Psychology and Aging
2010, Vol. 25, No. 1, 193–207

0882-7974/10/$12.00 DOI: 10.1037/a0018198

Aging and Flexible Remembering: Contributions of Conceptual Span, Fluid Intelligence, and Frontal Functioning

Alaitz Aizpurua
University of Minnesota and University of the Basque Country

Wilma Koutstaal
University of Minnesota

Aging attenuates the capacity to adaptively and flexibly use episodic memory at different levels of specificity. Older and younger adults were tested on a picture recognition task that required them to make episodic memory decisions at an item-specific (verbatim) versus category-based (gist-based) level on randomly intermixed trials. Specificity modulation was assessed using a measure of the likelihood that participants retrieved verbatim information in order to reject test items that were categorically related to studied items under item-specific recognition instructions (recollection rejection). We found that this measure positively correlated with conceptual span (an index of short-term semantic memory) and with level of fluid intelligence in older and younger adults. However, when we simultaneously considered each of four possible contributors (age, conceptual span, fluid intelligence, and frontal function), the only significant predictor of recollection rejection was the composite fluid intelligence measure (assessed by the Culture Fair Intelligence Test [Cattell & Cattell, 1960] and the Wechsler Adult Intelligence Scale—Revised Block Design subtest [Wechsler, 1981]). These findings suggest that interventions that facilitate adaptive specificity modulation in episodic memory may enhance the flexibility of thinking, and vice versa, in both older and younger adults.

Keywords: flexible remembering, conceptual span, fluid intelligence, memory, aging

A person's ability to adaptively deal with novel problems and to creatively respond to unexpected opportunities fundamentally depends on his or her capacity to mentally represent and access experiences at multiple levels of specificity. Although highly specific memory is crucial in some contexts (e.g., in eyewitness memory and medicine), the ability to recall and use knowledge in a more abstract or categorical manner is essential in enabling people to transfer what they learn to new situations, to making appropriate inferences, and to many other complex forms of thought, such as using analogies and metaphorical thinking (e.g., Luria, 1968/1987; Koedinger, Alibali, & Nathan, 2008).

Individuals can mentally represent and retrieve information at different levels of grain or specificity, varying from category-based or gist-based memory on one side to item-specific or verbatim memory on the other side (e.g., Ackerman & Goldsmith, 2008; Brainerd & Reyna, 1990; Brainerd, Reyna, & Mojardin, 1999; Goldsmith, Koriat, & Pansky, 2005; Reyna & Brainerd, 1995). Previous research has shown that people can strategically regulate the degree of specificity of the information on which they base their memory decisions (e.g., Goldsmith et al., 2005; Koutstaal, 2006; Malmberg & Xu, 2007; Reyna & Kiernan, 1994). Furthermore, it has been found that the tendency to rely on more or less specific representations in a given situation varies as a function of an individual's characteristics, such as age (e.g., Koutstaal, 2006; Reder, Wible, & Martin, 1986; Tun, Wingfield, Rosen, & Blanchard, 1998), emotional state (e.g., Raes et al., 2005; Williams et al., 2007), and expertise (e.g., Johnson & Mervis, 1997; Kulatunga-Moruzi, Brooks, & Norman, 2001). To date, however, little is known about how the ability to adaptively modulate the level of specificity in episodic memory is related to other complex cognitive processes.

There is suggestive evidence that a too-exclusive reliance on representations at one level of specificity—either primarily categorical or primarily item-specific—may lead to impairments in complex problem solving. For example, individuals who are clinically depressed or suicidal show impairments in social problem solving that are associated with their tendency to predominantly rely on categorical autobiographical memory (e.g., Raes et al., 2005; Williams et al., 2007). Conversely, some individuals with superior memory for verbatim information have been observed to experience difficulties in, for instance, understanding language and expressions that depend on more figurative and categorical aspects, such as metaphorical thinking and analogical reasoning (e.g., Luria, 1968/1987; Parker, Cahill, & McGaugh, 2006; Treffert & Christensen, 2005).

The main purpose of this study was to determine how specificity modulation in episodic memory is related to other cognitive aptitudes that contribute to complex thinking and how these relations

Alaitz Aizpurua, Department of Psychology, University of Minnesota and Psychology Faculty, University of the Basque Country, Gipuzkoa, Spain; Wilma Koutstaal, Department of Psychology, University of Minnesota.

This research was supported by a postdoctoral grant from the Education, Universities, and Research Department of the Basque Government to Alaitz Aizpurua. We thank research assistants who participated in the research described. We are grateful to Henk J. Haarmann, Eddy J. Davelaar, and Marius Usher for providing us with the stimuli for the original conceptual span tasks and for suggestions for the modified versions developed here.

Correspondence concerning this article should be addressed to Alaitz Aizpurua, Psychology Faculty, University of the Basque Country, Avendida Tolosa 70, Donostia-San Sebastian, 20.018, Gipuzkoa, Spain. E-mail: alaitz.aizpurua@ehu.es

may change with age. We focused particularly on three broad measures: conceptual span (an index of semantic short-term memory [STM]), fluid intelligence, and frontal-lobe functioning (FLF) level. We characterize each of these measures and our predictions below but first briefly describe the flexible remembering task.

Flexible Remembering and Recollection Rejection

To obtain a measure of specificity modulation in episodic memory, we employed the flexible remembering task (Koutstaal, 2006). In this task, participants are first shown pictures of common objects under incidental encoding conditions, and are then tested for their memory of the items under two different types of recognition instruction that vary in a pseudorandom, intermixed order. Test items include items that were initially shown during the encoding phase (*same exemplars*), alternate instances of the items that were shown at encoding (*different exemplars*), and new items not related to the initially shown items (*unrelated*). During testing, for some items participants are asked to make an identical or item-specific recognition judgment, designating an item as *old* only if they encountered exactly the same item before. For other items, participants are asked to make a conceptual or category-based recognition judgment, designating as *old* both previously presented items and conceptually similar items, that is, items belonging to the same semantic category (cf. Brainerd et al., 1999; Reder et al., 1986; Reyna & Kiernan, 1994). The ability to flexibly move between these two levels of episodic specificity in memory retrieval has been shown to be diminished—but not eliminated—in older compared with younger adults (Koutstaal, 2006).

A key measure that will be examined here is what has been termed *recollection rejection* or a process whereby retrieval of verbatim information from an earlier encounter enables an individual to appropriately reject (as not previously encountered) stimuli or events that are similar to previously studied items (Brainerd, Reyna, Wright, & Mojardin, 2003). In recollection rejection, individuals who are presented with a new stimulus (e.g., the word *sofa*) that is semantically or categorically similar to a studied target item (e.g., the word *couch*) are nonetheless able to appropriately reject the lure when required to make item-specific judgments because they recall the verbatim information that was presented (even though the probe itself is a good cue for gist memory), and can use that verbatim information to reject the probe. Recollection rejection—verbatim processing (designated as V)—was measured directly from the data, by applying the formula derived by Brainerd and Reyna (2002, p. 160, Equation 12), for use in situations involving only two recognition instruction conditions (conceptual and identical recognition). For comparison, we also calculated the corresponding index of gist memory processing, or G (Brainerd & Reyna, 2002, p. 160, Equation 13), though our primary interest relates to recollection rejection; we also report measures of sensitivity and response bias.

Conceptual Span

Despite broad consensus regarding the important roles played by STM and working memory in enabling effective cognitive functioning and general intelligence (e.g., Gray, Chabris, & Braver, 2003; Kane, Hambrick, & Conway, 2005; Jaeggi, Buschkuehl, Jonides, & Perrig, 2008), there is little evidence on the role of semantic or conceptual STM, or what has been termed *conceptual span*. Similarly, although there are some findings documenting the important contribution of semantic STM to language processing (Haarmann, Davelaar, & Usher, 2003), the relation between semantic STM and retention in long-term memory (LTM) remains unclear.

To test whether semantic STM capacity predicts specificity modulation in the flexible remembering task, we developed two modified versions of the conceptual span task. The conceptual span task was designed by Haarmann et al. (2003) to orient participant's memory maintenance and retrieval processes to stimulus meaning, with the aim to maximize the engagement of STM processes while minimizing the contribution from LTM. One version of the task, the clustered format, presents lists of 12 words, semantically clustered by category, with four consecutive words for each of three different semantic categories (e.g., *cloud, snow, drought, rain, kiwi, grape, lemon, cherry, bike, truck, taxi, subway*). In contrast, the second version, the nonclustered format, presents lists of nine randomly ordered words pertaining to three different semantic categories (e.g., *hockey, nickel, parrot, tennis, rugby, owl, peso, quarter, crow*). In both versions of the task, immediately after the list is presented, the participant is asked to recall only the words from one cued semantic category (e.g., *money*), providing the words in any order (i.e., regardless of input order).[1] Standard STM tasks (e.g., forward digit span) require the immediate literal–serial repetition of unrelated items in a fixed order and thus maximize the contribution of phonological rehearsal to performance. In contrast, the conceptual span task requires the immediate recall of words pertaining to one cued semantic category (out of several presented), without regard for the initial order of presentation of the words. This procedure thus maximizes the role of semantic processing, that is, the maintenance and processing of meanings, and reduces reliance on phonological STM (because verbatim sequential repetition of the input is not required).

Two additional aspects of the conceptual span task merit emphasis. First, the words are presented at a relatively fast rate (one item per second) in order to minimize the participants' ability to encode the items into a script-type representation in LTM, thereby more extensively engaging STM (Cowan, 2001; Haarmann & Usher, 2001). Nevertheless, this presentation rate is thought to be slow enough to enable semantic encoding of individual words in semantic STM (Potter, 1993). Second, participants are asked to read, out loud, all of the words twice immediately before starting the test, and during the test all of the items that are presented are taken from this fixed word pool. Such repeated presentation is likely to induce proactive interference, which affects retrieval from LTM more than from STM (Craik & Birtwistle, 1971; Halford, Maybery, & Bain, 1988) and which may, therefore, further help to promote the use of STM rather than LTM during the task. Following Cowan (2001; see also Postle, 2006; Lewis-Peacock & Postle, 2008), we regard STM as the capacity-limited, activated

[1] In the clustered version, there are four words per category to prevent ceiling levels of performance, which may result when small clusters of mutually supportive, adjacent items are stored in STM (cf. Haarmann & Usher, 2001).

part of LTM and assume that proactive interference affects retrieval of inactive representations in LTM.

Several studies have supported this two-component model of STM, providing strong evidence for the existence of a lexical-semantic component in STM (e.g., Haarmann & Usher, 2001; Martin & Freedman, 2001; Potter, 1993). Initial research using the conceptual span task found that semantic STM predicted successful context maintenance in the *AX*-Continuous Performance Test (Braver et al., 2001)—a common measure of cognitive control—and also predicted performance for on-line language comprehension and thinking tasks such as the detection of semantic anomalies and verbal problem solving (Haarmann, Ashling, Davelaar, & Usher, 2005; Haarmann et al., 2003; but see Kane & Miyake, 2007). Given that the flexible remembering task requires continuous maintenance of the task demands, and changing reliance on semantic representations in response to the instructions on each trial, we hypothesized that conceptual span also would be positively correlated with specificity modulation (recollection rejection) on the flexible remembering task.

Additionally, we expected significant differences between younger and older adults on the conceptual span tasks. Haarmann et al. (2005) reported higher average conceptual span in younger than in older adults in a nonclustered version of the task. However, as these authors also noted, conceptual span scores for the nonclustered version may be determined, at least in part, by an individual's ability to semantically group the relevant words during recall, that is, clustering ability, which may itself be influenced by aging (e.g., Taconnat et al., 2009). Successful performance on the nonclustered version requires reorganization of the randomly intermixed words at retrieval, whereas in the clustered version greater environmental support is provided because the words are already grouped by semantic category. In this experiment, we examined if there are age differences on the clustered version, which provides the clearest index of semantic STM capacity (Haarmann et al., 2003) but that, to date, has not been used with older adults. The clustered task was therefore administered before the nonclustered version. We also report the results from a nonclustered version of the task that was administered later in the experiment; however, given that a clean test of the magnitude of age differences on the clustered versus nonclustered tasks would require a between- rather than within-subjects manipulation of version type, with comparisons made for only the first-administered version, we do not here compare age effects on the nonclustered task.

Fluid Intelligence and Fluency Measures

General fluid intelligence (or Gf) is the ability to reason about and to solve new problems independently of previously acquired knowledge (Horn & Cattell, 1967; Stuart-Hamilton, 1996). This ability enables people to adapt their thinking to novel circumstances and is critical for a wide variety of cognitive and learning tasks (e.g., Colunga & Smith, 2008; Gray & Thompson, 2004). Fluid intelligence capacity generally declines with increasing age, whereas crystallized intelligence is typically well preserved (e.g., Horn & Cattell, 1967; Schaie & Willis, 1993). In the current study, we explored the relationship between flexible remembering of recently experienced events and measures of both fluid intelligence and fluency.

Fluid intelligence was measured by the Culture Fair Intelligence Test (Cattell & Cattell, 1960), a well-accepted index of analogical reasoning (Colom & Garcia-Lopez, 2003), and the Block Design subtest from the Wechsler Adult Intelligence Scale—Revised (WAIS-R; Wechsler, 1981), a well-documented measure of visual-spatial thinking that reflects general fluid ability (Kaufman, 2001). These two measures are strongly associated (e.g., Salthouse, Atkinson, & Berish, 2003; Tranter & Koutstaal, 2008). We also included three indices of different sorts of fluency, including visual–spatial fluency (Ruff Figural Fluency Test; Evans, Light, & Ruff, 1987), verbal (letter) fluency (Controlled Oral Word Association Test; Spreen, & Benton, 1977), and ideational fluency (Alternative Uses Task; Guilford, Christensen, Merrifield, & Wilson, 1978). In fluency tests, participants are typically asked to generate items under timed and limited search conditions, and the most common fluency measure derived is the number of correct items generated. Given that fluid intelligence and fluency tasks assess cognitive flexibility (e.g., Evans et al., 1987), we hypothesized that individuals achieving higher scores on these tests would demonstrate more effective specificity modulation on the flexible remembering task, independently of their age.

In addition to the total number of acceptable responses that are produced, fluency tasks also provide the opportunity to evaluate the number of errors (e.g., repetitions), and the number of times the individual switches between different strategies. Using a letter fluency task in an older adult sample, Koutstaal (2006) found a positive relation between the number of words generated to the letter *F* and a slightly different measure of specificity modulation on the flexible remembering task. Based on this finding, it was expected that there would be a significant positive correlation between the recollection rejection measure and letter fluency across age groups. Moreover, consistent with the idea that phonemic fluency is a multifactorial task that depends on both clustering and switching processes (e.g., Hughes & Bryan, 2002; Troyer, Moscovitch, & Winocur, 1997), we analyzed these two components in the verbal fluency task. Clustering refers to the number of words generated within a phonemic subcategory (e.g., *salad*, *salute*, *salt*). Switching refers to the shifting from one cluster of words to another and is thought to rely on strategic processes in the search for new and appropriate clusters (i.e., cognitive flexibility; Troyer, 2000). Thus, we speculated that there would be a positive correlation between the number of switches in the verbal fluency task and scores of recollection rejection, in both older and younger adults.

The Alternative Uses Task has been used as a measure of divergent thinking in the context of creativity and problem solving (e.g., Gilhooly, Fioratou, Anthony, & Wynn, 2007). In this task, we asked participants to produce as many uses as possible for three different common objects (e.g., a chair). Later, following an experimental procedure introduced by Gilhooly et al. (2007), we also asked participants to identify which of the generated uses were *old* (based on their previous experiences, and so drawing on LTM) and which of the uses they gave were *new* (newly generated for the first time during the experimental session). We speculated that flexibility in producing diversity of ideas could be related to flexibility in memory retrieval (Eslinger & Grattan, 1993). More specifically, given that generating new uses is believed to place high demands on executive processes (Gilhooly et al., 2007), we predicted that the frequency of new uses would be positively correlated with the measure of recollection rejection.

Finally, both avoiding errors (repetitions and/or intrusions) in the fluency tasks and avoiding false recognition responses to different exemplars on item-specific probe trials may depend on similar processes of cognitive control (e.g., Salthouse et al., 2003; Troyer et al., 1997). Thus, we predicted a positive association between number of errors in the figural and verbal fluency tasks and false recognition of different exemplars under the identical-recognition instructions.

Frontal Lobe Functioning

The frontal lobes are presumed to subserve such higher cognitive functions as strategic search, complex decision making, online maintenance of abstract rules, and mental flexibility (e.g., Duncan et al., 2000; Eslinger & Grattan, 1993; Hughes & Bryan, 2002; Troyer et al., 1997). With regard to memory, in particular, frontal cortex not only participates in encoding (e.g., binding item and contextual information; Mitchell, Johnson, Raye, Mather, & D'Esposito, 2000), but is central to criterion setting and monitoring during retrieval (e.g., Melo, Winocur, & Moscovitch, 1999; Schacter, Norman, & Koutstaal, 1998). For example, lesion evidence suggests that inhibitory processes rely on the integrity of frontal cortex (e.g., Floden & Stuss, 2006).

The flexible remembering task requires inhibition of the tendency to respond only on the basis of verbatim information if the current instructional cue requires categorically based recognition decisions, and, contrariwise, inhibition of the tendency to respond on the basis of gist information if the current instructional cue requires item-specific recognition decisions. As noted earlier, using a test of letter fluency that is known to reflect FLF, Koutstaal (2006) found a strong positive correlation between the number of items that participants generated to the letter *F* and a measure of specificity modulation on the flexible remembering task in older adults. This suggests that the ability to intentionally switch between using item-specific and category-based information could be associated with FLF level. In the current study, we employed a more extensive neuropsychological battery designed to assess FLF (Glisky, Polster, & Routhieaux, 1995), and hypothesized that recollection rejection would be higher in participants with higher FLF scores. In addition, considering the age-related deficits in verbatim recollection, we also expected differences between older and younger adults within high- and low-FLF groups.

Previous research has shown that the ability to remember in a category-based manner remains intact in older adults (e.g., Castel, 2005; Koutstaal, 2006; Koutstaal & Schacter, 1997; Koutstaal, Schacter, Gallucio, & Stofer, 1999; Reder et al., 1986), whereas on tests of precise recollection, older adults are more likely to accept as true false information that is similar to actually presented information (e.g., Aizpurua, Garcia-Bajos, & Migueles, 2009; Koutstaal, 2003; Koutstaal et al., 1999; Schacter, Koutstaal, Johnson, Gross, & Angell, 1997). This likelihood of false memory has been correlated with FLF not just in older adults (Butler, McDaniel, Dornburg, Price, & Roediger, 2004; see also LaVoie, Willoughby, & Faulkner, 2006), but also in younger adults (Chan & McDermott, 2007). In these studies, participants with relatively low-FLF level showed higher false-memory susceptibility than did high-FLF individuals. Specifically, Butler et al. (2004) found higher proportions of false recall in older adults with low-FLF compared with those with high-FLF, with high-FLF older adults' performance being comparable with that of younger adults. However, these researchers assumed that all younger adults uniformly have high-FLF. More recently, Chan and McDermott (2007) found that FLF scores correlated negatively with proportions of false recall in both older and younger adult groups. In the present study, we expected FLF level to predict the proportion of false alarms for the identical-recognition instruction trials (especially to different exemplars) in both older and younger adults.

Overview

In summary, we examined measures of flexible remembering in older and younger adults in relation to conceptual span, fluid intelligence–fluency, and frontal function. Our primary predictions were that (a) older adults would show lower conceptual span than younger adults, and conceptual span would correlate with recollection rejection on the flexible remembering task in both age groups; (b) fluid intelligence and measures of fluency (total fluency, switching, and newly generated items on the Alternative Uses Task) would be lower in older than in younger adults and would correlate with recollection rejection; and (c) recollection rejection would be greater in those achieving high than low frontal functioning scores and, within both high and low frontal-function groups, would be greater in younger than in older adults; furthermore, across age groups, frontal functioning would be correlated with false recognition on the identical-recognition instruction trials. Additionally, in a final across-task analysis, we sought to examine the relative effectiveness of four factors—age, conceptual span, fluid intelligence, and frontal function—in predicting levels of recollection rejection on the flexible remembering task.

Method

Participants

The participants were 36 younger adults recruited through posted flyers at the University of Minnesota and 36 older adults recruited, through e-mail and posted notices, from a Retirees Volunteer Center and Lifelong Learning Institute. All participants reported being native speakers of English and having normal or corrected-to-normal vision, color vision, and hearing; all were compensated $10/hr for taking part in the study.

As can be seen from Table 1, on average, younger adults (29 women) had fewer years of education, $t(70) = 5.85$, and lower scores on the WAIS-R (Wechsler, 1981) Vocabulary Test, $t(70) = 3.46$, than did older adults (26 women), indicating that older adults were generally high functioning. Both older and younger adults rated their subjective state of health as close to excellent on a 7-point scale ranging from 1 (*very poor*) to 7 (*excellent*): older adults, $M = 6.35$ ($SD = 0.56$); younger adults, $M = 6.25$ ($SD = 0.81$).

All participants were screened for depression and anxiety using the Brief Symptom Inventory (Derogatis & Melisaratos, 1983), and for several medical conditions that could affect their cognitive performance: open heart or bypass surgery, uncontrolled high blood pressure, Parkinson's or nervous system disease, stroke or Transient Ischemic Attack, loss of consciousness lasting more than

Table 1
Demographic Characteristics and Cognitive Performance of Older and Younger Adults

Characteristic	Older *M*	Older *SD*	Younger *M*	Younger *SD*
Age	69.94	5.51	21.31	2.35
Years of education	16.86	1.81	14.58	1.48
Vocabulary score	53.33	5.10	48.72	6.16
Mini-Mental State Exam	29.17	1.09	—	—

Note. Vocabulary score comes from the Wechsler Adult Intelligence Scale—Revised. Dashes indicate that data were not obtained.

5 min, diabetes, mental or emotional problems for which they were admitted to the hospital, and alcohol or drug abuse. Any participant reporting one or more of these conditions and/or obtaining a score of 11 or higher on the depression-related and/or anxiety-related items in the Brief Symptom Inventory was excluded. In addition, the Mini-Mental State Exam (Folstein, Folstein, & McHugh, 1975) was used as a screening for cognitive state of older adults, and only individuals who scored 27/30 or higher were included (M = 29.17, SD = 1.09).[2]

Materials

Flexible remembering task (Koutstaal, 2006). The stimuli were colored photographs or detailed line drawings of common objects and animals (see Figure 1). There were a total of 240 object–exemplar pairs (plus practice items); 120 items (one each from 120 pairs) were presented at study. The test consisted of 360 items, including 120 items from each of three item types: previously presented items (same exemplars), categorically related items (different exemplars), and entirely new (unrelated) items. Items of each type and the type of recognition-test instruction (item specific or category based) occurred in pseudorandom order, with equivalent numbers of each item and test type within each sixth of the test. For any one participant, the type of test instruction was held constant for a given exemplar pair (e.g., whistle-1 and whistle-2 would both be tested either using item-specific or category-based probes). The items presented at study versus at test, and the type of test instruction to which they were assigned, were counterbalanced across participants, and any given item was tested only once per participant. Brief breaks were given after each third of the test. Stimuli were presented on a personal computer with a color monitor.

Conceptual span task. Two approximately parallel forms of the conceptual span task were created, one administered in clustered format (and here always administered first, for the reasons outlined in the introduction), and the other administered in nonclustered format. The semantic category norms utilized in the original task (Battig & Montague, 1969), but updated by Van Overschelde, Rawson, and Dunlosky (2004), were used to choose the categories and items. Each version consisted of words drawn from six different semantic categories, with eight items per category. Specifically, modified versions of the original tasks developed by Haarmann et al. (2003) were created, in which several important aspects of the material were controlled. First, both the category names and all of the items within the categories were selected to have only one or two syllables. Second, the most common item in the category provided in the norms from Van Overschelde et al. (2004) was excluded (with one exception: for the category of *jobs*, the word *doctor* was used). In addition, category size was controlled, and generational differences between the 1969 and 2004 populations for categories were minimized. Furthermore, the frequency, familiarity, and concreteness of the items were controlled using the Thorndike–Lorge database. Differences were not statistically significant between the clustered versus nonclustered versions, respectively, in category size (M = 1.48, SD = 0.24 and M = 1.49, SD = 0.02), generational difference (M = −0.08, SD = 0.04 and M = −0.10, SD = 0.05), written frequency, (M = 192.76, SD = 121.05 and M = 147.58, SD = 83.19), familiarity (M = 523.05, SD = 43.76 and M = 513.14, SD = 29.75), or concreteness (M = 589.73, SD = 30.09 and M = 587.36, SD = 19.51). The two versions also were broadly matched for the semantic categories included (e.g., one version included mammals, whereas the other included birds). For the clustered version, changes in the categories and items from Haarmann et al. (2003) were minimized. The Appendix provides the categories and items for the clustered and nonclustered versions of the conceptual span task.

Measures of fluid intelligence and fluency. We examined five different measures of fluency of thinking and fluid intelligence: (a) Culture Fair Intelligence Test (Cattell & Cattell, 1960); (b) Block Design subtest from the WAIS-R (Wechsler, 1981); (c) Ruff Figural Fluency Test (Evans et al., 1987); (d) the Controlled Oral Word Association Test (COWAT; Spreen & Benton, 1977); and (e) a modified form of the Alternative Uses Task (Guilford et al., 1978).

The Culture Fair Intelligence Test includes visual tests that require the participant to perceive relationships between shapes and figures. In the present study, Forms A and B of Scale 2 were used, each consisting of four subtests that require 3, 4, 3, and 2.5 min for completion, respectively: Series, Classification, Matrices, and Conditions. For example, in the Classifications subtest, participants are shown 14 problems of five abstract shapes and figures and are asked to select which, out of five, does not match or belong with the others. The participant's score was calculated summing all the correct answers produced within the time limit.

For the Block Design subtest, the participant uses blocks to replicate two-color designs within a specified time limit. Each block has two white, two red, and two half-red and half-white sides. The 13 designs progress in difficulty from simple designs requiring four blocks (60 s allotted) to more complex designs requiring nine blocks (120 s allotted). The participant's score was based on the correctness and the time (number of seconds) taken to construct the designs.

The Ruff Figural Fluency Test consists of five parts, each containing a different spatial configuration of five dots, repeated throughout the page. The participant is asked to connect two or more of the five dots with straight lines, with the goal of being able

[2] Two older adults and three younger adults were excluded because they did not meet one or more of the screening criteria. Another two older and three younger adults completed only the first of the three experimental sessions and thus were replaced.

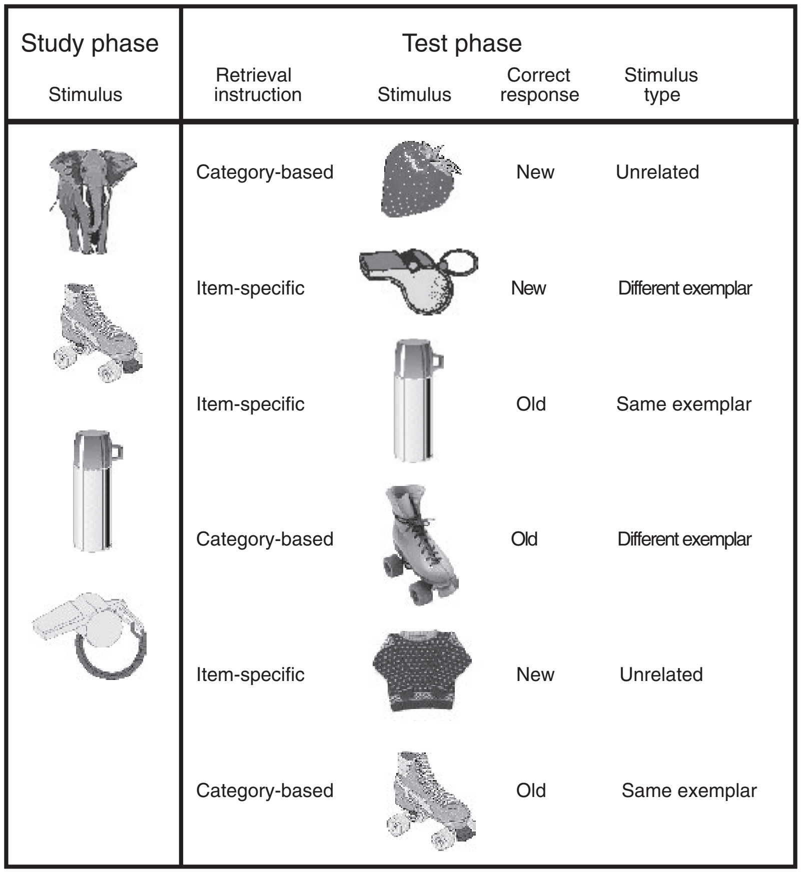

Figure 1. Schematic representation of the study and recognition tasks in the flexible remembering task.

to draw as many unique designs as possible during the allotted time (1 min per part). For purposes of this study, the total number of unique designs and the number of perseverative errors (i.e., repetitions of the designs) were calculated.

The ideational fluency task was a modified form of the Alternative Uses Task, in which individuals are given the name of a common object (e.g., *newspaper*) and are asked to generate alternative, nonstandard uses for the object (e.g., in addition to its common use as for reading, a newspaper also might be used to start a fire). The Alternative Uses Task used in this study included two phases. First, in the production phase, three common object items (chair, bed sheet, wooden pencil) were consecutively presented (on separate sheets); participants had 3 min per sheet to write down as many alternative uses for each object as possible that were both different from the common use and different from one another. An example was given before the task. Second, after completion of the production phase, participants were asked to review each of the uses they had generated and to indicate (by circling) those uses that they had first thought of while doing the task, that is, uses they had never seen or heard before, either in their own experience directly or in films, books, television, and so forth. The response protocols were scored using the task manual, and new uses generated across the three objects were also calculated.[3]

For the verbal fluency task participants were given 60 s to verbally generate as many English words as possible, beginning with either the letter *F*, *A*, or *S*. Participants' responses were audiotaped (provided the participant consented to being recorded) to allow later crosschecking and confirmation of the experimenter's written responses. Numbers, proper names, and variants of the same word (e.g., the same word with different suffixes, such as *thank* and *thankful*), as well as repetitions, were considered errors. The participant's score for the COWAT[4] was the total number of acceptable words generated for the three letters, but we also examined the number of clusters and switches, and the number of errors produced.

FLF battery. The FLF battery (see Glisky et al., 1995) included five neuropsychological test scores: the number of categories achieved on the computerized version of the Wisconsin Card Sorting Test (Hart, Kwentus, Wade, & Taylor, 1988); the total number of words generated for the letters *F*, *A*, and *S* on the COWAT (Spreen & Benton, 1977); the Mental Arithmetic score

[3] For *pencil*, a similar scoring key was created. The same rater scored all of the protocols, but two-thirds (80%) of the protocols were scored by a second independent rater. Interrater reliability, calculated with the Pearson product–moment correlation coefficient, was very high, $r(58) = .99$.

[4] The same rater scored all of the protocols, but two-thirds (80%) of the protocols were scored by a second independent rater. Interrater reliability was very high, $r(58) = .98$.

from the Wechsler Adult Intelligence Scale—III[5] (Wechsler, 1997a); and the Mental Control and Backward Digit Span scores from the Wechsler Memory Scale—III (Wechsler, 1997b). FLF scores were computed by taking the average of the five component test scores, with each of the test scores first standardized (*z* scored) on the total sample. To maximize the differences between the high and low frontal functioning groups, while also still maintaining a sufficient number of participants in each group, we contrasted the lower versus upper third of the overall distribution. The 33rd percentile of the FLF score was –.13, and the 66th percentile was .23 for the total sample. To form high and low frontal functioning groups, participants with FLF scores equal to, or above, the top third were placed in the high-FLF group ($M = 0.58$, $SD = 0.26$, range = 0.23 to 1.03), whereas participants with FLF scores below the bottom third were placed in the low-FLF group ($M = -0.64$, $SD = 0.40$, range = –1.57 to –.13). Table 2 presents the average scores on the five components for the four groups.[6]

Procedure

Participants were tested individually in three 90-min sessions (rest breaks were given when needed), and gave their written consent to participate in the experiment. Age-related differences may vary according to time of testing and self-reported time-of-day preference (e.g., May, 1999), so participants selected their own preferred times to be tested. The average time between each of the sessions was approximately 3.5 days, with no differences in the average interval for older and younger adults.

In the first session, the flexible remembering task was presented in two separate phases. First, in the encoding phase, participants performed a size-judgment task, in which they indicated whether the real-world referent of the object shown was larger than a 1-ft (30.48-cm) square box (an example box was provided). The items were presented for 2 s. This incidental encoding task was followed by a brief filler task not involving pictures. Second, for the episodic recognition test, participants were informed about the two types of recognition decisions that they would be asked to make, and were told that the type of decision required would be indicated by an instruction cue stating either "identical" or "conceptual," presented immediately before, and concurrently with, the presentation of each item. Examples of identical and conceptually related items were provided. The recognition test was self-paced. All responses were indicated using designated keys on the computer keyboard. Finally the Mental Control and Block Design subtests were administered.

In the second session, first the Mental Arithmetic test was administered, followed by the conceptual span task, in the clustered format. Then the participant completed the Vocabulary subtest, the Culture Fair Intelligence Test, and, finally, the Alternative Uses Task.

In the third session, first the conceptual span task, in the nonclustered format, was administered, followed by the COWAT. Then the participant completed the Ruff Figural Fluency Test, the Digit Span subtest, and, finally the Wisconsin Card Sorting Test.

Results

For all statistical analyses, the significance level (α) was set at $p < .05$, unless otherwise noted.

Age-Related Differences in the Flexible Remembering Task

Older and younger adults were equally accurate in the size-judgment task ($M_{older} = 0.91$, $SD_{older} = 0.04$; $M_{younger} = 0.90$, $SD_{younger} = 0.04$). Recollection rejection (*V*) was significantly lower in older adults ($M = 0.38$, $SD = 0.18$) than it was in younger adults ($M = 0.58$, $SD = 0.16$), $F(1, 69)$[7] = 19.77, suggesting that older adults were less able to retrieve verbatim information regarding the studied items to enable rejection of categorically related objects. In contrast, the measure of *G*, or the estimated probability that distractors would elicit retrieval of gist traces, was much higher in older adults ($M = 0.88$, $SD = 1.06$) than in younger adults ($M = 0.25$, $SD = 0.37$), $F(1, 69) = 11.27$.

Hits. Table 3 provides hits, false alarms, measures of sensitivity and response criteria, and the measure of verbatim recollection separately for older and younger adults and for the two types of test instructions. There were not significant differences between older and younger adults in the proportion of hits either for the identical or for the conceptual test probes (*F*s < 1). Thus, there were no age-related differences in correct category-based recognition. For the conceptual probes, there was a recognition advantage for same exemplars versus different exemplars (.89 vs. .84, respectively); this same-exemplar advantage replicates previous findings (e.g., Brainerd et al., 1999; Koutstaal & Cavendish, 2006; Reyna & Kiernan, 1994), and may partially reflect what Brainerd and Reyna (1993) termed a "verbatim exit bias" (p. 52), or a tendency of individuals to make similarity judgments based on surface form (rather than meaning) for studied items. Notably, both age groups showed significantly higher meaning-based recognition for identical items than for categorically related items, $F(1, 68) = 61.18$. This equivalent same-exemplar advantage across age groups is consistent with previous findings (Koutstaal, 2003, 2006), and indicates that older adults were as able as were younger adults to encode at least some of the item-specific details associated with the pictures.

False alarms. False alarms could be assessed for different exemplars on the identical test-instruction trials, and for unrelated items for both identical and conceptual test-instruction trials. Overall, participants were more likely to commit false alarms for different exemplars than for unrelated items (.21 vs. .06, respectively), indicating that conceptual and/or perceptual similarity between lures and targets elicited higher false recognition in both age groups, $F(1, 68) = 228.02$. However, false alarms to related lures on identical test-instruction trials were nearly twice as frequent in older compared with younger adults, suggesting (as shown by the analyses of *V* and *G*) impairments in item-specific differentiation and greater reliance on gist-based information in older adults, *F*(1,

[5] The version of Mental Arithmetic used here differs from the WAIS-R version (Wechsler, 1981) in that it has one new item, and three other items are different. However, this version has one of the highest correlations with Working Memory of all the tests in the Wechsler Adult Intelligence Scale—III battery (Wechsler, 1997a).

[6] There were no significant age-related differences on any of the subcomponent measures.

[7] On the flexible remembering task, the scores for one younger adult were treated as outliers and were not included in any of the analyses involving *V* or *G*.

Table 2

Mean Performance on the Neuropsychological Tests That Contributed to the Frontal-Lobe Functioning Score, on Measures of Fluid Intelligence, on Conceptual Span, and on Recollection Rejection in Older and Younger Adults

	Older				Younger			
	High (n = 12)		Low (n = 13)		High (n = 12)		Low (n = 11)	
Test	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Frontal battery score	.60	.27	−.72	.42	.56	.25	−.55	.37
Wisconsin Card Sorting Test	6.00	0.00	5.31	1.10	6.00	0.00	5.82	0.60
Controlled Oral Word Association Test	53.50	6.56	36.92	11.81	50.58	9.36	36.00	6.89
Mental Arithmetic	17.25	2.09	11.77	2.42	15.92	3.03	12.91	4.04
Mental Control	7.92	0.29	7.46	0.66	8.17	0.39	7.09	1.04
Backward Digit Span	6.92	1.24	4.69	1.11	7.08	1.17	5.09	1.22
Culture Fair Intelligence Test	67.75	5.74	56.85	3.51	75.42	5.92	71.45	8.87
Wechsler Adult Intelligence Scale—Revised Block Design	44.42	11.25	32.31	11.35	56.42	9.26	49.55	13.61
Conceptual span	.72	.09	.64	.07	.83	.07	.74	.06
V[a]	.45	.23	.31	.14	.60	.17	.50	.19

[a] Recollection rejection *(V)* was calculated from the flexible remembering data by adding the differences between (a) designating as *old* different exemplars under conceptual versus identical recognition instructions, and (b) designating as *old* new unrelated exemplars under conceptual versus identical recognition instructions (see Brainerd & Reyna, 2002, p. 160).

68) = 29.75, for the effect of age. Considering false alarms for new items, older adults also showed a significantly higher proportion of false alarms than did younger adults on the identical-test probes, $F(1, 68) = 7.64$, but not on the conceptual-test probes, $F(1, 68) < 1.63$.

We used signal detection theory to compute analyses for measures of sensitivity (A') and corresponding measures of response bias (B''_D). The two measures of sensitivity and response bias were (a) a comparison of hits for different exemplars versus false alarms to novel (unrelated) items, termed A'_{gist} and $B''_{D\text{-}gist}$, respectively, and (b) a comparison of hits for same exemplars versus false alarms to different (related) exemplars, termed A'_{fine} and $B''_{D\text{-}fine}$, respectively. A' scores (Snodgrass & Corwin, 1988) vary from 0 to 1.0. Scores of .5 indicate chance performance, and higher scores reflect greater accuracy. B''_D scores (Donaldson, 1992) vary from –1.0 to +1.0. Zero scores indicate neutral responding, negative values indicate more lenient responding (tendency to respond *old*), and positive scores indicate more stringent responding (tendency to respond *new*).

A′ scores. Sensitivity to gist was higher under conceptual- than under identical-test probes (.87 vs. .67, respectively), in both age groups, $F(1, 68) = 281.85$. There were no age-related differences in gist sensitivity on conceptual-test probes, suggesting that older adults were as accurate as younger adults in conditions where gist information should be used (to designate as *old* both same and different exemplars, and calling *new* only the unrelated items).

Table 3

Mean Proportions of Hits and False Alarms, Mean Scores of Sensitivity and Response Criteria, and Mean Scores of Recollection Rejection in the Flexible Remembering Task

	Identical				Conceptual			
	Older		Younger		Older		Younger	
Variable	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Hits								
Same items[a]	.68	.12	.72	.11	.88	.11	.90	.08
Different items[b]					.83	.10	.85	.09
False alarms								
Different items[c]	.27	.10	.15	.09				
Unrelated items[d]	.08	.06	.04	.04	.25	.15	.20	.15
A'_{fine}	.78	.08	.85	.07	.57	.07	.58	.07
$B''_{D\text{-}fine}$	.06	.29	.34	.30	−.80	.22	−.82	.16
A'_{gist}	.71	.06	.64	.09	.86	.05	.88	.06
$B''_{D\text{-}gist}$	.88	.11	.88	.14	−.16	.42	−.10	.45
V	.38	.18	.58	.16				

Note. A'_{fine} and $B''_{D\text{-}fine}$ compare *old* responses to same versus different exemplars, and A'_{gist} and $B''_{D\text{-}gist}$ compare *old* responses to different versus new exemplars.

[a] Designating Whistle 1 as old when Whistle 1 was studied. [b] Designating both Whistle 1 and Whistle 2 as old when any whistle was studied. [c] Designating Whistle 2 as old when actually Whistle 1 was studied. [d] Designating as old an umbrella when no umbrellas were studied.

Sensitivity to fine-grained information was higher on identical- than on conceptual-test probes (.81 vs. .58, respectively), in both age groups, $F(1, 68) = 384.48$. Under identical retrieval instructions, older adults were more sensitive than were younger adults to gist information, $F(1, 68) = 11.77$; they also were less accurate than younger adults when discriminating between same and different items, $F(1, 68) = 14.04$. Thus older adults were less able to retrieve or use item-specific details to differentiate between same and different exemplars.

B''_D scores. Participants adopted much stricter criteria under identical- than conceptual-test instructions, both in differentiating same and different exemplars (.20 vs. –.81, respectively), $F(1, 68) = 674.08$, and in discriminating different from unrelated items (.88 vs. –.13, respectively), $F(1, 68) = 435.62$. There were no differences between older and younger adults in the response criteria adopted when discriminating between different and unrelated exemplars either under the conceptual- or the identical-test probes (Fs < 1). However, younger adults were more conservative than were older adults on the identical-test probes when discriminating between same and different exemplars, $F(1, 68) = 15.95$, indicating that older adults were more liberal under conditions where gist should not be used.

Recollection Rejection and the Conceptual Span Task

Figure 2 shows the average proportion recalled on the clustered and nonclustered versions of the conceptual span task for older and younger adults. On the key clustered task—that provides the purest measure of conceptual span—older adults ($M = 43.58$, $SD = 5.20$) recalled, on average, 5.5 fewer words than did younger adults ($M = 49.08$, $SD = 5.40$), $F(1, 70) = 19.37$. Recollection rejection significantly correlated with scores on the clustered version, $r(69) = .47$ (see Figure 3), confirming that conceptual span plays an important role in recollection rejection ability.

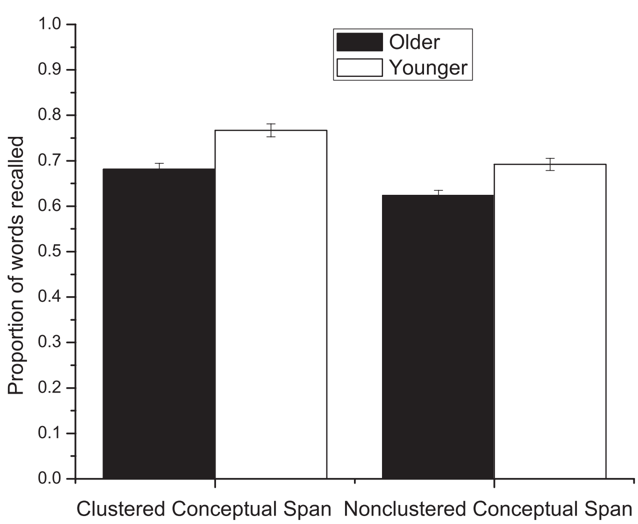

Figure 2. Mean proportion of words recalled in the clustered (left two bars) and nonclustered (right two bars) version of the conceptual span task, shown separately by age group (older vs. younger). Error bars show standard errors of the mean. Maximum scores were 64 in the clustered version and 48 in the nonclustered version.

Flexible Remembering Task and Measures of Fluid Intelligence and Fluency

Table 4 provides the average scores for the measures of fluid intelligence and fluency in older and younger adults. As shown in Table 4, older adults scored significantly lower than did younger adults on the Culture Fair Intelligence Test and WAIS-R measures of fluid intelligence and also on the figural fluency measure; additionally, they made more errors on the figural and verbal fluency tasks.

As predicted, there was a significant positive correlation between recollection rejection and the average score on the Culture Fair Intelligence Test, $r(69) = .58$, and on the Block Design subtest, $r(69) = .67$[8] (see Figure 4). In contrast, recollection rejection did not significantly correlate with Vocabulary, $r(69) = –.07$, or with Mental Arithmetic, though the latter correlation was numerically greater than the former, $r(69) = .25$. These results suggest that recollection rejection is more strongly related to thinking processes involving novel, on-line problem solving than to long-term knowledge.

Recollection rejection was also associated in the expected direction with the quantity of items produced on the fluency tasks. Specifically, it was significantly associated with the number of designs in the Ruff Figural Fluency Test, $r(69) = .35$, but only modestly and not significantly correlated with the number of words provided on the verbal fluency test, $r(69) = .18$,[9] and the number of new uses in the ideational fluency task, $r(69) = .07$, with more items produced being associated with higher recollection rejection. In the verbal fluency task, the association between the number of switches and recollection rejection was also positive but likewise not significant, $r(69) = .19$.[10] However, clustering on the COWAT and recollection rejection were significantly correlated, $r(69) = –.24$, as were clustering and false alarms for different exemplars to item-specific recognition probes, $r(70) = .31$, with more words generated in a cluster associated with less recollection rejection and more false recognition. In the ideational fluency task, since we found that number of new uses was not normally distributed, we also divided the sample into high and low performers in generating new uses; there were no differences between high and low performers in recollection rejection (.48 vs. .45, respectively) $t(69) = –.80$.

Finally, the number of errors on both the figural fluency, $r(70) = .34$, and the verbal fluency, $r(70) = .36$ test, were positively associated with the proportion of false alarms for different exemplars, indicating that more errors on these fluency tasks predicted more false recognition of related items in the flexible remembering task.

[8] Since performance on the Culture Fair and Block Design task strongly correlated, $r(70) = .77$, these scores were combined to increase power for further analysis. The correlation between the composite fluid intelligence measure and recollection rejection was $r(69) = .66$.

[9] With letter F, $r(69) = .12$; with letter A, $r(69) = .20$; and with letter S, $r(69) = .14$; with letter F in older adults, $r(34) = .10$ (all correlations *ns*).

[10] This correlation was stronger and significant when considering only letter S, $r(69) = .24$.

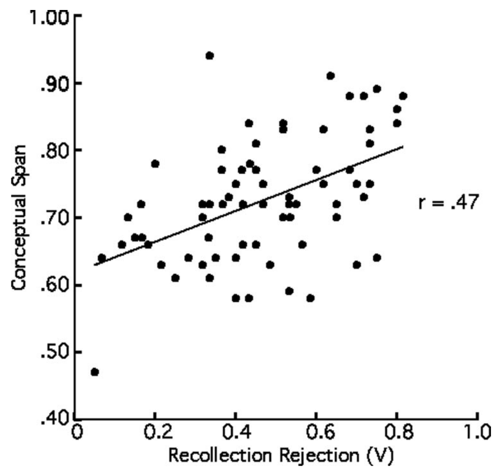

Figure 3. Correlation between recollection rejection and score on the conceptual span task (clustered version).

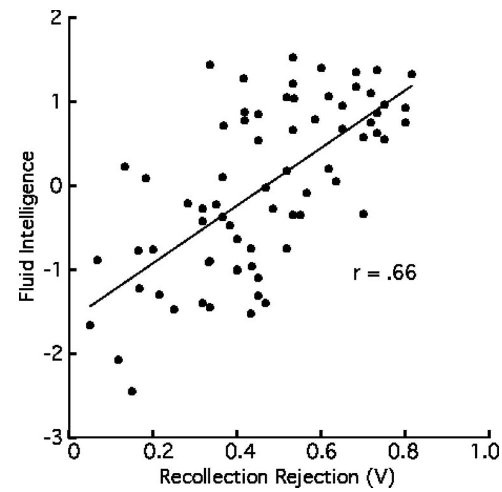

Figure 4. Correlation between recollection rejection and *z* score on fluid intelligence (combining Culture Fair Intelligence Test and the Block Design subtest of the Wechsler Adult Intelligence Scale—Revised).

Flexible Remembering and Frontal Functioning Level

Recollection rejection positively correlated with frontal function, $r(69) = .30$, such that participants with higher scores on the composite assessment of frontal functioning demonstrated an enhanced ability for recollection rejection. There were significant differences in recollection rejection between high- and low-FLF groups (.53 and .40, respectively), $F(1, 46) = 5.17$. Within each age group, high-FLF participants showed numerically higher recollection rejection than did low-FLF participants, respectively, for both older (.45 and .31) and younger (.60 and .50) adults, but these differences were not significant ($p = .07$). Age-related differences in recollection rejection also were in the predicted direction for older and younger adults, respectively, in the high-FLF group (.45 vs. .60), $F(1, 22) = 3.17$, $p = .089$, but significant only in the low-FLF group (.31 vs. .50), $F(1, 22) = 7.61$, with younger adults showing higher recollection rejection in both comparisons.

In the recognition task, FLF score was modestly but not significantly negatively correlated with false alarms to different exemplars under item-specific probes, $r(70) = -.22$, $p = .07$. No significant differences were found as a function of FLF level. However, under item-specific instructions, where attention and cognitive processes related to frontal function would be most relevant, significant age-related differences were observed within both the high- and low-FLF groups. Specifically, in the high-FLF group, older adults showed lower sensitivity to fine-grained information than did younger adults, respectively (A'_{fine}, .77 vs. .85), $F(1, 34) = 7.17$, and higher sensitivity to category-based information (A'_{gist}, .71 vs. .64), $F(1, 34) = 6.97$; they also adopted more liberal response criteria when discriminating between same and different exemplars ($B''_{D\text{-}fine}$, .10 vs. .32), $F(1, 34) = 5.26$. In the low-FLF group, results for the item-specific probes followed the same pattern on A'_{fine} (.77 vs. .84), $F(1, 34) = 10.41$, on A'_{gist} (.71 vs. .65), $F(1, 34) = 6.56$, and on $B''_{D\text{-}fine}$ (.06 vs. .36), $F(1, 34) = 8.26$ for older and younger adults, respectively.

Multiple Regression Analysis

Table 5 summarizes the simple correlations between age, recollection rejection, conceptual span, fluid intelligence, and frontal

Table 4
Mean Scores on Measures of Fluid Intelligence and Fluency in Older and Younger Adults

Measure	Older *M*	Older *SD*	Younger *M*	Younger *SD*
Culture Fair Intelligence Test**	62.81	7.01	73.58	7.00
Wechsler Adult Intelligence Scale—Revised Block Design**	38.69	12.15	54.28	10.94
Ruff Figural Fluency Test**	75.83	16.26	96.33	22.67
Errors*	11.47	15.48	5.5	4.59
Verbal fluency	43.28	11.69	42.08	10.58
Clustering	.38	.19	.36	.19
Switching	33.47	8.98	32.97	8.43
Errors*	3.39	2.63	2.17	1.58
Ideational fluency (new uses)	5.03	4.27	6.58	5.60

Note. The Controlled Oral Word Association Test (Spreen & Benton, 1977) was used as an index of verbal fluency. A modified version of the Alternative Uses Task (Guilford et al., 1978) was used as an index of ideational fluency. For differences between older and younger adults: * $p < .05$. ** $p < .001$.

Table 5
Pearson Product–Moment Correlations Between Age, Recollection Rejection, Conceptual Span, Fluid Intelligence, and Frontal Function

Variable	1	2	3	4	5
Overall					
1. Age	—				
2. Recollection rejection	−.40***	—			
3. Conceptual span	−.48***	.47***	—		
4. Fluid intelligence	−.64***	.66***	.50***	—	
5. Frontal function	−.05	.30*	.38**	.41***	—
Older adults					
1. Age	—				
2. Recollection rejection	−.09	—			
3. Conceptual span	−.31	.40*	—		
4. Fluid intelligence	−.26	.55**	.37*	—	
5. Frontal function	.04	.34*	.35*	.55**	—
Younger adults					
1. Age	—				
2. Recollection rejection	−.14	—			
3. Conceptual span	.13	.24	—		
4. Fluid intelligence	−.04	.52**	.21	—	
5. Frontal function	.13	.25	.41*	.33*	—

* $p < .05$. ** $p < .01$. *** $p < .001$.

function both overall (top panel), and within each age group separately (bottom panels). To determine their relative contribution to recollection rejection, we conducted a simultaneous multiple regression including age, conceptual span, fluid intelligence, and frontal function as predictors. The strongest and only significant predictor of recollection rejection was fluid intelligence ($\beta = .55$), $t(69) = 4.15$; the contribution of age was not significant ($\beta = .03$), $t(69) = .25$.

Discussion

We assessed the influence of aging, conceptual span, FLF, and fluid intelligence on an individual's capacity to adaptively and flexibly use episodic memory at different levels of specificity. The responses of older and younger adults during a flexible remembering picture recognition task were used to obtain a measure of recollection rejection. Recollection rejection occurs when an individual retrieves verbatim information in response to items or events that are categorically or semantically related to previously encountered items; the mismatch between the retrieved verbatim information and the similar event enables rejection of related stimuli when item-specific recognition is required (Brainerd & Reyna, 2002; Brainerd et al., 2003). Consistent with earlier findings regarding age-related differences in specificity modulation in episodic memory (Koutstaal, 2006), we found that older adults showed significantly attenuated recollection rejection compared with their younger counterparts. In addition, individual differences in recollection rejection ability were significantly associated with measures of conceptual span, frontal function, and level of fluid intelligence and fluency. Critically, when examining the predictive utility of each of these measures, we found that the measure of fluid intelligence (as assessed by the Culture Fair Intelligence Test and the WAIS-R Block Design subtest) was the most strongly and uniquely predictive contributor. Simultaneous regression showed that the contributions of age, conceptual span, and FLF to recollection rejection ability were not significant above and beyond the contribution of fluid intelligence.

Conceptual span, a measure of the capacity to actively maintain and integrate semantic representations in short-term memory, was significantly associated—when considered on its own—with recollection rejection ability. Previous research found that conceptual span predicted individual differences in thinking and language comprehension (Haarmann et al., 2005, 2003); we can conclude that it also explains variation in the ability to effectively retrieve verbatim information to reject semantically or categorically similar, but nonexperienced, events in episodic memory retrieval. This finding is consistent with fuzzy trace theory's explanation of recall of categorized lists in that verbatim memory is involved in both recall and recognition (e.g., Brainerd, Wright, Reyna, & Payne, 2002). Furthermore, conceptual span was reduced in older adults, suggesting that age-related declines in recollection rejection could be derived from lessened memory for identity information of previous stimuli (Haarmann et al., 2005). However, additional analyses revealed that, across age-groups, conceptual span correlated strongly and positively with appropriate sensitivity to gist (A'_{gist}) on the category-based recognition trials, $r(69) = .42$, but not as strongly with inappropriate gist sensitivity on the item-specific trials, $r(69) = -.17$. Thus, conceptual span was most strongly associated with differences in the component of episodic specificity modulation that involves the effective processing of meanings in response to instructions requiring categorically or conceptually based episodic recognition decisions. This finding fits well with evidence from patients with semantic dementia, in whom selective degradation of semantic knowledge is observed, and who may show impaired meaning-based recognition for perceptually different examples of the studied items but relatively preserved recognition of perceptually identical items (Graham, Simons, Pratt, Patterson, & Hodges, 2000).

Frontal functioning, when considered on its own, was also significantly positively related to recollection rejection on the flexible remembering task. Given that effective recollection rejection requires intentional alternations in the specificity of memory recognition decisions in response to changing contextual cues, this finding is consistent with the idea that the frontal lobes as a structure, and executive functions as a process, are central to an individual's ability to intentionally guide his or her own behavior (e.g., Goldberg & Podell, 2000; Pansky, Goldsmith, Koriat, & Pearlman-Avnion, 2009). As expected, frontal function was negatively correlated with the probability of false recognition, with a higher frontal-function level associated with lower probability of false alarms to different exemplars on the identical recognition probes. This outcome is generally consistent with the findings of Butler et al. (2004) and of Chan and McDermott (2007), but extends their findings to a recognition task involving visually and conceptually similar objects and incidental encoding.

We found that the ability to perform well on novel visual–spatial and reasoning tasks strongly correlated with recollection rejection, such that individuals who scored higher on measures of fluid intelligence and a measure of figural fluency showed greater recollection rejection. Given our earlier findings (Koutstaal, 2006), we also expected to find a stronger correlation between recollection rejection and verbal fluency, particularly when switching, as the component most reflecting cognitive flexibility (cf. Troyer, 2000), was examined. A across-experiment comparison revealed that variability on the verbal fluency test in older adults was higher in our earlier study than among the older participants in this study (for the letter *F*; $SD = 5.70$ vs. 4.02, respectively). This may partially explain the stronger association found in the earlier study. Also, considering measures of errors, broadly in line with the earlier findings, we found that the frequency of errors on the verbal and figural fluency tasks correlated negatively with recollection rejection and positively with false recognition.

Finally, and more importantly, when simultaneously including each of the possible contributors, we found that fluid intelligence capabilities constituted the main and only significant predictor of recollection rejection ability in both older and younger adults. This suggests that, independently of age, analogical thinking and nonverbal problem-solving skills substantially contribute to the ability to flexibly and effectively move between different levels of episodic memory retrieval. Also, although conceptual span and frontal function individually correlated with recollection rejection, when taken together, they did not here help to explain additional variation on that measure beyond that accounted for by the composite measure of fluid intelligence. Previous researchers have particularly emphasized measures of reasoning and memory as predictors of everyday cognition (e.g., Allaire & Marsiske, 1999; Kirasic, Allen, Dobson, & Binder, 1996). Although performance on tests representing both crystallized and fluid intelligence significantly relate to measures of everyday cognition, indices of fluid intelligence explain larger proportions of the variance in many studies (e.g., Cornelius & Caspi, 1987; Kirasic et al., 1996). Moreover, recent experiments have shown that participating in novel mentally stimulating activities in older adults (Tranter & Koutstaal, 2008), or tasks that systematically challenge working memory capacity in younger adults (Jaeggi et al., 2008), can lead to significant enhancements on indices of fluid intelligence. It has also been shown that training older adults to rely more extensively on recollection (rather than familiarity- or gist-based responding) in a memory recognition task led to more generalized improvements on nontrained transfer tasks, including a working memory (*n*-back) task, self-ordered pointing, source monitoring, and the Digit Symbol Substitution Test (Jennings, Webster, Kleykamp, & Dagenbach, 2005). An important practical and theoretical question therefore is whether challenging individuals with flexible remembering tasks—and thereby extending their capability to adaptively move between differing levels of specificity—would likewise facilitate flexible thinking, and enhance learning and problem solving.

In conclusion, this work has provided strong initial evidence for the contributions to flexible remembering in older and younger adults, and has shown that there is a robust positive correlation between the ability to use episodic memory at different levels of specificity and the ability to flexibly and adaptively approach particularly novel problems. However, given that the younger and older participants were mostly university educated and Caucasian, future work testing participants from a broader range of backgrounds and ability levels is needed to determine whether these results generalize to other populations. In addition, the lack of more complete screening for mild cognitive impairment could be a confounding factor in studies involving older adults, although this age-associated impairment is considered an aspect of typical aging (Petersen, 2000). It might also be noted that in this study the different tasks were administered across three experimental sessions. To the extent that cognitive flexibility fluctuates from moment to moment (e.g., Leber, Turk-Browne, & Chun, 2008) or as a result of various contextual factors such as positive affect (e.g., Dreisbach & Goschke, 2004), this could attenuate the correlations between the measures. Finally, since correlational data cannot establish causation, future studies exploring the effects of experimental manipulations of flexible remembering on flexibility of thinking—and vice versa—would be especially valuable.

References

Ackerman, R., & Goldsmith, M. (2008). Control over grain size in memory reporting: With and without satisficing knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 1224–1245.

Aizpurua, A., Garcia-Bajos, M., & Migueles, M. (2009). False memories for a robbery in young and older adults. *Applied Cognitive Psychology, 23,* 174–187.

Allaire, J. C., & Marsiske, M. (1999). Everyday cognition: Age and intellectual ability correlates. *Psychology and Aging, 14,* 627–644.

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

Brainerd, C. J., & Reyna, V. F. (1990). Gist is the grist: Fuzzy-trace theory and the new intuitionism. *Developmental Review, 10,* 3–47.

Brainerd, C. J., & Reyna, V. F. (1993). Memory independence and memory interference in cognitive development. *Psychological Review, 100,* 42–67.

Brainerd, C. J., & Reyna, V. F. (2002). Recollection rejection: How children edit their false memories. *Developmental Psychology, 38,* 156–172.

Brainerd, C. J., Reyna, V. F., & Mojardin, A. H. (1999). Conjoint recognition. *Psychological Review, 106,* 160–179.

Brainerd, C. J., Reyna, V. F., Wright, R., & Mojardin, A. H. (2003). Recollection rejection: False-memory editing in children and adults. *Psychological Review, 110,* 762–784.

Brainerd, C. J., Wright, R., Reyna, V. F., & Payne, D. G. (2002). Dual-retrieval processes in free and associative recall. *Journal of Memory and Language, 46,* 120–152.

Braver, T. S., Barch, D. M., Keys, B. A., Carter, C. S., Cohen, J. D., Kaye, J. A., . . . Mumenthaler, M. S. (2001). Context processing in older adults: Evidence for a theory relating cognitive control to neurobiology in healthy aging. *Journal of Experimental Psychology: General, 130,* 523–536.

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.

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

Cattell, R. B., & Cattell, A. K. S. (1960). *Handbook for the individual or group Culture Fair Intelligence Test.* Champaign, IL: IPAT.

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

Colom, R., & Garcia-Lopez, O. (2003). Secular gain in fluid intelligence: Evidence from the Culture-Fair intelligence test. *Journal of Biosocial Science, 35,* 33–39.

Colunga, E., & Smith, L. B. (2008). Flexibility and variability: Essential to human cognition and the study of human cognition. *New Ideas in Psychology, 26,* 174–192.

Cornelius, S. W., & Caspi, A. (1987). Everyday problem solving in adulthood and old age. *Psychology and Aging, 2,* 144–153.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24,* 87–185.

Craik, F. I. M., & Birtwistle, J. (1971). Proactive inhibition in free recall. *Journal of Experimental Psychology, 91,* 210–213.

Derogatis, L. R., & Melisaratos, N. (1983). The Brief Symptom Inventory: An introductory report. *Psychological Medicine, 13,* 595–605.

Donaldson, W. (1992). Measuring recognition memory. *Journal of Experimental Psychology: General, 121,* 275–277.

Dreisbach, G., & Goschke, T. (2004). How positive affect modulates cognitive control: Reduced perseveration at the cost of increased distractibility. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 343–353.

Duncan, J., Seitz, R. J., Kolodny, J., Bor, D., Herzog, H., Ahmed, A., . . . & Emslie, H. (2000, July 21). A neural basis for general intelligence. *Science, 289,* 457–460.

Eslinger, P. J., & Grattan, L. M. (1993). Frontal lobe and frontal-striatal substrates for different forms of human cognitive flexibility. *Neuropsychologia, 31,* 17–28.

Evans, R. W., Light, R. H., & Ruff, R. M. (1987). The Ruff Figural Test: A normative study with adults. *Developmental Neuropsychology, 3,* 37–51.

Floden, D., & Stuss, D. T. (2006). Inhibitory control is slowed in patients with right superior medial frontal damage. *Journal of Cognitive Neuroscience, 18,* 1843–1849.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-mental state: A practical method for grading the cognitive state of patients for the clinical. *Journal of Psychiatric Research, 12,* 189–198.

Gilhooly, K. J., Fioratou, E., Anthony, S. H., & Wynn, V. (2007). Divergent thinking: Strategies and executive involvement in generating novel uses for familiar objects. *British Journal of Psychology, 98,* 611–625.

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

Goldberg, E., & Podell, K. (2000). Adaptive decision making, ecological validity, and the frontal lobes. *Journal of Clinical and Experimental Neuropsychology, 22,* 56–68.

Goldsmith, M., Koriat, A., & Pansky, A. (2005). Strategic regulation of grain size in memory reporting over time. *Journal of Memory and Language, 52,* 505–525.

Graham, K. S., Simons, J. S., Pratt, K. H., Patterson, K., & Hodges, J. R. (2000). Insights from semantic dementia on the relationship between episodic and semantic memory. *Neuropsychologia, 38,* 313–324.

Gray, J. R., Chabris, C. F., & Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Nature Neuroscience, 6,* 316–322.

Gray, J. R., & Thompson, P. M. (2004). Neurobiology of intelligence: Science and ethics. *Nature Reviews Neuroscience, 5,* 471–482.

Guilford, J. P., Christensen, P. R., Merrifield, P. R., & Wilson, R. C. (1978). *Alternate uses manual.* Orange, CA: Sheridan Psychological Services.

Haarmann, H. J., Ashling, G. E., Davelaar, E. J., & Usher, M. (2005). Age-related declines in context maintenance and semantic short-term memory. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 58*(A), 34–53.

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

Haarmann, H. J., & Usher, M. (2001). Maintenance of semantic information in capacity-limited item short-term memory. *Psychonomic Bulletin & Review, 8,* 568–578.

Halford, G. S., Maybery, M. T., & Bain, J. D. (1988). Set-size effects in primary memory: An age-related capacity limitation. *Memory & Cognition, 16,* 480–487.

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.

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

Hughes, D. L., & Bryan, J. (2002). Adult age differences in strategy use during verbal fluency performance. *Journal of Clinical and Experimental Neuropsychology, 24,* 642–654.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceeding of the National Academy of Sciences, USA, 105,* 6829–6833.

Jennings, J. M., Webster, L. M., Kleykamp, B. A., & Dagenbach, D. (2005). Recollection training and transfer effects in older adults: Successful use of a repetition-lag procedure. *Aging, Neuropsychology and Cognition, 12,* 278–298.

Johnson, K. E., & Mervis, C. B. (1997). Effects of varying levels of expertise on the basic level of categorization. *Journal of Experimental Psychology: General, 126,* 248–277.

Kane, M. J., Hambrick, D. C., & Conway, A. R. A. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin, 131,* 66–71.

Kane, M. J., & Miyake, T. M. (2007). The validity of "conceptual span" as a measure of working memory capacity. *Memory & Cognition, 35,* 1136–1150.

Kaufman, A. S. (2001). WAIS–III IQs, Horn's theory, and generational changes from young adulthood to old age. *Intelligence, 29,* 131–167.

Kirasic, K. C., Allen, G. L., Dobson, S. H., & Binder, K. S. (1996). Aging, cognitive resources, and declarative learning. *Psychology and Aging, 11,* 658–670.

Koedingera, K. R., Alibalib, M. W., & Nathan, M. J. (2008). Trade-offs between grounded and abstract representations: Evidence from algebra problem solving. *Cognitive Science, 32,* 366–397.

Koutstaal, W. (2003). Older adults encode—but do not always use—perceptual details: Intentional versus unintentional effects of detail on memory judgments. *Psychological Science, 14,* 189–193.

Koutstaal, W. (2006). Flexible remembering. *Psychonomic Bulletin & Review, 13,* 84–91.

Koutstaal, W., & Cavendish, M. (2006). Using what we know: Consequences of intentionally retrieving gist versus item-specific information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 778–791.

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

Koutstaal, W., Schacter, D. L., Galluccio, L., & Stofer, K. A. (1999).

Reducing gist-based false recognition in older adults: Encoding and retrieval manipulations. *Psychology and Aging, 14,* 220–237.

Kulatunga-Moruzi, C., Brooks, L. R., & Norman, G. R. (2001). Coordination of analytic and similarity-based processing strategies and expertise in dermatological diagnosis. *Teaching and Learning in Medicine, 13,* 110–116.

LaVoie, D. L., Willoughby, L., & Faulkner, K. (2006). Frontal lobe dysfunction and false memory susceptibility in older adults. *Experimental Aging Research, 32,* 1–21.

Leber, A. B., Turk-Browne, N. B., & Chun, M. M. (2008). Neural predictors of moment-to-moment fluctuations in cognitive flexibility. *Proceedings of the National Academy of Sciences, USA, 105,* 13592–13597.

Lewis-Peacock, J. A., & Postle, B. R. (2008). Temporary activation of long-term memory supports working memory. *Journal of Neuroscience, 28,* 8765–8771.

Luria, A. R. (1987). *The mind of a mnemonist: A little book about a vast memory* (Lynn Solotaroff, Trans.). Cambridge, MA: Harvard University Press. (Original work published 1968)

Malmberg, K. J., & Xu, J. (2007). On the flexibility and fallibility of associative memory. *Memory & Cognition, 35,* 545–556.

Martin, R. C., & Freedman, M. (2001). Verbal working memory: The ins and outs of phonological and lexical-semantic retention. In H. L. Roediger III, J. S. Nairne, I. Neath, & A. M. Surprenat (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 331–349). Washington, DC: American Psychological Association.

May, C. P. (1999). Synchrony effects in cognition: The costs and benefit. *Psychonomic Bulletin & Review, 6,* 142–147.

Melo, B., Winocur, G., & Moscovitch, M. (1999). False recall and false recognition: An examination of the effects of selective and combined lesions to the medial temporal lobe diencephalon and frontal lobe structures. *Cognitive Neuropsychology, 16,* 343–359.

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

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.

Parker, E. S., Cahill, L., & McGaugh, J. L. (2006). A case of unusual autobiographical memory. *Neurocase, 12,* 35–49.

Petersen, R. C. (2000). Aging, mild cognitive impairment, and Alzheimer's disease. *Neurologic Clinics, 18,* 789–805.

Postle, B. R. (2006). Working memory as an emergent property of the brain and mind. *Neuroscience, 139,* 23–38.

Potter, M. C. (1993). Very short-term conceptual memory. *Memory & Cognition, 21,* 156–161.

Raes, F., Hermans, D., Williams, J. M. G., Demyttenaere, K., Sabbe, B., Pieters, G., & Eelen, P. (2005). Reduced specificity of autobiographical memory: A mediator between rumination and ineffective social problem-solving in major depression? *Journal of Affective Disorders, 87,* 331–335.

Reder, L. M., Wible, 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.

Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences, 7,* 1–75.

Reyna, V. F., & Kiernan, B. (1994). Development of gist versus verbatim memory in sentence recognition: Effects of lexical familiarity, semantic content, encoding instructions, and retention interval. *Developmental Psychology, 30,* 178–191.

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 Experimental Psychology: General, 132,* 566–594.

Schacter, D. L., Koutstaal, W., Johnson, M. K., Gross, M. S., & Angell, K. E. (1997). False recognition induced by photographs: A comparison of older and younger adults. *Psychology and Aging, 12,* 203–215.

Schacter, D. L., Norman, K. A., & Koutstaal, W. (1998). The cognitive neuroscience of constructive memory. *Annual Review of Psychology, 49,* 289–318.

Schaie, K. W., & Willis, S. L. (1993). Age-difference patterns of psychometric intelligence in adulthood: Generalizability within and across ability domains. *Psychology and Aging, 8,* 44–55.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117,* 34–50.

Spreen, O., & Benton, A. L. (1977). *Neurosensory Center Comprehensive Examination for Aphasia* (*NCCEA*). Victoria, British Columbia, Canada: University of Victoria, Neuropsychology Laboratory.

Stuart-Hamilton, I. (1996). Intellectual changes in late life. In R. T. Woods (Ed.), *Handbook of the clinical psychology of ageing* (pp. 23–41). Oxford, England: Wiley.

Taconnat, L., Raz, N., Tocze, C., Bouazzaoui, B., Sauzéon, H., Fay, S., & Isingrini, M. (2009). Ageing and organisation strategies in free recall: The role of cognitive flexibility. *European Journal of Cognitive Psychology, 21,* 347–365.

Tranter, L., & Koutstaal, W. (2008). Age and flexible thinking: An experimental demonstration of the beneficial effects of increased cognitively stimulating activity on fluid intelligence in healthy older adults. *Aging, Neuropsychology, and Cognition, 15,* 184–207.

Treffert, D. A., & Christensen, D. D. (2005). Inside the mind of a savant. *Scientific American, 293,* 108–113.

Troyer, A. K. (2000). Normative data for clustering and switching on verbal fluency tasks. *Journal of Clinical and Experimental Neuropsychology, 22,* 370–378.

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

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.

Van Overschelde, J. P., Rawson, K. A., & Dunlosky, J. (2004). Category norms: An updated and expanded version of the Battig and Montague (1969) norms. *Journal of Memory and Language, 50,* 289–335.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale—Revised manual.* New York, NY: Psychological Corporation.

Wechsler, D. (1997a). *Wechsler Adult Intelligence Scale—III manual.* San Antonio, TX: Psychological Corporation.

Wechsler, D. (1997b). *Wechsler Memory Scale—III manual.* San Antonio, TX: Psychological Corporation.

Williams, J. M. G., Barnhofer, T., Crane, C., Herman, D., Raes, F., Watkins, E., & Dalgleish, T. (2007). Autobiographical memory specificity and emotional disorder. *Psychological Bulletin, 133,* 122–148.

Appendix

Materials for the Clustered and Nonclustered Versions of the Conceptual Span Task

Clustered version
FRUITS: mango, cherry, grape, lemon, kiwi, pear, peach, plum
MAMMALS: lion, cow, bear, horse, deer, rabbit, pig, tiger
MUSIC: trumpet, bass, tuba, cello, oboe, harp, flute, organ
TRANSPORT: bus, truck, train, bike, boat, subway, taxi, van
JOBS: teacher, lawyer, nurse, doctor, dentist, cook, banker, athlete
WEATHER: rain, snow, hail, flood, thunder, drought, cloud, blizzard
Nonclustered version
TREES: maple, pine, birch, elm, palm, willow, cedar, aspen
BIRDS: robin, blue jay, hawk, crow, parrot, sparrow, dove, owl
MONEY: penny, dime, nickel, quarter, yen, euro, pound, peso
BUILDING: door, floor, wall, roof, stairs, ceiling, hall, entrance
SPORTS: cricket, baseball, tennis, hockey, swimming, golf, rugby, wrestling
EARTH: river, ocean, lake, valley, hill, desert, beach, cave

Received November 26, 2008
Revision received June 18, 2009
Accepted June 29, 2009 ■