American Psychological Association

Psychology and Aging

0882-7974/18/$12.00

2018, Vol. 33, No. 8, 1105–1114
http://dx.doi.org/10.1037/pag0000315

Individual Differences in Executive Functions and Retrieval Efficacy in Older Adults

Fergus I. M. Craik
Rotman Research Institute at Baycrest and University of Toronto

Eldar Eftekhari
Rotman Research Institute at Baycrest

Ellen Bialystok
York University and Rotman Research Institute at Baycrest

Nicole D. Anderson
Rotman Research Institute at Baycrest and University of Toronto

Two prominent aspects of memory problems in older adults are a difficulty in retrieving recent episodic events and an often transient inability to retrieve names and other well-known facts from semantic memory. The question addressed in the present studies was whether these age-related difficulties reflect a common cause—a retrieval problem related to inefficient executive functions (EF). In the first study, 50 older adults were given 4 tests of EF; a derived composite measure correlated strongly with a measure of retrieval efficacy in free recall, less strongly with paired-associate recall, and nonsignificantly with retrieval of general knowledge. A second study used somewhat different measures of EF and also different measures of retrieval from semantic memory, and this study did find significant relations between EF, episodic memory, and knowledge retrieval. Changes in the specific tests representing both EF and memory retrieval changed the relations between them, suggesting that no one task is a pure measure of the theoretical constructs of either EF or episodic and semantic memory. Taken together, the 2 studies showed that individual differences in EF in older adults are correlated with retrieval efficacy in both episodic and semantic memory but also that these relations depend on the specific measures chosen to represent both EF and memory retrieval.

Keywords: retrieval, executive functions, older adults, episodic memory, semantic memory

Age-related impairments in memory performance are well established in the literature on cognitive aging (Craik & Rose, 2012; McDaniel, Einstein, & Jacoby, 2008; Park & Reuter-Lorenz, 2009); the interesting questions now are related to a finer-grained analysis of the strengths and weaknesses of performance in older adults, descriptions of causation in cognitive terms, and explanations of how such causative factors relate to age-related changes in brain function. The present report focuses on memory retrieval abilities in older adults, how these abilities vary as a function of material and task differences, and how they relate to individual differences in executive functions.

Retrieval difficulties in older adults are well documented (e.g., Burke & Light, 1981; Craik & Jennings, 1992; Moscovitch & Winocur, 1992). There is some agreement that tasks requiring the greatest amounts of self initiation (Craik, 1983, 1986) show the greatest age-related impairments and that these related losses may be attributed to the declining efficiency of executive functions (Bouazzaoui et al., 2013, 2014; Hedden, Lautenschlager, & Park, 2005; Moscovitch & Winocur, 1992; Shimamura, 1995, 2002). Executive functions are generally held to reflect the integrity of frontal lobe functioning (Miyake et al., 2000; Stuss & Benson, 1986), and there is good converging evidence to relate age decrements in frontal lobe integrity to a parallel decline in executive control (Buckner, 2004; Cabeza & Dennis, 2013; Raz, 2000).

One interesting question in this general context concerns possible differences between the ability to retrieve information from episodic and semantic memory. Two of the most common complaints of older adults are forgetting of recent episodic events and a transient failure to retrieve well-known information (especially names of people and objects) from semantic memory. Are these age-related inefficiencies attributable to some common mechanism? One line of argument suggests that they are; individual

This article was published Online First December 3, 2018.

Fergus I. M. Craik, Rotman Research Institute at Baycrest, and Department of Psychology, University of Toronto; Eldar Eftekhari, Rotman Research Institute at Baycrest; Ellen Bialystok, Department of Psychology, York University, and Rotman Research Institute at Baycrest; Nicole D. Anderson, Rotman Research Institute at Baycrest, and Departments of Psychology and Psychiatry, University of Toronto.

This research was supported by a grant (A8261) from the Natural Sciences and Engineering Research Council of Canada (to Fergus I. M. Craik).

Study 1 was presented at a symposium in honor of Lynn Hasher held at University College, University of Toronto in November 2017.

We thank Tim Shallice for valuable comments on an earlier version of this article and Kevin Tang for his help in conducting Study 2.

Correspondence concerning this article should be addressed to Fergus I. M. Craik, Rotman Research Institute at Baycrest, 3560 Bathurst Street, Toronto, ON M6A 2E1, Canada. E-mail: fcraik@research.baycrest.org

differences in frontal lobe integrity in older adults should map onto differences in efficiency of executive functions and thus to differences in retrieval effectiveness. Retrieval problems in older adults are characterized by an intact ability to retrieve general global features of the sought-for information (e.g., a person's appearance and the context of meeting them) but an impaired ability to recollect specific details (e.g., the person's name and their place of employment; Craik & Bialystok, 2008; Shing, Werkle-Bergner, Brehmer, Müller, Li, & Lindenberger, 2010). This pattern appears to hold for both episodic and semantic information and has been described as a loss of resolving power associated with impaired frontal lobe functioning (Craik & Grady, 2002; Fuster, 2002). On the other hand, if episodic and semantic memory are indeed different memory systems as suggested by Endel Tulving (Tulving, 1972; Tulving & Schacter, 1990), it seems possible that retrieval mechanisms are different for episodic and semantic information.

The present two studies addressed this problem by obtaining measures of executive functions and also measures of retrieval from episodic and semantic memory in groups of older adults. In one measure of semantic memory retrieval, we used a general knowledge test, but this method has the problem of differentiating between cases of retrieval failure for information the participant knows from cases in which the participant does not know the correct answer. To address this problem of accessibility versus availability, we first asked participants various general knowledge questions in a free responding manner and then gave the same questions again in the form of a multiple-choice recognition test. The measure of retrieval efficacy was taken to be the ratio of items recalled correctly in the first phase to the total number of items correctly recognized in the multiple-choice test; only items that were both correctly recalled and recognized were included in the numerator. Correct recognition was assumed to be a rough measure of what the participant actually knows, so the recall/recognition ratio was taken to reflect the participant's success in recalling known information without further hints or clues. To obtain parallel measures of episodic retrieval efficacy, we conducted similar manipulations for verbal free recall and paired-associate learning. In both cases a recall phase was followed by a recognition phase, and the measure of retrieval efficacy was taken as the ratio of recalled words (provided they were also recognized in the later test) to the total number of words recognized. In summary, these procedures yielded three measures of retrieval efficacy (ability to recollect information they knew), two representing episodic retrieval and one representing retrieval from semantic memory.

The same participants were also given a short series of tests to measure executive function (EF) ability. The tests were chosen to reflect different aspect of EF, partly following the scheme suggested by Miyake et al. and Friedman et al. (e.g., Friedman et al., 2008; Miyake et al., 2000) and also the suggestions of Oberauer, Süß, Schulze, Wilhelm, and Wittmann (2000) regarding the structure of working memory. The tests were also chosen to reflect participants' ability to deal with a variety of different types of information. Our final selection included a version of the Stroop task (Stroop, 1935) to measure the ability to inhibit prepotent responses, as used by Friedman et al. (2008) and Miyake et al. (2000). We also included the Star Counting task (Das-Smaal, de Jong, & Koopmans, 1993), which involves keeping track of ongoing counting operations and set shifting in the course of these operations. In Study 1 we included the Alpha Span test (Craik, 1986; Craik, Bialystok, Gillingham, & Stuss, 2018) in which participants must mentally rearrange short lists of words into correct alphabetic order. Finally, Study 1 included the Consonant Updating task (Morris & Jones, 1990), used by Friedman et al. (2008) and Miyake et al. (2000) under the title Letter Memory; this task gives a measure of the updating aspects of EF. These four EF tasks used in Study 1 thus involve a variety of processes typically ascribed to working memory and EF, including the manipulation of letters, words and numbers, naming speed, and ability to inhibit prepotent stimuli. The tests are described below; following evidence for correlations among the tests, they were combined to give a composite measure of EF for each participant.

Study 1

Method

Participants. The participants were 50 adult volunteers (17 males, 33 females) aged between 60 and 70 years (mean age = 66.5 years), recruited initially through newspaper advertisements and word of mouth. On average they had received 16.8 years of formal education (SD = 3.9 years). All participants were in good mental and physical health; they lived independently in the community and traveled to our laboratory to participate in the study. They were paid $16 CAD per hour to compensate them for their time and participation. The study received ethical approval from the relevant body of the University of Toronto.

Materials and procedure.

Tests of executive function. All participants performed four tasks to measure EF: the Stroop Test, Star Counting, Consonant Updating, and Alpha Span. Scores on these tasks were then standardized and averaged to give a composite measure of executive functioning for each participant.

Stroop test. Participants were shown colored stimuli on a computer monitor and their task was to name the color as rapidly as possible. The stimuli were either color patches or color names; the four colors used were red, green, yellow, and blue. There were three conditions: (a) colored patches, (b) congruent words in which color names (red, green, yellow, or blue) were displayed in their own color, and (c) incongruent words in which color names were displayed in a different color (e.g., the word red in blue font). In all cases the participant named the color of the patch or the word's font by saying it aloud; reaction time (RT) was measured by voice key. The stimuli were blocked by condition, with 24 stimuli appearing in each of two runs of the three conditions (for a total of 144 stimuli divided over six runs). The conditions were randomized separately for each participant, and the order of stimuli within each block of 24 stimuli was also randomized. Stimuli remained on the screen until the participant vocalized a response; the next stimulus was presented by the experimenter approximately 1 s after each response. The average RT for each condition was calculated from correct responses in both conditions, yielding measures for color patch and congruent and incongruent words. In the present study, only two of these measures were used—incongruent RT minus patch RT—to yield a measure of the Stroop Effect.

Star counting. Participants were presented with a series of eight cards, each containing a starting number (e.g., 61) followed

by a series of stars and interspersed plus and minus signs. Starting from the original number, the task was to add or subtract 1 for each star as they progressed through the series, depending on the latest symbol (+ or −) encountered. Participants counted out loud (61, 62, 63, . . . etc.) as well as pointing with their index finger as they progressed through the sequence. More difficult variations included switching the sign of the symbols (i.e., treating + as − and − as +), adding by 2 but subtracting by 1, and both switching symbols and adding by 2 while subtracting by 1. Each of the four variations (including counting normally) was tested on two cards, for a total of eight cards. The order of conditions was randomized separately for each participant, and the time to complete each card was measured. The score on each card was calculated by assigning a point for each number stated correctly after each switch point (+ or − sign). However, if participants made a mistake while counting, they could still receive points after subsequent switches if they were correct with respect to their current total. For example, if a participant incorrectly said 82 instead of 83 after a switch point, they would receive no point for that switch but would still receive a point for saying 81 after encountering the next −1 symbol. Over all eight cards, the total possible score was 56. To take solution times into account, each participant's total score was divided by his or her average time per card in seconds. This procedure yielded final Star Counting scores ranging between 0.29 and 2.01 (see *Results*), whereby larger scores indicated better performance.

Consonant updating. Participants were visually presented with a series of consonants at a 2-s rate; their score was based on the correct recall of the last four letters at the end of the series (Friedman et al., 2008; Morris & Jones, 1990). At the beginning of the series, the participant stated the first letter, then the first two, first three, and first four; thereafter they stated the most recent four letters and finally recalled the last four on presentation of a recall screen. One point was awarded for each correct letter, regardless of response order. Each series contained five, seven, nine, or 11 letters; there were two trials at each length for a total of eight trials and a maximum score of 32. List length was randomized, and participants were not informed of the length before presentation.

Alpha span. Participants were presented with a short series of common concrete one-syllable nouns at a 1-s rate and asked to repeat them back in correct alphabetical order (Craik et al., 2018). Two trials were given at each list length, starting at a list length of two, and participants progressed to the next length if they were correct on at least one of the two trials. List length of the current trial was announced before each presentation. One point was awarded for each word recalled in its appropriate alphabetic sequence, so the maximum score for each trial was simply the sequence length for that trial, and the maximum possible score for the whole test (alpha score) was 70. To give credit for responses that were partially correct, participants were awarded one point for each word recalled as part of a correctly ordered pair; additionally one point was awarded for the single first word (if recalled first) and also the single last word (if recalled last). As an illustration for the correct sequence, cloud, dog, king, oak, rock, street, wine, the response cloud, king, oak, dog, street, wine would be awarded five points (for cloud, king, oak, street, and wine); the response king, dog, cloud, oak, rock, wine would receive three points (for oak, rock, and wine); the response cloud, king, rock, wine would receive two points (for cloud and wine). Alpha score was the total number of points received, including those from partially correct sequences up to and including the list length at which the participant was incorrect on both trials.

Tests of retrieval efficacy. The three tests involved different forms of memory—free recall of words, verbal paired associates, and general knowledge of facts—but shared the characteristic that the measure was retrieval expressed as a proportion of what the participant knew as indexed by successful recognition memory.

Episodic retrieval. On each of three trials, participants were shown a series of 12 common concrete two-syllable nouns, presented on the computer screen for 3 s each; 12 different words were presented on each trial. Participants were then given 45 s to recall (orally) as many words as possible in any order (free recall). Following the recall phase, participants were shown a list of 28 words (again all common, concrete two-syllable nouns) including the original 12 words plus 16 distractors, mixed randomly. The test was yes/no self-paced recognition. The sequence of three recall/recognition trials was preceded by a practice trial. In both sections of the test, the presentation of words was randomly ordered individually for each participant, and the order of the three trials was also randomized. To eliminate recency effects and ensure that retrieval was from secondary memory, only the first eight words from the original presentation list were scored. On each trial the number of correctly recognized words (of these eight) was first obtained; the participant's recalled words were then examined to determine the number of correctly recognized words that were also recalled. The very few false-positive errors were ignored in this calculation. The total possible score was thus eight/eight for each of three lists, and 24/24 for the whole test. The episodic retrieval (ER) score was the proportion of correctly recognized words previously recalled, for example, $16/20 = 0.80$.

Verbal associates. On each of three trials, participants were shown a sequence of eight unrelated word pairs (common concrete two-syllable nouns) for 4 s per pair; different word pairs were presented on each trial. Immediately following presentation, participants were shown the first words from each pair (in a different random order from presentation order) and asked to recall the associated second words. Following this self-paced recall phase, participants were given a sheet of paper showing two eight-word lists, the first containing only the first words and the second containing only the second words from each pair; word order was again randomized with respect to the original order. Participants were asked to recognize the original pairs by matching the corresponding words. Following a practice trial, participants were given three trials (a different list on each trial), so the maximum scores for recall and recognition were 24 in both cases. As with the episodic retrieval test, the correctly paired associates in the second (recognition) phase were first obtained, and the verbal associates (VA) score was again the proportion of correctly recognized items that were also recalled.

General knowledge test. In the first half of the test, participants were asked 29 questions that were a mixture of general knowledge facts and vocabulary definitions. Examples are: What was the name of the World War II British Prime Minister? What is an infant whale called? What is the name of the tube connecting the mouth to the stomach? The questions were asked by the experimenter, and the participant responded orally. Following this unpaced recall phase, participants were given a sheet with the same questions in multiple-choice format (four possible answers for each question) and were asked to circle an answer in all cases, even

if that required guessing. For the three example questions quoted above, the respective alternatives were Lloyd George, Macmillan, Churchill, or Attlee, elver, calf, dolphin, or pup, and esophagus, sphincter, pylorus, or duodenum. Following the model of the previous two retrieval tests, the general knowledge (GK) score was the proportion of correct choices in the multiple-choice phase that the participant answered correctly in the first phase. In this case the order of questions in both phases was the same for all participants.

Overall procedure. Each participant performed the four EF tasks and three retrieval tasks in a different random order involving all seven tests. Testing took approximately 1½ hr.

Results

The means and standard deviations of the various measures are shown in Table 1. The table shows both raw scores and derived scores for the cases in which the main experimental measure was derived from two raw scores. In general the means and distributions of scores portray a reasonable amount of variability in a group of 60-year-old adults. There are ceiling effects in some raw scores; for example in Star Counting in which the mean is 51.9 of a possible 56. The test was originally designed for children and is therefore easy for adults. However, the time to complete the trials varied greatly, so the derived measure used to indicate ability (total score divided by time per trial) has a satisfactory degree of variation. Ceiling effects are also clearly present in Consonant Updating with a mean of 26.6 of 32, and in Episodic Recognition with a mean of 22.4 of 24. In the latter case, the derived measure used in further analyses (ER) is satisfactory, however, given the large variation in episodic recall scores. Cronbach's alpha test (Cronbach, 1951) was calculated for the GK scale to assess its internal consistency, and this procedure yielded a value of $\alpha = .72$, which is regarded as acceptable (Tavakol & Dennick, 2011). The calculation was based on 28 items because one question was both recognized and recalled by all 50 participants.

The next steps were first to see how the four potential components of EF (Star Counting, Consonant Updating, Alpha Span, and the Stroop Effect) interrelated and second to see how the EF components correlated with the derived retrieval variables ER, VA, and GK. The correlations among the four EF variables were all significant at the $p < .05$ level or better (values of r ranged from 0.31 to 0.55), confirming the assumption that they measure one common latent variable. With regard to correlations between the EF variables and the retrieval variables, ER correlated significantly with all four EF variables (values of r were between 0.45 and 0.53); VA correlated reliably with Stroop, $r = -0.33$, $p < .05$, and with Alpha Span, $r = .38$, $p < .01$; GK correlated reliably only with Star Counting, $r = .32$, $p < .05$. Finally, with regard to relations among the retrieval measures, ER and VA correlated significantly, $r = .48$, $p < .01$, but neither measure related to GK. It seems from this pattern of correlations that ER and VA tap similar abilities and that both related to EF variables to some degree. Contrary to prediction, GK does not appear to form part of this cluster, however.

Table 1
Means and Standard Deviations for Study 1

Executive function tests		Retrieval tests	
	M (*SD*)		*M* (*SD*)
Stroop test		Episodic retrieval	
Color patches (ms)	623 (83)	Recall	12.4 (4.6)
Incongruent (ms)	892 (156)	Recognition	22.4 (2.1)
Stroop effect (ms)	269 (102)	Measure (ER)	.54 (.19)
Star counting		Verbal associates	
Score	51.9 (4.1)	Recall	11.4 (6.7)
Time (s, per trial)	48.0 (16.6)	Recognition	16.4 (6.4)
Measure	1.18 (.34)	Measure (VA)	.60 (.23)
Consonant updating		General knowledge	
Measure	26.6 (4.0)	Recall	16.8 (4.1)
Alpha span		Recognition	23.1 (2.5)
Measure	27.6 (10.2)	Measure (GK)	.72 (.13)

The EF measures were then converted into z-scores to give equal weighting to each component and then combined for each participant to form a composite measure. We then assessed the degree to which the composite EF measure predicted retrieval ability in ER, VA, and GK; the results are shown in the left-hand panel of Table 2. The interim conclusion that GK is independent of the ER, VA, and EF variables is given further weight by the correlations between the EF composite measure and the three retrieval variables shown in Table 2. The composite measure related strongly to ER ($r = .64$), less strongly to VA ($r = .39$), and nonsignificantly to GK ($r = .25$). Steiger's z test transformations revealed that the EF-ER correlation was statistically greater than the EF-VA correlation ($z = 2.11$, $p < .04$), and the EF-ER correlation was significantly greater than the EF-GK correlation ($z = 2.66$, $p < .01$). However, the ER-VA correlation did not differ significantly from the ER-GK correlation ($z = 0.77$, $p = .44$).

We also carried out a principal components analysis (PCA) to further assess relations among all seven variables (four EF variables plus three retrieval variables). Table 3 shows the loadings on a two-component solution after oblique rotation (oblimin with Kaiser normalization). Only two factors had eigenvalues greater than 1.0 (eigenvalues were 3.18 and 1.08 for Components 1 and 2, respectively), so only the first two components were considered. Table 3 shows that all four EF variables loaded on the first component with values greater than 0.60; ER and VA also loaded on this component with loadings of 0.79 and 0.70, respectively. General knowledge retrieval ability did not load on this first component, however (loading = 0.03), but it was the only variable that did load substantially on the second component, with a loading of 0.86. The first component accounted for 45.5% of the overall variance, with the second component contributing a further 15.5%. The PCA thus confirms findings from the previous correlation analyses, pointing to the conclusion that the four EF variables cluster together and form a coherent component with retrieval ability from verbal ER and from VA. However, retrieval ability from factual semantic memory (GK) apparently reflects different abilities.

Discussion

The purpose of this study was to develop a composite measure of executive functioning in older adults and then assess the degree to which this composite measure predicted individual differences in retrieval abilities. Specifically, we explored the possibility that retrieval of general knowledge facts known to the participant would relate to retrieval of episodic verbal items and to verbal

Table 2
Product-Moment Correlations Among the Composite Measure of EF and Measures of Retrieval in Study 1 and Study 2

Variables	Study 1				Study 2			
	EF	ER	VA	GK	EF	ER	VA	GK
EF composite	1.0	**.64**	**.39**	.25	1.0	**.44**	**.40**	**.34**
Episodic retrieval		1.0	**.48**	.24		1.0	**.52**	**.28**
Verbal associates			1.0	.07			1.0	**.39**
General knowledge				1.0				1.0

Note. EF = executive function; ER = Episodic retrieval; VA = Verbal associates; GK = General knowledge. $N = 50$ in all cases. Correlations significant at the $p = .01$ level (two tailed) are printed in bold type.

associates and also that this common retrieval ability would correlate with the composite measure of executive function. Table 3 provides good evidence that the four postulated executive function components tap one common factor and justify their combination into a composite EF measure (see also Friedman et al., 2008; Miyake et al., 2000) despite the point that the individual tests involve different materials (numbers, letters, words, color patches) and different operations (counting, updating, mental rearrangement and inhibition).

The hypothesis that EF abilities would predict performance equivalently on retrieval tests of episodic items, episodic associative information, and general knowledge was clearly not upheld, however. Table 2 shows that the EF composite measure predicted ER well ($r = .64$), VA to a lesser degree ($r = .39$), and GK at a nonsignificant level ($r = .25$). This pattern is reinforced by the results of the PCA (see Table 3), showing that ER and VA loaded on the same factor as the four EF variables but with GK loading on a separate component from the other six variables. The clear independence of the GK measure was unexpected. One interpretation is that episodic and semantic memory systems are indeed independent, using different retrieval procedures. Although agreeing that the episodic/semantic distinction is useful and necessary, we have argued previously against a strong separation of the two systems and in favor of continuity between the two modes of representation (e.g., Craik, 2007).

One possible account of the present results is that whereas retrieval of recently presented unrelated verbal items or verbal associates requires processes that involve domain-general executive functions, the retrieval of known names or facts relies on different processes. For example, retrieval of known information that is already highly organized and structured may utilize different processes from those involved in the retrieval of recently presented arbitrary and unstructured events. A simpler account, however, suggests that measures of executive function do not all tap directly into the same abstract construct but rather assess somewhat different processes and abilities. In the present case, our choice of EF measures included two tests (Alpha Span and Consonant Updating) that clearly involve episodic recall of verbal material, so it is possible that the relatively strong correlations between EF on the one hand and ER and VA on the other reflect this fact rather than relations to a more abstract form of executive control. We therefore conducted a second study using a different selection of EF tests and an expanded set of semantic memory tests to provide a more reliable measure of fact and name retrieval.

Study 2

For the EF variables in the second study, we dropped Alpha Span and Consonant Updating but added the Trail Making Test (Reitan, 1958), a nonverbal test used widely in neuropsychological assessment that relates strongly to individual differences in speed and fluid cognitive abilities (Salthouse, 2011). Thus, EF was assessed by three tests: Star Counting, Stroop, and Trails. The ER and VA tests were unchanged from Experiment 1; we also retained the verbal general knowledge test (now referred to as GK.verbal) but added further items to take it from 29 to 40 questions. We also added a separate pictorial knowledge test (GK.photo) described below. Finally, we included a version of the word retrieval task first used by Freedman and Loftus (1971), also described below. Thus, the ability to retrieve factual information and words from semantic memory was now assessed by three tests: GK.verbal, GK.photo, and Word Retrieval.

Method

Participants. The participants were 50 adults (12 males, 38 females) aged between 60 and 70 years (mean age = 66.6 years). On average they had received 16.8 ($SD = 3.16$) years of formal education. They were in good mental and physical health, lived independently in the community, and traveled to our laboratory to participate in the study. They were volunteers who formed part of a pool of research participants recruited and organized by the Rotman Research Institute of Baycrest in Toronto. They were paid $12 CAD per hour plus transportation costs to compensate them for their time and participation. The study received ethical approval from the Baycrest Research Ethics Board.

Materials and procedure.

Tests of executive function. All participants performed three tasks to measure EF: the Stroop Test, Star Counting, and the Trail Making Test, Parts A and B (Reitan, 1958). The Stroop and Star Counting tests were those described in Study 1.

Trail making test. We administered the Reitan (1958) version of the Trail Making Test, consisting of Part A (connecting the numbers 1–25 in sequence) and Part B (connecting alternating numbers and letters, from 1 to L). In our version participants vocalized the numbers and letters as they proceeded through the sequence. Errors were corrected, so the total completion time

Table 3
Principal Component Analysis of Seven Variables in Study 1

Variables	Component	
	1	2
Episodic retrieval	.793	.084
Stroop effect	−.756	.303
Verbal associates	.697	−.273
Alpha score	.693	.276
Star counting	.628	.385
Consonant updating	.620	.301
General knowledge	.027	.860

Note. All variables are derived from the original raw scores except Alpha score and Consonant Updating (raw scores). The oblique rotation method used was oblimin with Kaiser normalization. Correlation between the two factors was $r = .16$. Rotation converged in 18 iterations.

included times to correct errors and continue from the correct symbol in the sequence. The score calculated was time to complete Trails B minus the time to complete Trails A. We thus assumed that Trails A measured processing speed and that the extra time to complete Trails B reflected the EF abilities of working memory and task switching (Lamberty, Putnam, Chatel, Bieliauskas, & Adams, 1994).

Tests of retrieval efficacy. The ER and VA tests were repeated from Study 1. The general knowledge test (now called GK.verbal) was also repeated but with three changes. First, the number of items was increased from 29 to 40 to provide a more satisfactory range of difficulty. Second, the recognition phase now included five alternatives per question rather than four in Study 1. Third, whereas participants in Study 1 had been asked to always choose an alternative in the recognition phase, participants in Study 2 were asked to choose an answer only if they were sure or fairly sure it was correct; in this way we hoped to achieve a more valid measure of what they actually knew. Additionally, retrieval from general knowledge (semantic memory) was now also represented by a pictorial form of the GK test and by a test of word retrieval (WR).

General knowledge test–photographic. This test (GK.photo) was similar in form to the GK.verbal test used in the present study except that the materials were pictorial. In the first half of the test, participants were shown 40 photographs of objects, places, or people and asked to provide the name or some fact about each photograph. In the second half, participants were shown the same photographs with their associated questions and five possible answers. Participants were asked to select the correct answer but only if they were sure or reasonably sure it was correct. The score was calculated by dividing the total number of correctly answered questions in the first half (counting only those items that were also correctly answered in the multiple-choice recognition test) by the total number of correctly answered questions in the second half. As with the GK.verbal task, we assumed that this score reflected the ability to spontaneously recall information that the participants did actually know.

Word retrieval (WR). On each of 40 trials, participants were shown a category name and the first letter of an exemplar from that category; for example, fruit—A, or profession—T. The task was to state the exemplar's name (e.g., apple or teacher) as rapidly as possible. Participants spoke their answers into a microphone, which triggered a voice key, giving a value of retrieval time for each trial. The category name plus first letter appeared for 5 s, and this was followed by a plus sign for 2 s and then the next category and letter. A trial was considered to be valid provided that the voice key had not been triggered prematurely by a cough or by the participant thinking aloud. Each valid trial was scored correct or incorrect depending on the participant's success in stating a correct word within 5 s, and the final score for the test was the number of correct responses divided by the total number of valid trials $\times$ 100.

Procedure. Each participant performed the three EF tasks and five retrieval tasks in a different random order involving all eight tests. Testing took approximately 2 hr.

Results and Discussion

The means and standard deviations are shown in Table 4. Comparing Tables 1 and 4 for the five tests that were common to the two studies (Star Counting, Stroop, ER, VA, and GK.verbal), the means and standard deviations are very similar, showing that the two samples of 50 older adults were broadly comparable in their abilities. Differences between the studies in the final measures for each test of EF or retrieval efficacy were assessed by t tests. The obtained values of t ranged from $t = 0.00$ in the case of Star Counting to $t = 1.59$ in the case of GK.verbal, $p > .05$ in all cases.

Correlations among the eight measures showed that the three EF measures (Star Counting, Stroop, and Trails) all correlated significantly with each other (values of r ranged from 0.39 to 0.67),

Table 4
Means and Standard Deviations for Study 2

Executive function tests		Retrieval tests	
Stroop test		Episodic retrieval	
Color patches (ms)	624 (81)	Recall	11.8 (4.3)
Incongruent (ms)	882 (181)	Recognition	22.4 (1.8)
Stroop effect (ms)	258 (133)	Measure (ER)	.52 (.18)
Star counting		Verbal associates	
Score	52.5 (4.1)	Recall	10.3 (7.0)
Time (s, per trial)	49.9 (16.5)	Recognition	15.0 (7.1)
Measure	1.18 (.42)	Measure (VA)	.56 (.25)
Trail making test		General knowledge verbal	
Part A	35.4 (17.4)	Recall	20.0 (8.0)
Part B	74.7 (33.2)	Recognition	27.2 (7.5)
B – A	39.3 (24.1)	GK.verbal	.67 (.18)
		General knowledge photo	
		Recall	16.4 (7.6)
		Recognition	25.7 (6.7)
		GK.photo	.60 (.18)
		Word retrieval	
		Mean (s)	2.53 (.69)
		Median (s)	2.17 (1.01)
		Measure (WR)	41.6 (13.7)

confirming the expectation that they measure a common latent variable of EF. The ER and VA verbal retrieval tasks also correlated significantly, $r = .52$, $p < .001$, and the three semantic memory retrieval tasks (GK.verbal, GK.photo, and Word Retrieval) all correlated strongly (average value of $r = .64$). This last result thus provides evidence for a latent variable measuring the efficiency of retrieval from semantic memory. The three EF variables correlated modestly but reliably with ER and VA for the most part (values of r are between 0.25 and 0.42), echoing the result from Study 1. However, the correlations between the EF measures and the semantic retrieval measures are strikingly different in the present case in that eight of the nine correlations are now significant.

As in Study 1, the measures were then transformed to z-scores and combined to form composite measures of EF (Star Counting, Stroop, and Trails) and of semantic retrieval efficacy (GK.verbal, GK.photo, and WR). The measures ER and VA remained as separate components. Table 2 (right-hand panel) shows the pattern of correlations among the four variables. Comparing these values of r with those found in Study1, the obvious differences are first that the correlation between the EF composite and ER is lower in Study 2 and second that the correlation between the EF composite and the GK composite is now a statistically significant $r = .34$, as opposed to a nonsignificant $r = .25$ in Study 1. The relation of the EF composite measure to the other three measures maintains the same relative order (as in Study 1) of ER > VA > GK, but in the present study, Steiger's z test transformations revealed no significant differences between the three correlations (all p values > .50).

Following our procedure in Study 1, we also carried out a PCA on all eight variables, and Table 5 shows the loadings on a two-component solution after oblique rotation (oblimin with Kaiser normalization). Again two factors had eigenvalues greater than 1.0 (eigenvalues were 3.98 and 1.08 for Components 1 and 2, respectively). The table shows that the three semantic retrieval variables dominate the first component with loadings greater than 0.80; this component may therefore be identified as a semantic retrieval factor, predicted to some degree by Star Counting (loading = 0.53). The second component contains three variables with loadings greater than 0.60—Stroop, ER, and VA—and this cluster is predicted to some degree by Trails (0.58) and Star Counting (0.37). Trails and Stroop have a different sign from the other variables in most cases because better performance is indexed by smaller scores for these two measures. The cluster represented by Component 2 may be tentatively identified as episodic retrieval and EF. Thus, the PCA provides some evidence for separate semantic and episodic retrieval factors, with the semantic factor predicted by one EF measure (Star Counting) and the episodic factor predicted principally by the Stroop Effect and to a lesser degree by Trails.

Table 5
Principal Component Analysis of Eight Variables in Study 2

Variable	Component 1	Component 2
Word retrieval	.888	.093
General knowledge, verbal	.862	−.029
General knowledge, photo	.831	−.046
Star Counting	.531	−.367
Stroop Effect	.143	.890
Episodic retrieval	−.027	−.763
Verbal associates	.151	−.607
Trails	−.210	.582

Note. All variables are derived from the original raw scores. The oblique rotation method used was oblimin with Kaiser normalization. Correlation between the two factors was $r = -.50$. Rotation converged in seven iterations.

The main conclusions to be drawn from Study 2 are first that semantic memory retrieval does seem to constitute a reliable ability that embraces both retrieval of facts (GK.verbal and GK .photo) and retrieval of single words (WR). Furthermore, the PCA results suggest that semantic memory retrieval and episodic memory retrieval are somewhat distinct, although both are predicted by components of EF. The main difference from Study 1 is that eight of the nine correlations between the three semantic retrieval variables and the three EF variables are significant, showing that EF does play a role in retrieval from semantic memory. Speculative reasons for the discrepancy between the studies are offered in the *General Discussion*.

General Discussion

The two experiments reported in this article explored the role of executive functions in retrieval from episodic and semantic memory in older adults. In overview, Study 2 showed that EF is involved in retrieval from both episodic and semantic memory, but a consideration of the two studies revealed an unexpected complexity. The PCAs (Tables 3 and 5) both showed that episodic and semantic memory are somewhat distinct but that the two types of memory are related to different components of EF.

We proposed a measure of retrieval efficacy in which participants first attempted to recall specified information in an unaided fashion and then, as a measure of participants' knowledge of the material, tried to recognize that information from a series of options. Retrieval efficacy was calculated as the proportion of known information recognized that was successfully recalled in the first phase. Such measures were calculated for the recall of word lists (ER), for paired-associates (VA), and for factual knowledge in which the material presented was either verbal (GK.verbal) or pictorial (GK.photo). The dependence of retrieval efficacy on individual differences in executive function abilities was then assessed through correlations with a variety of EF measures.

Study 1 showed that a composite measure of EF, comprising Star Counting, Consonant Updating, Alpha Span, and Stroop tests, related strongly to ER, less strongly to VA, and nonsignificantly to GK.verbal. It was tentatively concluded that individual differences in executive functions predicted episodic retrieval abilities, especially in tests such as free recall that require large amounts of self-initiated processing, but played a comparatively small role in retrieval of factual knowledge. The notion that the processes associated with retrieval from episodic memory are substantially different from those associated with semantic memory is an interesting one, especially if it could be shown that executive processes play a major role in the former but not the latter case. However, an alternative possibility is that the specific EF measures used in Study 1 were more congruent with the specific processes involved in the retrieval of verbal information (as reflected by the ER and VA measures) than factual information (represented by the GK

test). This point was explored by conducting a second experiment with somewhat different measures. In this second study, EF was assessed by three tests—Star Counting, Stroop, and Trails, none of which involve verbal retrieval; the ER and VA tasks were repeated from Study 1, and retrieval efficacy from semantic memory was measured by an expanded version of the verbal general knowledge test from Study 1 (GK.verbal) plus a new pictorial version (GK .photo) and a word retrieval task (WR). The results showed that the relations of the new EF variables to measures of retrieval efficacy were considerably different from the relations found in the first experiment. In particular, the composite measure of general knowledge retrieval (semantic memory) now did correlate significantly with the EF composite and also with ER and VA. The different pattern of results between the studies is also marked by the different composition of the PCAs shown in Tables 3 and 5.

The major differences in the designs of Studies 1 and 2 are first the different composition of the EF factor—the replacement of the verbal memory tests Alpha Span and Consonant Updating with the largely nonverbal Trails test—and second, the more adequate representation of a general knowledge factor by three tests. It therefore seems that changes in the composition of an EF factor on the one hand and changes in the material to be retrieved on the other hand result in markedly different answers to the question, "Do individual differences in retrieval efficacy depend on individual differences in executive functions?" On reflection, our initial question may have been too global and altogether too simple; rather than regard EF as one thing and also retrieval processes from episodic and semantic memory as separable entities, it may be necessary to look at components of these general factors and draw rather cautious conclusions with respect to the larger-scale theoretical constructs. One case in point is that the relation between EF and ER efficacy appeared to depend on the tests chosen to represent EF. The four EF tests in Study 1 and the three somewhat different tests in Study 2 both clustered well to provide latent variables of executive functioning, but the different relations between these two composite measures and ER (0.64 in Study 1 and 0.44 in Study 2) suggest that inclusion of two verbal retrieval tasks in the EF composite in Study 1 boosted the relation to ER in the first study. Speculatively, the greater demand on self-initiated activities in free recall may make ER more dependent on similar components in the EF composite.

A study by Hedden et al. (2005) explored the relations between verbal processing capacity and the episodic tasks of free recall, cued recall, and recognition memory. The results from a structural equation modeling analysis showed that the processing construct related to free recall, cued recall, and recognition with path coefficients of 0.85, 0.73, and 0.63, respectively, in a group of older adults ($N = 173$; aged 55–92 years), a result in line with the notion that processing ability relates most strongly to performance on tasks requiring most self-initiation and least environmental support (Craik, 1983, 1986). The authors also found that a construct of verbal knowledge was essentially unrelated to free recall in older adults (path coefficient = 0.04). Thus, both the Hedden experiment (Hedden et al., 2005) and the present two studies are in agreement that unaided free recall performance benefits from high EF abilities, and the present results at least suggest that this sensitivity to executive functions is especially beneficial when the EF components contain processes similar to those needed for effective free recall.

Similarly, the relation between free recall and EF (or processing capacity) may depend substantially on how these constructs are measured. The larger point illustrated by the two present studies is that one EF composite factor may be considerably different from another EF composite, depending on their components. Despite the finding that the individual EF tasks correlated in both studies, perhaps there is no one Platonic ideal latent measure of EF at the pragmatic level of cognitive tasks but only composite measures with a family resemblance to a higher abstract construct that exists solely in the theoretical domain (see also Miyake et al., 2000; Stuss, 2011). In the same vein, relations between EF and other composite measures will presumably depend on the specific components of both composites.

The PCAs shown in Tables 3 and 5 suggest that retrieval from episodic and semantic memory involve different processes to some extent at least. Does this speculative conclusion concede the point that retrieval processes in episodic and semantic memory are essentially different and that the two types of memory therefore represent independent systems as suggested by Tulving (1972)? An alternative proposition (Craik, 2007) is that episodic and semantic memory form a continuum of representations, running from the highly contextually dependent episodic representations of recently experienced events, through a hierarchy of increasingly abstract representations to entirely context-free representations of factual knowledge about aspects of experienced events—that is, semantic memory. One possible resolution of this debate is the suggestion that memory is structured as levels of representation running continuously from largely episodic to largely semantic but that effective retrieval processes shift in nature, depending on the type of information sought. For example, ER may depend more on EF processes and semantic retrieval may depend more on relevant knowledge (see Bouazzaoui et al., 2013, 2014; Hedden et al., 2005 for further thoughts and speculations).

Findings from the neuroimaging literature may eventually serve to clarify the relations between frontally based executive processes and retrieval from episodic and semantic memory involving different types of material. In a study using positron emission tomography imaging, Tulving, Kapur, Craik, Moscovitch, and Houle (1994) showed that whereas episodic memory retrieval was associated with activations in right prefrontal areas, retrieval of lexical-semantic information was associated with activation in left prefrontal regions. This general pattern has also been shown by more recent work using functional magnetic resonance imaging methods (e.g., Düzel, Habib, Guderian, & Heinze, 2004; Wiggs, Weisberg, & Martin, 1999). A meta-analysis by Costafreda, Fu, Lee, Everitt, Brammer, and David (2006) confirmed the observation that activation of the left inferior frontal gyrus is associated with retrieval from semantic memory but that phonologic verbal fluency tasks activated areas that were significantly more dorsal than areas associated with semantic verbal fluency tasks. Thus when the specific retrieval task within semantic memory is changed, the area within the frontal lobes (and speculatively the type of executive control processing associated with that area) is also changed. In the same vein, Burianova and Grady (2007) compared the neural activation patterns associated with autobiographical, episodic, and semantic memory and found evidence for neural patterns common to all types of memory retrieval but also regions that responded uniquely to each type of retrieval. This type of work may thus have the potential to clarify differences between different aspects of

memory retrieval and also to relate different types of executive processing to these retrieval differences.

In conclusion, what have the present studies shown about retrieval processes in older adults? The initial question was whether retrieval of information from episodic and semantic memory in older adults both relied on the involvement of executive functions. Specifically, do individual differences in retrieval from these two types of memory both depend on individual differences in EF abilities? Consideration of the results from two studies led us to conclude that they do, but with the unexpected qualification that different aspects of EF predict retrieval performance differentially in the two types of memory. A second qualification is that the processes involved in retrieval from episodic and semantic memory may not themselves be constant from one situation to another but depend on the nature of the material to be retrieved. Age-related impairments in retrieval from episodic and semantic memory are to some extent attributable to inefficiencies in executive functions, but the nature of these executive functions is somewhat different between the two types of retrieval.

References

Bouazzaoui, B., Angel, L., Fay, S., Taconnat, L., Froger, C., & Isingrini, M. (2014). Does the greater involvement of executive control in memory with age act as a compensatory mechanism? *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 68,* 59–66. http://dx.doi.org/10.1037/cep0000005

Bouazzaoui, B., Fay, S., Taconnat, L., Angel, L., Vanneste, S., & Isingrini, M. (2013). Differential involvement of knowledge representation and executive control in episodic memory performance in young and older adults. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 67,* 100–107. http://dx.doi.org/10.1037/a0028517

Buckner, R. L. (2004). Memory and executive function in aging and AD: Multiple factors that cause decline and reserve factors that compensate. *Neuron, 44,* 195–208. http://dx.doi.org/10.1016/j.neuron.2004.09.006

Burianova, H., & Grady, C. L. (2007). Common and unique neural activations in autobiographical, episodic, and semantic retrieval. *Journal of Cognitive Neuroscience, 19,* 1520–1534. http://dx.doi.org/10.1162/jocn.2007.19.9.1520

Burke, D. M., & Light, L. L. (1981). Memory and aging: The role of retrieval processes. *Psychological Bulletin, 90,* 513–514. http://dx.doi.org/10.1037/0033-2909.90.3.513

Cabeza, R., & Dennis, N. A. (2013). Frontal lobes and aging: Deterioration and compensation. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (2nd ed., pp. 628–652). New York, NY: Oxford University Press.

Costafreda, S. G., Fu, C. H., Lee, L., Everitt, B., Brammer, M. J., & David, A. S. (2006). A systematic review and quantitative appraisal of fMRI studies of verbal fluency: Role of the left inferior frontal gyrus. *Human Brain Mapping, 27,* 799–810. http://dx.doi.org/10.1002/hbm.20221

Craik, F. I. M. (1983). On the transfer of information from temporary to permanent memory. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences, 302,* 341–359. http://dx.doi.org/10.1098/rstb.1983.0059

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

Craik, F. I. M. (2007). Encoding: A cognitive perspective. In H. L. Roediger, III, Y. Dudai, & S. M. Fitzpatrick (Eds.), *Science of memory: Concepts* (pp. 129–135). New York, NY: Oxford University Press.

Craik, F. I. M., & Bialystok, E. (2008). Lifespan cognitive development: The roles of representation and control. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (3rd ed., pp. 557–601). New York, NY: Psychology Press.

Craik, F. I. M., Bialystok, E., Gillingham, S., & Stuss, D. T. (2018). Alpha Span: A measure of working memory. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 72,* 141–152. http://dx.doi.org/10.1037/cep0000143

Craik, F. I. M., & Grady, C. L. (2002). Aging, memory, and frontal lobe functioning. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 528–540). New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780195134971.003.0031

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

Craik, F. I. M., & Rose, N. S. (2012). Memory encoding and aging: A neurocognitive perspective. *Neuroscience and Biobehavioral Reviews, 36,* 1729–1739. http://dx.doi.org/10.1016/j.neubiorev.2011.11.007

Cronbach, L. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16,* 297–334. http://dx.doi.org/10.1007/BF02310555

Das-Smaal, E. A., de Jong, P. F., & Koopmans, J. R. (1993). Working memory, attentional regulation and the Star Counting Test. *Personality and Individual Differences, 14,* 815–824. http://dx.doi.org/10.1016/0191-8869(93)90094-J

Düzel, E., Habib, R., Guderian, S., & Heinze, H. J. (2004). Four types of novelty-familiarity responses in associative recognition memory of humans. *European Journal of Neuroscience, 19,* 1408–1416. http://dx.doi.org/10.1111/j.1460-9568.2004.03253.x

Freedman, J. L., & Loftus, E. F. (1971). Retrieval of words from long-term memory. *Journal of Verbal Learning and Verbal Behavior, 10,* 107–115. http://dx.doi.org/10.1016/S0022-5371(71)80001-4

Friedman, N. P., Miyake, A., Young, S. E., Defries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology: General, 137,* 201–225. http://dx.doi.org/10.1037/0096-3445.137.2.201

Fuster, J. M. (2002). Physiology of executive functions: The perception-action cycle. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 96–108). New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780195134971.003.0006

Hedden, T., Lautenschlager, G., & Park, D. C. (2005). Contributions of processing ability and knowledge to verbal memory tasks across the adult life-span. *Quarterly Journal of Experimental Psychology, 58,* 169–190. http://dx.doi.org/10.1080/02724980443000179

Lamberty, G. J., Putnam, S. H., Chatel, D. M., Bieliauskas, L. A., & Adams, K. M. (1994). Derived Trail Making test indices: A preliminary report. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology, 7,* 230–234.

McDaniel, M. A., Einstein, G. O., & Jacoby, L. L. (2008). New considerations in aging and memory. In F. I. M. Craik & T. A. Salthouse (Eds.), *Handbook of aging and cognition* (3rd ed., pp. 251–310). New York, NY: Taylor and Francis.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41,* 49–100. http://dx.doi.org/10.1006/cogp.1999.0734

Morris, N., & Jones, D. M. (1990). Memory updating in working memory: The role of the central executive. *British Journal of Psychology, 81,* 111–121. http://dx.doi.org/10.1111/j.2044-8295.1990.tb02349.x

Moscovitch, M., & Winocur, G. (1992). The neuropsychology of memory and aging. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 315–372). Hillsdale, NJ: Erlbaum.

Oberauer, K., Süß, H.-M., Schulze, R., Wilhelm, O., & Wittmann, W. W. (2000). Working memory capacity—Facets of a cognitive ability con-

struct. *Personality and Individual Differences, 29,* 1017–1045. http://dx.doi.org/10.1016/S0191-8869(99)00251-2

Park, D. C., & Reuter-Lorenz, P. (2009). The adaptive brain: Aging and neurocognitive scaffolding. *Annual Review of Psychology, 60,* 173–196. http://dx.doi.org/10.1146/annurev.psych.59.103006.093656

Raz, N. (2000). Aging of the brain and its impact on cognitive performance: Integration of structural and functional findings. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 1–90). Mahwah, NJ: Erlbaum.

Reitan, R. M. (1958). Validity of the Trail Making Test as an indicator of organic brain damage. *Perceptual and Motor Skills, 8,* 271–276. http://dx.doi.org/10.2466/pms.1958.8.3.271

Salthouse, T. A. (2011). What cognitive abilities are involved in trail-making performance? *Intelligence, 39,* 222–232. http://dx.doi.org/10.1016/j.intell.2011.03.001

Shimamura, A. P. (1995). Memory and frontal lobe function. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 803–813). Cambridge, MA: MIT Press.

Shimamura, A. P. (2002). Memory retrieval and executive control processes. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 210–220). New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780195134971.003.0013

Shing, Y. L., Werkle-Bergner, M., Brehmer, Y., Müller, V., Li, S.-C., & Lindenberger, U. (2010). Episodic memory across the lifespan: The contributions of associative and strategic components. *Neuroscience and Biobehavioral Reviews, 34,* 1080–1091. http://dx.doi.org/10.1016/j.neubiorev.2009.11.002

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643–662. http://dx.doi.org/10.1037/h0054651

Stuss, D. T. (2011). Functions of the frontal lobes: Relation to executive functions. *Journal of the International Neuropsychological Society, 17,* 759–765. http://dx.doi.org/10.1017/S1355617711000695

Stuss, D. T., & Benson, D. F. (1986). *The frontal lobes*. New York, NY: Raven Press.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2,* 53–55. http://dx.doi.org/10.5116/ijme.4dfb.8dfd

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 381–403). New York, NY: Academic Press Inc.

Tulving, E., Kapur, S., Craik, F. I. M., Moscovitch, M., & Houle, S. (1994). Hemispheric encoding/retrieval asymmetry in episodic memory: Positron emission tomography findings. *Proceedings of the National Academy of Sciences of the United States of America, 91,* 2016–2020. http://dx.doi.org/10.1073/pnas.91.6.2016

Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science, 247,* 301–306. http://dx.doi.org/10.1126/science.2296719

Wiggs, C. L., Weisberg, J., & Martin, A. (1999). Neural correlates of semantic and episodic memory retrieval. *Neuropsychologia, 37,* 103–118. http://dx.doi.org/10.1016/S0028-3932(98)00044-X

Received September 26, 2017
Revision received October 3, 2018
Accepted October 8, 2018 ■