was on the list,* 4 = *sure the item was on the list*).

Participants were next asked to complete another recall test for each of the six lists. The second recall test was always administered under free report cued recall instructions, so participants were asked to recall as many items as possible without guessing. The category names were again presented in the same order on individual sheets, and participants had 2 min to write down as many items as possible from each category. Participants were asked to make confidence ratings for each item produced by relying on the same scale used in Recall Test 1.

Finally, participants were given a final source monitoring recognition test and asked to indicate the source of each item presented on the test. Specifically, they were asked to indicate whether they remembered the item from the study lists ("lists"), they had recalled it on a previous recall test ("self"), or they did not remember the item as having occurred in the context of the current experiment ("neither"). Participants could check multiple boxes for each word (e.g., "list" and "self") if they believed they had studied the word and also recalled it previously.

TABLE 1. Mean Proportion of List and Critical Items Recalled and Mean Proportion of Highest Confidence Ratings (CR4) Produced by Young Adults, High-Frontal Older Adults, and Low-Frontal Older Adults on Free Report Cued Recall or Forced Report Cued Recall Test 1, Experiment 1

		Older adults	
	Young adults	High frontal	Low frontal
	Recall (CR4)	Recall (CR4)	Recall (CR4)
Free report cued recall			
List	.44 (.40)	.45 (.37)	.41 (.34)
Critical	.17 (.06)	.15 (.04)	.36 (.12)
Forced report cued recall			
List	.66 (.45)	.58 (.39)	.55 (.39)
Critical	.66 (.11)	.68 (.25)	.67 (.32)

RESULTS

Recall Test 1

Table 1 reveals the mean proportion of list and critical items produced on the first recall test by young adults and high-frontal and low-frontal older adults under instructions of free report cued recall and forced report cued recall. Data in parentheses are the proportion of items for which participants expressed the highest level of confidence (ratings of 4). We consider the overall number of items produced in the first set of analyses and then analyze the high confidence data.

Free report cued recall and forced report cued recall were analyzed separately because guessing is encouraged in the second but not the former test. The level for statistical significance was set at $p \leq .05$ unless otherwise noted.

LIST ITEMS

No differences appeared in free report cued recall of the lists across participant groups. A univariate ANOVA with participant group as an independent factor revealed that young adults, high-frontal older adults, and low-frontal older adults recalled similar proportions of list items on free report cued recall, $F < 1.0$, possibly because of the strong retrieval cues (category names) present at recall (cf. Craik, 1983; Craik & McDowd, 1987). Of course, this analysis does not correct for any inflation that may result from older adults being more error prone than younger adults.

In contrast to the pattern obtained on the free report cued recall test, the participant groups differed in the proportion of list items produced under forced report cued recall instructions, $F(2, 42) = 14.36$, $MSE = .01$. Specifically, young adults produced more list items than both high-frontal older adults, $t(28) = 3.55$, $SEM = .02$, and low-frontal older adults, $t(28) = 5.02$, $SEM = .02$. Low- and high-frontal older adults did not differ from each other, $t(28) = 1.65$, $SEM = .02$, $p > .05$. This outcome in relation to the free report cued recall results may seem odd, because a reasonable expectation might have been for older adults to do more poorly on free report cued recall, in which they needed to monitor their recollections and only produce list items, than on forced report cued recall, in which they could produce any response.

CRITICAL ITEMS

Older adults intruded more critical items than younger adults on the free report cued recall test, but this outcome was confined to the older adults who scored low on the frontal battery. A univariate ANOVA revealed a significant main effect of participant group, $F(2, 42) = 8.17$, $MSE = .03$. Subsequent analyses confirmed that false recall of high-frontal older adults did not differ from false recall of young adults, $t < 1.0$. However, low-frontal older adults produced more critical items than both high-frontal older adults, $t(28) = 3.33$, $SEM = .06$, and young adults, $t(28) = 3.24$, $SEM = .06$, by a factor of two. This finding is consistent with previous research (e.g., Butler et al., 2004) showing that aging effects in false recall are driven primarily by low-frontal older adults.

No group differences were obtained in the proportion of critical lures produced under forced report cued recall instructions, $F < 1.0$. Participants in this condition were required to guess, and so it is not surprising that participants in all groups were equally likely to produce the most common exemplars from each category as guesses and to the same degree.

CONFIDENCE RATINGS

For each item produced on the first recall test, participants indicated how confident they were that the item had been presented on the list. Confidence ratings ranged from 1 to 4, with 4 indicating the highest level of confidence. Table 1 presents the mean proportion of items produced that participants gave a confidence rating of 4 in parentheses (the pattern does not change when confidence ratings of both 3 and 4 are included). Analyses of confidence ratings are not reported for free report cued recall because participants were asked not to guess and so mostly assigned confidence ratings of 3 and 4 to produced items.

The three participant groups did not show reliable variability in high confidence responses to list items in forced report cued recall, $F < 1$, although older adults did perform numerically worse than younger adults on this measure. However, the three participant groups did vary in the proportion of high confidence responses to critical items, $F(2, 45) = 5.57$, $MSE = .03$. Subsequent analyses revealed that young adults were highly confident in fewer errors than were high-frontal older adults, $t(28) = 2.35$, $SEM = .06$, and low-frontal older adults, $t(28) = 3.53$, $SEM = .06$, but that the two groups of older adults did not differ, $t < 1.0$. Thus, even though the three groups did not differ in overall proportions of errors produced in the first phase of forced report cued recall, older adults judged more than twice as many critical items as having appeared in the list. This finding is consistent with work by Kelley and Sahakyan (2003), who found that older adults have a lower correspondence between confidence ratings and objective correctness of items (see Mitchell, Johnson, & Mather, 2003). Our findings suggest that on tests of forced report cued recall, both high-frontal and low-frontal older adults demonstrate reduced subjective monitoring (cf. Pansky et al., 2009).

Recall Test 2

After the first test phase, all participants were tested again under free report cued recall instructions with a warning against guessing. The primary interest in this analysis is in seeing whether participants who were forced to guess in the first test phase will show greater levels of false recall than those who were not forced to guess. Table 2 presents the mean proportion of list and critical items recalled on Recall Test 2 as a function of participant group and prior retrieval condition.

LIST ITEMS

A 3 (young adults or high-frontal older adults or low-frontal older adults) × 2 (prior free report cued recall or prior forced report cued recall) ANOVA computed on the mean proportion of list items recalled on Test 2 revealed no significant main effect of participant group, $F(2, 84) = 2.48$, $MSE = .02$, $p > .05$, or prior retrieval condition, $F < 1.0$; however, a reliable interaction between prior retrieval condition and participant group occurred, $F(2, 84) = 3.86$, $MSE = .02$. Further analyses revealed that all participants produced equivalent proportions of list items after free report cued recall, *t*s < 1.0. After prior forced report cued recall, however, young adults recalled a greater proportion of list items than did high-frontal older adults, $t(28) = 2.73$, $MSE = .05$, and low-frontal older adults, $t(28) = 3.76$, $SEM = .04$. The two older groups (high- and low-frontal) did not differ from each other, $t < 1.0$. The finding that older adults were less likely to produce list items after prior forced report cued recall might be explained by their reduced source monitoring ability or their reduced subjective monitoring ability. Because forced recall requires participants to produce both list items and guesses, it becomes an especially difficult source monitoring task to subsequently determine which items produced previously were list items and which were guesses. Older adults' source monitoring deficits or reduced subjective monitoring exaggerates this difficulty, such that their recall of both list items and critical items was affected by prior forced report cued recall.

TABLE 2. Mean Proportion of List and Critical Items Recalled and Mean Proportion of Highest Confidence Ratings (CR4) Produced by Young Adults, High-Frontal Older Adults, and Low-Frontal Older Adults on Recall Test 2, Experiment 1

		Older adults	
	Young adults	**High frontal**	**Low frontal**
	Recall (CR4)	Recall (CR4)	Recall (CR4)
Prior free report cued recall			
List	.42 (.39)	.45 (.37)	.43 (.35)
Critical	.18 (.08)	.16 (.05)	.47 (.17)
Prior forced report cued recall			
List	.51 (.44)	.38 (.32)	.37 (.31)
Critical	.24 (.09)	.37 (.23)	.40 (.27)

CRITICAL ITEMS

Primary interest centers on two aspects of critical item recall, especially with regard to group differences after free report cued recall and forced report cued recall. Indeed, older adults showed greater false recall than younger adults, although the particular pattern that occurred was not as predicted. An ANOVA conducted for the mean proportion of critical items produced on Recall Test 2 revealed a significant main effect of participant group, $F(2, 84) = 11.92$, $MSE = 2.79$. Low-frontal older adults ($M = .43$) produced more critical items than did both young adults ($M = .21$), $t(28) = 4.79$, $SEM = .05$, and high-frontal older adults ($M = .26$), $t(28) = 3.75$, $SEM = .04$. High-frontal older adults showed levels of false recall that did not differ significantly from that of young adults, $t < 1.2$, $p > .05$. Consistent with Butler et al. (2004), this finding supports the idea that the battery of frontal tests is correlated with false recall.

Surprisingly (and in contrast to other research), the effect of prior retrieval condition on false recall was only marginally significant; participants falsely recalled a mean proportion of .27 critical items after prior free report cued recall and .33 items after prior forced report cued recall), $F(1, 84) = 3.14$, $MSE = .03$, $p = .08$. This result is surprising given that prior forced report cued recall often leads to errors (Ackil & Zaragoza, 1998; Meade & Roediger, 2006; Roediger, Wheeler, & Rajaram, 1993). Unlike prior research, we gave confidence ratings during the first test, and it is possible that the process of making ratings on Recall Test 1 helped participants monitor guesses (see Henkel, 2004; Meade & Roediger, 2006).

Most interestingly, the interaction between participant group and prior retrieval condition was significant, $F(2, 84) = 4.79$, $MSE = .03$. Subsequent analyses on the interaction revealed that young adults' false recall did not differ as a function of prior retrieval condition, $t < 1.0$, nor did the false recall of low-frontal older adults differ, $t < 1.5$, with false recall in this group being 11% greater on the second free report cued recall test than the first (Table 1). Low-frontal participants actually produced more errors after free report cued recall than after forced report cued recall, contrary to our prediction. In contrast, high-frontal older adults demonstrated a significant effect of prior retrieval condition, $t(28) = 3.30$, $SEM = .07$; after forced report cued recall, they were much more likely to produce critical items on the subsequent test than after free report cued recall. Free report cued recall instructions allowed participants to use their normal memory strategies, and they usually set the criterion for responding to be relatively conservative (Koriat & Goldsmith, 1996). However, when they were instructed to guess during forced report cued recall, participants' preferred retrieval strategies may be disrupted. The requirement to guess disproportionately affects high-frontal older adults and may overwhelm their source monitoring or subjective monitoring capabilities; low-frontal older adults are apparently affected this way even by successive free report cued recall tests. The fact that task demands have a disproportionately large impact on the high-frontal older adults supports the idea that there may be strategic differences between high- and low-frontal older adults.

CONFIDENCE RATINGS

As for Recall Test 1, participants were asked to rate how confident they were in their responses on Recall Test 2. The mean proportion of 4 responses (the highest confidence rating) are presented in Table 2. Recall Test 2 was always administered under free report cued recall instructions asking participants not to guess. Consequently, the analyses of high confidence responses are concordant with conclusions from all responses.

Recognition

At the end of the experiment, participants were given a final recognition test. Keep in mind that performance on this test undoubtedly is compromised by the prior two recall tests. The mean proportion of veridical and false recognition is displayed in Table 3. Veridical recognition was defined as the proportion of times that participants attributed a list item to having been presented in the list ("lists" alone or in conjunction with "self" responses). False recognition was defined as the proportion of times participants attributed a critical item as having been presented in the list ("lists" alone or in conjunction with "self" responses). We report all the data although we were primarily interested in veridical and false recognition.

Examining first veridical recognition, a 3 (young adults, high-frontal older adults, or low-frontal older adults) × 2 (prior free report cued recall or prior forced report cued recall) ANOVA revealed no significant main effect of participant group, $F(2, 84) = 2.79$, $MSE = .03$, $p > .05$, no main effect of prior retrieval condition, $F < 1.0$, $p > .05$, and no interaction between participant group and prior retrieval condition, $F < 1.0$, $p > .05$. Regardless of prior free report cued recall or forced report cued recall, young and older adults (both high- and low-frontal) correctly recognized a similar proportion of list items. Finding no age difference on the recognition test is not atypical (cf. Balota et al., 2000).

For false recognition, however, large differences were observed between groups. After forced report cued recall, young adults showed the lowest false recognition, whereas both groups of older adults showed higher levels of false recognition. After free report cued recall, high-frontal older adults showed the lowest false recognition. An ANOVA revealed a significant main effect of participant group, $F(2, 84) = 4.99$, $MSE = .07$. Low-frontal older adults ($M = .74$ collapsing across retrieval condition) falsely recognized the highest proportion of critical items relative to high-frontal older adults ($M = .54$ collapsing across retrieval condition), $t(28) = 2.75$, $SEM = .07$, and young adults ($M = .60$ collapsing across retrieval condition), $t(28) = 2.36$, $SEM = .06$. High-frontal older adults and young adults did not differ from each other, $t < 1.0$. Furthermore, there was a main effect of retrieval condition, $F(1, 84) = 4.15$, $MSE = .07$, demonstrating that participants were more likely to falsely recognize critical items after prior forced report cued recall ($M = .68$ collapsing across participant group) than after prior free report cued recall ($M = .57$ collapsing across participant group). Finally, the interaction between participant group and retrieval condition was signifi-

TABLE 3. Mean Proportion of Veridical and False Recognition for Correct and Critical Items Produced by Young Adults, High-Frontal Older Adults, and Low-Frontal Older Adults, Experiment 1

	Prior free report cued recall			Prior forced report cued recall		
		Older			Older	
Source judgments	Young	High	Low	[illegible]	High	Low
			List items			
List only	.39	.21	.30	.21	.19	.26
Both list and self	.49	.54	.53	.63	.58	.54
Total veridical recognition	.88	.76	.84	.84	.77	.81
Self only	.00	.00	.02	.07	.08	.03
Neither list nor self	.12	.24	.15	.08	.15	.17
			Critical items			
List only	.41	.19	.22	.18	.15	.21
Both list and self	.23	.22	.44	.37	.51	.61
Total false recognition	.64	.41	.66	.55	.67	.82
Self only	.01	.00	.05	.30	.18	.06
Neither list nor self	.36	.58	.29	.15	.16	.13

cant, $F(2, 84) = 3.59$, $MSE = .07$. Further analyses on the interaction revealed that after prior free report cued recall instructions, high-frontal older adults actually demonstrated lower levels of false memory than young adults, $t(28) = 2.35$, $SEM = .10$, and low-frontal older adults, $t(28) = 2.31$, $SEM = .10$, who did not differ from each other, $t < 1.0$. After prior forced report cued recall, high-frontal older adults did not perform reliably worse than young adults, $t < 1.5$, but low-frontal older adults falsely recognized more items than young adults, $t(28) = 3.14$, $SEM = .09$, and high-frontal older adults, although this last difference was only marginally significant, $t(28) = 1.97$, $SEM = .08$, $p = .057$.

DISCUSSION

The results of Experiment 1 confirm that older adults who score low on tests of frontal functioning show a greater tendency to false recall and false recognition than either older adults who score high on frontal tests or young adults, confirming prior research in other paradigms (Butler et al., 2004; Roediger & Geraci, 2007). Furthermore, both groups of older adults showed higher levels of false recognition when their initial test was forced report cued recall rather than free report cued recall. Low-frontal older adults recalled and recognized many items in error on the later tests after either type of initial test, whereas high-functioning older adults showed high levels of false responding only when the first test used the forced report cued recall procedure. Our interpretation of these data is that low-frontal older adults are challenged by source monitoring difficulties even by a free report cued recall test, whereas high-frontal older adults' performance deteriorates only when they are faced with the challenge of a forced report cued recall test (which increases source monitoring problems).

One surprise in the results of Experiment 1 is that forced report cued recall did not have dramatic effects on later responding, as are sometimes seen in prior work (Ackil & Zaragoza, 1998; Roediger et al., 1993), although the data trended in the expected direction. We attribute the differing effects to the fact that we used confidence ratings in our first test (free report cued recall or forced report cued recall) and prior researchers did not. The use of confidence ratings sensitizes participants to source monitoring difficulties during recall and may provide sufficient information to avoid errors on future tests. We examine this idea in the second experiment by removing the requirement for confidence ratings.

EXPERIMENT 2

Experiment 1 revealed that low-frontal older adults demonstrated greater false recall and recognition than young adults after both prior free report cued recall and prior forced report cued recall. However, high-frontal older adults demonstrated greater false recall than young adults only after forced report cued recall. After free report cued recall, erroneous recall of high-frontal older adults was equivalent to that of young adults. These findings extend previous work on older adults using the categorized list paradigm (Meade & Roediger, 2006, 2009) and also confirm research demonstrating that frontal status mediates false recall in older adults (Butler et al., 2004; Roediger & Geraci, 2007; Chan & McDermott, 2007), at least on test of free report cued recall. Frontal status does not necessarily mediate older adults' susceptibility to false recall on tests of forced report cued recall.

The purpose of Experiment 2 was to replicate the differences between high- and low-frontal older adults but without using confidence ratings on the first test. Although confidence ratings were important in eliminating guesses on the first recall test, they also helped participants monitor which items were guesses and therefore may have minimized the impact of forced report cued recall on subsequent tests (Henkel, 2004; Meade & Roediger, 2006). Experiment 2 is necessary in order to reconcile the results of Experiment 1 with the larger forced report cued recall literature. Furthermore, the confidence ratings in Experiment 1 may have differentially affected high-frontal and low-frontal older adults, given that previous research by Roediger and Geraci (2007) demonstrated that these groups differ in their ability to monitor information to reduce errors. In Experiment 2, we predicted that differences between free and forced report cued recall would emerge when participants were not required to monitor their forced report cued recall performance in their initial test. In particular, the prediction was that forced report cued recall would greatly inflate errors on later tests for low-frontal older adults relative to an initial free report test.

An additional purpose of Experiment 2 was more subtle. In Experiment 1, we found that recall of list items did not differ across groups (see Table 1). However, accepting these data at face value ignores the fact that the groups differed in error rates on critical items, and so, if some correction were possible for erroneous responding, differences may well appear between groups. No such correction is possible in the data of Experiment 1 because the critical items were always the first 5 items in the category norms, and the studied items were items 6–22. We changed the method in Experiment 2 to assess more directly memorability of the lists across participant groups. In particular, we altered the nature of the categorized lists in order to allow a comparison of items when they were presented relative to when they were not presented. Rather than leaving out the first 5 exemplars, as was done in Experiment 1, we selected the first 10 most common exemplars and divided them into sets of even and odd items. We then rotated the two sets of exemplars across participants so that each set was presented to some participants and not presented to others. In this way, we could assess performance on the same item when it was presented relative to when it was not presented, thus controlling for guessing (see Huff, Meade, & Hutchison, 2011; Meade & Roediger, 2009; and Roediger, 1973, for similar corrections).

METHOD

Participants

The participants were 32 Washington University undergraduates who participated in the experiment in partial fulfillment of a class requirement or for $10 cash, and 64 older adults (32 high-frontal older adults and 32 low-frontal older adults) who received $15 for their participation. Because the population of high-frontal older adults was limited, the same older adults who had participated in Experiment 1 also participated in Experiment 2, albeit at a delay of several months or more. Even though the pool of low-frontal older adults was substantially larger than the pool of high-frontal older adults, the low-frontal participants in Experiment 2 were also the same ones who participated in Experiment 1. In addition, 2 new high-frontal and 2 new low-frontal older adults were added to the participant sample for Experiment 2 in order to accommodate counterbalancing associated with the new list structure. Readers may worry about practice effects from Experiment 1 influencing the data for Experiment 2. However, we took several precautions to minimize the potential impact of any practice effects in Experiment 2. Participants from the free report cued recall condition in Experiment 1 participated in the forced report cued recall condition in Experiment 2 and vice versa. In addition, participants received no feedback on their performance in Experiment 1, and the minimum time lag between participation in Experiments 1 and 2 was 3 months. Importantly, no participants indicated remembering any details of having participated in Experiment 1 (these participants had participated in multiple experiments in the same building). The young adults who participated in Experiment 2 were different from those in Experiment 1 because most served for credit in classes, and the pool of students differed between semesters. Finally, the set of materials used in Experiment 2 was different from that used in Experiment 1. With these procedures in place, we think that it is unlikely that practice effects from Experiment 1 would explain the pattern of data seen in Experiment 2.

Design

The design consisted of a 3 × 2 between-subject design. Participant group (young adults, high-frontal older adults, and low-frontal older adults) and retrieval condition (free report cued recall or forced report cued recall on the first recall test) were manipulated between participants, such that there were 16 in each group. The dependent variables were veridical and false recall and recognition.

Materials

Materials were selected from Meade and Roediger (2009). The first 22 exemplars of the following categories were selected from Battig and Montague (1969) to create new lists: kitchen utensils, occupations or professions, sports, parts of a building, musical instruments, and fish. For each list, the first 10 exemplars were items for the measurement purpose described earlier, so they were separated into two sets of 5 each. One set contained exemplars numbered 1, 3, 5, 7, and 9, and the other set contained exemplars numbered 2, 4, 6, 8, and 10. One set was presented for each participant along with 12 other items, for a total of 17 items in each category. The two versions of each list were created by alternating the set of presented exemplars (5 included items and 5 omitted items). This list structure contrasts with that of Experiment 1, where exemplars numbered 1–5 were always nonpresented critical items. Thus all lists contained 17 studied items (presented in random order), with 5 nonpresented critical items.

Procedure

The procedure of Experiment 2 was identical to the procedure of Experiment 1 except that participants did not provide confidence ratings on Recall Test 1. Briefly, after preliminary instructions, participants saw 102 items belonging to common categories (17 items in each of 6 categories) with lists presented grouped by category (but items randomly presented within the category). They then took either an initial free report cued recall test (with a warning against guessing) or a forced report cued recall test (requiring them to guess). No confidence ratings were made on this test. Then all participants received a second test under free report cued recall instructions and then took a final recognition test that required them to monitor for sources of presentation for each target word, as in Experiment 1. Participants were thanked and debriefed at the end of the experiment.

RESULTS AND DISCUSSION

One purpose of Experiment 2 was to derive a more sensitive measure of recall by using two alternative sets of items. As a measure of corrected recall, we took the difference in proportion of items recalled in the two sets of five items. That is, if the odd items were presented, we subtracted the error rates of the even items from correct recall of the odd items (i.e., recall of the nonstudied critical items was subtracted from recall of the presented critical items). Besides this corrected recall measure, the data were also analyzed in the same manner as in Experiment 1, with items collapsed across list items and presented critical items. The two different analyses yielded similar results across Recall Tests 1 and 2.

Recall 1

Table 4 presents the mean proportion of list items, presented critical items, nonpresented critical items, and corrected recall for young adults and high- and low-frontal older adults on Recall Test 1 under free report cued recall or forced report cued recall instructions. In free report cued recall, the three groups of participants did not differ much in recall of list items, but older adults intruded more nonpresented critical items than did young adults. In forced report cued recall where guessing was encouraged, the three groups did not differ in recall of list items or nonpresented critical items.

LIST ITEMS

A univariate ANOVA with participant group as a between-participant factor was computed on the mean proportion of list items recalled under free report cued recall instructions. Replicating Experiment 1, young adults, high-frontal older adults, and low-frontal older adults all recalled similar numbers of list items, $F < 1.4, p > .05$. Similarly, an ANOVA on the proportion of list items produced under forced report cued recall instructions revealed no difference as a function of participant group, $F < 1.0, p > .05$. This finding is in contrast to our finding in Experiment 1, in which younger adults produced a higher proportion of list items under forced report cued recall instructions. This difference may be explained by the change in taxonomic frequency of the critical items between experiments. In Experiment 2, by guessing some of the top associates, participants may have improved their list recall and reduced critical recall (in contrast to Experiment 1, in which the first items in the norms were critical items). The change in list structure across experiments should affect all participants, but the impact may be greater for older adults, who are more likely than young adults to guess on forced report tests (Huff et al., 2011) and also show declines in output monitoring (Pansky et al., 2009). Note that the numerical trends are similar in both experiments.

PRESENTED CRITICAL ITEMS

The first 10 items from the norms were counterbalanced across participants so that 5 of the most common exemplars were presented to participants (presented critical items), and 5 were not presented to participants (nonpresented critical items). Separate univariate ANOVAs revealed that under free report cued recall, participant groups did not differ in the number of presented critical items recalled, $F < 1.0, p > .05$. However, the participant groups did differ in the proportion of presented critical items produced under forced report cued recall instructions, $F(2, 45) = 4.54, MSE = .04$. Specifically, under forced report cued recall instructions, young adults produced a greater proportion of presented critical items than did both low-frontal older adults, $t(30) = 3.01, SEM = .03$, and high-frontal older adults, $t(30) = 2.11, SEM = .03$. High-frontal older adults did not differ reliably from low-frontal older adults, $t < 1.0, p > .05$.

NONPRESENTED CRITICAL ITEMS

Under free report cued recall instructions, differences emerged between participant groups in false recall, $F(2, 45) = 7.10, MSE = .02$. Further examination of this main effect revealed that high-frontal older adults did not differ reliably from young adults, $t(30) = 1.92, SEM = .04, p > .05$, or low-frontal older adults, $t(30) = 1.81, SEM = .06, p > .05$, although the means were in the expected direction. The main effect was driven by the difference between young adults and low-frontal older adults, $t(30) = 3.90, SEM = .05$. Note that in Experiment 1, high-frontal older adults' false recall was equated with young adults' false recall. However, in Experiment 2, when the critical items had lower taxonomic frequency (lower output dominance items, on average) and the initial test was taken without confidence ratings, high-frontal older adults' false recall fell between that of young adults and low-frontal older adults.

TABLE 4. Mean Proportion of List and Critical Items Recalled by Young Adults, High-Frontal Older Adults, and Low-Frontal Older Adults on Free Report Cued Recall or Forced Report Cued Recall Test 1, Experiment 2

		Older adults	
	Young adults	High frontal	Low frontal
	Recall	Recall	Recall
Free report cued recall			
List	.33	.30	.27
Presented critical	.63	.59	.60
Nonpresented critical	.13	.21	.31
Corrected	.50	.38	.29
Forced report cued recall			
List	.48	.47	.45
Presented critical	.83	.77	.73
Nonpresented critical	.66	.58	.59
Corrected	.17	.19	.14

Under forced report cued recall instructions, an ANOVA computed on the mean proportion of critical items produced revealed a significant main effect of participant group, $F(2, 45) = 5.28$, $MSE = .01$. Young adults produced more critical items than did both high-frontal older adults, $t(30) = 2.86$, $SEM = .03$, and low-frontal older adults, $t(30) = 2.99$, $SEM = .02$. No differences were obtained between the high- and low-frontal older adults, $t < 1.0$.

CORRECTED RECALL

Corrected recall was measured so as to control for guessing of the most typical categorical exemplars. Corrected recall was derived by subtracting the proportion of nonpresented critical items produced from the proportion of presented critical items. Looking first at free report cued recall, a univariate ANOVA revealed a significant main effect of participant group, $F(2, 45) = 7.21$, $MSE = .03$. More specifically, young adults demonstrated greater corrected recall than did high-frontal older adults, $t(30) = 2.13$, $SEM = .04$, and low-frontal older adults, $t(30) = 3.44$, $SEM = .03$. High-frontal and low-frontal older adults did not differ from each other, $t(30) = 1.87$, $SEM = .03$, $p = .07$.

A separate ANOVA was conducted on the mean corrected recall of items produced under forced report cued recall instructions. The ANOVA revealed no difference, $F < 1.0$, suggesting that young adults and high- and low-frontal older adults showed equivalent corrected recall under forced report cued recall.

TABLE 5. Mean Proportion of List and Critical Items Recalled and Mean Proportion of Highest Confidence Ratings (CR4) Produced by Young Adults, High-Frontal Older Adults, and Low-Frontal Older Adults on Recall Test 2, Experiment 2

		Older adults	
	Young adults	High frontal	Low frontal
	Recall (CR4)	Recall (CR4)	Recall (CR4)
Prior free report cued recall			
List	.32 (.28)	.29 (.21)	.27 (.20)
Presented critical	.60 (.49)	.57 (.44)	.57 (.45)
Nonpresented critical	.12 (.03)	.22 (.06)	.30 (.16)
Corrected	.48	.35	.27
Prior forced report cued recall			
List	.32 (.26)	.36 (.27)	.30 (.23)
Presented critical	.60 (.48)	.62 (.55)	.60 (.48)
Nonpresented critical	.23 (.03)	.30 (.14)	.31 (.13)
Corrected	.37	.32	.29

Recall Test 2

The mean proportion of list and critical items produced on Recall Test 2 along with the proportion of highest confidence ratings are presented in Table 5 as a function of participant group and prior retrieval condition. Recall Test 2 was always a free report cued recall test, and we were interested in the relative impact of prior free report cued recall or prior forced report cued recall on this subsequent test.

LIST ITEMS

No differences emerged in list item recall (a similar conclusion was reached from data aggregated across list items and presented critical items). A 3 (young adults, high-frontal older adults, or low-frontal older adults) × 2 (prior free report cued recall or prior forced report cued recall) between-subject ANOVA computed on the mean proportion of list items recalled on Test 2 revealed no effect of participant group, $F(2, 90) = 1.60$, $MSE = .01$, $p > .05$, prior retrieval condition, $F(2, 90) = 2.39$, $SEM = .01$, $p > .05$, or any interaction between participant group and prior retrieval condition, $F < 1.0$, $p > .05$. List item recall was not influenced by participant group after free report cued recall, an outcome that replicates Experiment 1. On the other hand, list item recall did not differ between participant groups after forced report cued recall, a finding that differs from that obtained in Experiment 1.

PRESENTED CRITICAL ITEMS

The proportion of presented critical items recalled did not vary as a function of participant group or prior retrieval condition, $F\text{s} < 1.2$, $p\text{s} > .05$. Young adults and high- and low-frontal older adults all recalled similar proportions of presented critical items after prior free report cued recall and forced report cued recall.

NONPRESENTED CRITICAL ITEMS

Participants were more likely to recall critical items on Recall Test 2 after prior forced report cued recall ($M = .28$) than after prior free report cued recall ($M = .21$), $F(2, 90) = 5.22$, $MSE = .02$. Furthermore, there was a main effect of participant group, F(2,

90) = 7.78, *MSE* = .02. Low-frontal older adults (*M* = .30 collapsing across retrieval condition) and high-frontal older adults (*M* = .26 collapsing across retrieval condition) both were more likely to falsely recall the critical items than were young adults (*M* = .17), $t(30) = 4.10$, *SEM* = .03; $t(30) = 2.96$, *SEM* = .03, respectively. Contrary to our predictions, no difference emerged between the high-frontal older adults and the low-frontal older adults, $t < 1.5$, $p > .05$, although the means were in the expected direction, with greater false recall rates for low-frontal older adults than for high-frontal older adults. There was no interaction between participant group and prior retrieval condition, $F < 1.0$, $p > .05$. Low-frontal older adults were not dramatically more disadvantaged by having taken a first test under forced report cued recall conditions compared with high-frontal older adults.

CORRECTED RECALL

A separate 3 (young adults, high-frontal older adults, or low-frontal older adults) × 2 (prior free report cued recall or prior forced report cued recall) between-subject ANOVA computed on corrected recall revealed a main effect of participant group, $F(2, 90) = 6.00$, *MSE* = .03. Low-frontal older adults (*M* = .28 collapsing across retrieval condition) and high-frontal older adults (*M* = .34 collapsing across retrieval condition) both had lower corrected recall than young adults (*M* = .43), $t(30) = 3.30$, *SEM* = .05; $t(30) = 2.11$, *SEM* = .04, respectively. Also, there was no difference between the high-frontal older adults and the low-frontal older adults, $t < 1.5$, $p > .05$. There was no main effect of retrieval condition, $F < 1.1$, $p > .05$. Participants were just as likely to produce critical items on Test 2 after prior free report cued recall (*M* = .37) as after prior forced report cued recall (*M* = .33). Finally, the interaction was not significant, $F < 1.5$.

CONFIDENCE RATINGS

For each item produced on the free report cued recall test, participants were asked to indicate their confidence that the item had appeared in the study list. The mean proportion of items assigned a confidence rating of 4 (the highest level confidence) is also displayed in Table 5. Because the second recall test was a free report cued recall test with instructions not to guess, participants typically assigned items produced with high confidence ratings. The results from the confidence ratings follow the same pattern as the list items recalled and so are not discussed further.

Recognition

The mean proportions of veridical and false recognition as a function of participant group and prior retrieval condition are presented in Table 6. As in Experiment 1, veridical recognition was defined as the proportion of times that participants attributed list items to having occurred in the study list ("list" alone or in conjunction with "self" responses). False recognition is the proportion of times that participants attributed critical items as having occurred in the study list ("list" alone or in conjunction with "self" responses). A 3 (young adults, high-frontal older adults, or low-frontal older adults) × 2 (prior free report cued recall or prior forced report cued recall) between-subject ANOVA conducted on the proportion of veridical recognition revealed no significant main effect of participant group, $F(2, 90) = 2.90$, *MSE* = .01, $p > .05$; no effect of prior retrieval condition, $F < 1.0$, $p > .05$; and no interaction, $F < 1.0$, $p > .05$. The failure to find differences between participant groups in recognition replicates Experiment 1 and is consistent with other research demonstrating little or no age difference on recognition tests (e.g., Balota et al., 2000).

Turning to recognition of critical items, group differences did emerge, $F(2, 90) = 5.43$, *MSE* = .06. Replicating Experiment 1, low-frontal older adults (*M* = .72 collapsing across prior retrieval condition) falsely recognized significantly more critical items than did high-frontal older adults (*M* = .58 collapsing across prior retrieval condition), $t(30) = 2.66$, *SEM* = .06, and younger adults (*M* = .54 collapsing across prior retrieval condition), $t(30) = 4.15$, *SEM* = .05. Interestingly, there was no main effect of prior retrieval condition, $F < 1.0$, suggesting that participants were about as likely to falsely recognize critical items after free report cued recall (*M* = .60 collapsing across participant group) as they were to falsely recognize critical items after forced report cued recall (*M* = .62 collapsing across participant groups). This finding contradicts the results of Experiment 1 and suggests that any lasting impact of generating intrusions on a forced report cued recall test may depend on the nature of the intrusions, because the critical items in Experiment 2 were not the most common ex-

TABLE 6. Mean Proportion of Veridical and False Recognition for Correct and Critical Items Produced by Younger Adults, High-Frontal Older Adults, and Low-Frontal Older Adults, Experiment 2

	Prior free report cued recall			Prior forced report cued recall		
		Older			Older	
Source judgments	Young	High	Low	Young	High	Low
			List items			
List only	.23	.34	.27	.17	.17	.33
Both list and self	.64	.54	.66	.69	.74	.60
Total veridical recognition	.87	.88	.93	.87	.91	.93
Self only	.00	.01	.00	.09	.04	.03
Neither list nor self	.13	.11	.07	.05	.05	.04
			Critical items			
List only	.39	.29	.37	.19	.18	.30
Both list and self	.14	.25	.34	.35	.42	.42
Total false recognition	.53	.54	.72	.54	.61	.72
Self only	.03	.00	.03	.28	.19	.11
Neither list nor self	.46	.44	.25	.18	.21	.16

emplars for each category, as they were in Experiment 1. High–taxonomic frequency categories may cause greater confusion or reality monitoring difficulties because they are more easily generated and highly accessible.

GENERAL DISCUSSION

The current experiments examined potential differences in veridical and false memory between young adults, high-frontal older adults, and low-frontal older adults across free report cued recall and forced report cued recall tests. Under free report cued recall instructions, participant groups did not differ on veridical recall, but they did vary on false recall. Across experiments, low-frontal older adults demonstrated the highest levels of false recall and false recognition, findings consistent with past studies demonstrating that older adults' increased susceptibility to false memory may be mediated by frontal function (e.g., Butler et al., 2004). Interestingly, the relative performance of high-frontal older adults varied across experiments. In Experiment 1 (when critical items had high taxonomic frequency and confidence ratings were reported on the initial test), high-frontal older adults' false recall was identical to younger adults' false recall, whereas in Experiment 2 (when critical items had lower taxonomic frequency and confidence ratings were not reported), high-frontal older adults' level of false recall was between that of young adults and low-frontal older adults. One possible interpretation for high-frontal older adults' lower false recall rates in Experiment 1 may be that the process of providing confidence ratings helped minimize errors (cf. Henkel, 2004), assuming that high-frontal older adults were especially able to use the confidence ratings to reduce errors relative to low-frontal older adults (cf. Roediger & Geraci, 2007).

When the first test was a forced report cued recall test, high-frontal older adults produced the same numbers of veridical and false items as low-frontal older adults. In other words, the frontal differences evident on the initial free report cued recall test were not evident on the initial forced report cued recall test, although age differences were evident on forced report cued recall (both high-frontal and low-frontal older adults produced fewer veridical items than young adults in Experiment 1 and more false items than young adults in Experiment 2). Most important for the current purposes is that under forced report instructions, frontal function did not mediate older adults' production of critical items. When the results of free report cued recall and forced report cued recall are considered together, it appears that low-frontal older adults show source monitoring difficulties even on a free report cued recall task, whereas high-frontal older adults show source monitoring difficulties only on the more demanding forced report cued recall task. Note that such conclusions are based on cued recall, and it remains an open and interesting question whether similar results would be obtained on free recall tests of categorized word lists. We suspect the answer is "yes," because in a related study by Butler et al. (2004) on frontal differences in free recall of associated word lists, low-frontal older adults showed greater false recall than high-frontal older adults and young adults.

Also interesting is the effect of prior free report cued recall or forced report cued recall on recall and recognition. As suggested by Meade and Roediger (2006), source discriminations after forced report cued recall may be especially difficult because list items and critical items were both generated and recalled on a previous test. Therefore, the fact that the item was written down is not necessarily indicative that the item is old. To the extent that source monitoring after forced report cued recall is difficult and thus entails recruitment of frontal regions, we expected to see larger differences between high-frontal older adults and low-frontal older adults after forced report cued recall than after free report cued recall. Yet contrary to our predictions, high-frontal older adults were the only group in Experiment 1 to demonstrate the expected difference in false recall between free report and forced report procedures. In other words, after prior free report cued recall, high-frontal older adults demonstrated the same levels of false recall as young adults, but after prior forced report cued recall, high-frontal older adults demonstrated the same levels of false recall as low-frontal older adults (and higher than for young adults). In Experiment 2, once guessing was controlled and initial confidence ratings were eliminated, high- and low-frontal older adults still did not differ from each other in the level of false recall after the forced report cued recall task. Across experiments, then, we failed to obtain evidence that low-frontal older adults were especially disadvantaged by prior forced report cued recall. These findings suggest that source monitoring difficulties alone may not fully explain the role of frontal functioning on older adults' susceptibility to false recall in the categorized list paradigm.

An alternative interpretation of our findings may be that high-frontal older adults are more accurate on the initial free report cued recall test because here they are allowed to use their own strategies for recalling items on a memory test. In this view, we would expect to see frontal differences on free report cued recall. However, forced report cued recall should eliminate any frontal differences because item production is determined by the experimenter rather than by each participant (Koriat & Goldsmith, 1996). The data obtained in Experiment 1 (showing that high-frontal older adults were the only group to differ in critical recall on Test 2) are consistent with such an idea and suggest that high-frontal older adults may differ in their initial accuracy of items produced or their subjective monitoring of this output, especially when, as in Experiment 1, they are allowed to use confidence judgments to improve accuracy (cf. Roediger & Geraci, 2007).

The idea that strategic differences in retrieval may be responsible for the relationship between older adults' frontal function and their susceptibility to false memory is further supported by the recognition data obtained in the current study. Across both experiments, low-frontal older adults demonstrated the highest levels of false recognition after both free report cued recall and forced report cued recall. This finding replicates the results from Roediger and Geraci (2007) in showing that high-frontal older adults can use source cues to reduce memory errors on a source monitoring recognition test, whereas low-frontal older adults cannot reduce errors. Also relevant to our examination of strategy differences is the relationship between high-frontal older adults and younger adults. After prior free report cued recall instructions, high-frontal older adults actually demonstrated lower levels of false recognition than did young adults in Experiment 1 and levels of false recognition equivalent to those of young adults in Experiment 2. The finding that high-frontal older adults demonstrated lower false recognition than young adults suggests that they might be conservative in the items they call old. Assuming high-frontal older adults can rely on their own retrieval strategies, they may show very accurate memory performance, but under forced report cued recall instructions, high-frontal older adults can no longer use their preferred retrieval strategies, and so they may perform more similarly to low-frontal older adults.

In summary, the current studies were the first to compare the influence of frontal functioning on older adults' susceptibility to false memories on free report cued recall and forced report cued recall tests. Consistent with previous research (Butler et al., 2004; Chan & McDermott, 2007; Roediger & Geraci, 2007), results from free report cued recall revealed that high-frontal older adults had levels of false recall similar to those of younger adults. This difference persisted across a subsequent recall and source monitoring recognition test, thus suggesting that the age effect on false recall was driven largely by low-frontal older adults. Importantly, the frontal differences evident under free report cued recall instructions differed from the frontal patterns obtained under forced report cued recall instructions. After forced report cued recall, high- and low-frontal older adults performed similarly to each other, and both demonstrated higher false recall levels than young adults. On recognition, however, the high-frontal older adults were again able to use the source cues to reduce errors. Considered together, the results support the idea that frontal functioning plays an important role in older adults' susceptibility to false memory and further illuminate possible differences in frontal function as they relate to source monitoring and strategic processing.

NOTES

This research was supported by National Institute of Aging grant AG17481 to the third author. We thank Keith Hutchison and Rick Block for helpful comments.

Address correspondence about this article to Michelle L. Meade, Department of Psychology, Montana State University, P.O. Box 173440, Bozeman, MT 59717 (e-mail: mlmeade@montana.edu).

REFERENCES

Ackil, J. K., & Zaragoza, M. S. (1998). Memorial consequences of forced confabulation: Age differences in susceptibility to false memories. *Developmental Psychology, 34,* 1358–1372.

Balota, D. A., Cortese, M. J., Duchek, J. M., Adams, D. A., Roediger, H. L., McDermott, K. B., & Yerys, B. E. (1999). Veridical and false memories in healthy older adults and in dementia of the Alzheimer's type. *Cognitive Neuropsychology, 16,* 361–384.

Balota, D. A., Dolan, P. O., & Duchek, J. M. (2000). Memory changes in healthy young and older adults. In E. Tulving & F. I. M. Craik (Eds.), *Handbook of memory* (pp. 395–410). New York, NY: Oxford University Press.

Battig, W. F., & Montague, W. E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology Monograph, 80,* 1145.

Benton, A. L., & Hamsher, K. de S. (1976). *Multilingual aphasia examination manual.* Iowa City: University of Iowa.

Brewer, W. F. (1977). Memory for the pragmatic implications of sentences. *Memory & Cognition, 5,* 673–678.

Butler, K. M., McDaniel, M. A., Dornburg, C. C., Price, A. L., & Roediger, H. L. III. (2004). Age differences in veridical and false recall are not inevitable: The role of frontal lobe function. *Psychonomic Bulletin & Review, 11,* 921–925.

Chan, J. C. K., & McDermott, K. B. (2007). The effects of frontal lobe functioning and age on veridical and false recall. *Psychonomic Bulletin & Review, 14,* 606–611.

Craik, F. I. M. (1983). On the transfer of information from primary to secondary memory. *Philosophical Transactions of the Royal Society of London, B302,* 341–359.

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.

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

Delis, D. C., Kramer, J., Kaplan, E., & Ober, B. A. (1987). *The California Verbal Learning Test.* San Antonio, TX: Psychological Corp.

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

Glisky, E. L., 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.

Hart, R. P., Kwentus, J. A., Wade, J. B., & Taylor, J. R. (1988). Modified Wisconsin Card Sorting Test in elderly normal, depressed, and demented patients. *Clinical Neuropsychologist, 2,* 49–56.

Hashtroudi, S., Johnson, M. K., & Chrosniak, L. D. (1989). Aging and source monitoring. *Psychology and Aging, 4,* 106–112.

Henkel, L. A. (2004). Erroneous memories arising from repeated attempts to remember. *Journal of Memory and Language, 50,* 26–46.
Henkel, L. A. (2007). The benefits and costs of repeated memory tests for young and older adults. *Psychology and Aging, 22,* 580–595.
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.
Huff, M. J., Meade, M. L., & Hutchison, K. A. (2011). Age-related differences in guessing on free and forced recall tests. *Memory, 19,* 317–330.
Jacoby, L. L. (1999). Ironic effects of repetition: Measuring age-related differences in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 3–22.
Johnson, M. K., Foley, M. A., & Leach, K. (1988). The consequences for memory of imagining in another person's voice. *Memory & Cognition, 16,* 337–342.
Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114,* 3–28.
Kelley, C. M., & Sahakyan, L. (2003). Memory, monitoring, and control in the attainment of memory accuracy. *Journal of Memory and Language, 48,* 704–721.
Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review, 103,* 490–517.
Loftus, E. F. (1992). When a lie becomes memory's truth. *Current Directions in Psychological Science, 1,* 121–123.
Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 19–31.
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.
McDermott, K. B., & Chan, J. C. K. (2006). Effects of repetition on memory for pragmatic inferences. *Memory & Cognition, 34,* 1273–1284.
McIntyre, J. S., & Craik, F. I. M. (1987). Age differences in memory for item and source information. *Canadian Journal of Psychology Special Issue: Aging and Cognition, 41,* 175–192.
Meade, M. L., & Roediger, H. L. III. (2006). The effect of forced recall on illusory recollection in younger and older adults. *American Journal of Psychology, 119,* 433–462.
Meade, M. L., & Roediger, H. L. III. (2009). Age differences in collaborative memory: The role of retrieval manipulations. *Memory & Cognition, 37,* 962–975.
Mitchell, K. J., Johnson, M. K., & Mather, M. (2003). Source monitoring and suggestibility to misinformation: Adult age-related differences. *Applied Cognitive Psychology, 17,* 107–119.
Multhaup, K. S. (1995). Aging, source, and decision criteria: When false errors do and do not occur. *Psychology and Aging, 10,* 492–497.
Pansky, A., Goldsmith, M., Koriat, A., & Pearlman-Avnion, S. (2009). Memory accuracy in old age: Cognitive, metacognitive, and neurocognitive determinants. *European Journal of Cognitive Psychology, 21,* 303–329.
Raz, N. (2000). Aging of the brain and its impact on cognitive performance: Integration of structural and functional findings. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging & cognition* (2nd ed., pp. 1–90). Mahwah, NJ: Erlbaum.
Roediger, H. L. III. (1973). Inhibition in recall from cueing with recall targets. *Journal of Verbal Learning and Verbal Behavior, 12,* 644–657.
Roediger, H. L. III, & Geraci, L. (2007). Aging and the misinformation effect: A neuropsychological analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 321–334.
Roediger, H. L. III, & McDaniel, M. A. (2007). Illusory recollections in older adults: Testing Mark Twain's conjecture. In M. Garry & H. Hayne (Eds.), *Do justice and let the sky fall: Elizabeth F. Loftus and her contributions to science, law, and academic freedom* (pp. 105–136). Hillsdale, NJ: Erlbaum.
Roediger, H. L. III, & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 803–814.
Roediger, H. L., Wheeler, M. A., & Rajaram, S. (1993). Remembering, knowing and reconstructing the past. In D. L. Medin (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 30, pp. 97–134). New York, NY: Academic Press.
Wechsler, D. (1981). *Wechsler Adult Intelligence Scale: Revised manual.* New York, NY: Psychological Corp.
Wechsler, D. (1987). *Wechsler Memory Scale: Revised manual.* New York, NY: Psychological Corp.