Psychology and Aging
2000, Vol. 15, No. 3, 527–541

0882-7974/00/$5.00 DOI: 10.1037//0882-7974.15.3.527

Aging and Reflective Processes of Working Memory: Binding and Test Load Deficits

Karen J. Mitchell, Marcia K. Johnson, Carol L. Raye, and Mara Mather
Princeton University

Mark D'Esposito
University of Pennsylvania Medical Center

It was hypothesized that age-related deficits in episodic memory for feature combinations (e.g., B. L. Chalfonte & M. K. Johnson, 1996) signal, in part, decrements in the efficacy of reflective component processes (e.g., M. K. Johnson, 1992) that support the short-term maintenance and manipulation of information during encoding (e.g., F. I. M. Craik, R. G. Morris, & M. L. Gick, 1990; T. A. Salthouse, 1990). Consistent with this, age-related binding deficits in a working memory task were found in 2 experiments. Evidence for an age-related test load deficit was also found: Older adults had greater difficulty than young adults when tested on 2 features rather than 1, even when binding was not required. Thus, disruption of source memory in older adults may involve deficits in both encoding processes (binding deficits) and monitoring processes (difficulty accessing multiple features, evaluating them, or both).

An *episodic memory* is a memory that includes details that can be used to attribute it to a source (e.g., the time, place, and modality of its acquisition). Older adults often experience substantial difficulty compared with young adults in remembering many types of event details (see Burke & Light, 1981; Kausler, 1994; Spencer & Raz, 1995, for reviews). For example, there is evidence of age-related memorial deficits for spatial information (e.g., Light & Zelinski, 1983), voice (e.g., Kausler & Puckett, 1981), modality (e.g., McIntyre & Craik, 1987), color (e.g., Park & Puglisi, 1985; but see Chalfonte & Johnson, 1996), temporal cues (e.g., Schacter, Kaszniak, Kihlstrom, & Valdiserri, 1991; but see Bayen & Murnane, 1996), person (e.g., Ferguson, Hashtroudi, & Johnson, 1992), and cognitive operations (e.g., Rabinowitz, 1989).

Although it is reasonably clear that there are age-related deficits in memory for source information, there has been little work to date attempting to separate age-related deficits in feature memory, per se, and deficits in memory for associations between the individual features of an event (i.e., a *feature binding* deficit). That is, older adults might remember as well as young adults which objects they saw and which locations had objects in them (i.e., they may have intact memory for individual features) but have more difficulty than young adults remembering which object was in which location (i.e., they may have a deficit for bound object and location information). Chalfonte and Johnson (1996) directly compared feature memory (colors, objects, and locations) and feature binding in young and older adults and found evidence for an age-related deficit in feature binding (i.e., remembering object + location and object + color combinations).[1] For example, even though older adults' ability to remember what colors they have seen and what objects they have seen is equivalent to that of young adults, older adults show a deficit in their ability to remember object + color combinations. In addition, Chalfonte and Johnson found evidence for an age-related long-term deficit in memory for location.

The failure to remember the associations between features of an event could result from processing deficits at acquisition (e.g., binding during encoding), during the retention interval (e.g., self-initiated rehearsal or reactivations; Johnson & Chalfonte, 1994), or at test (e.g., revival and decision), or some combination of these. Although their design could not distinguish between these alternatives, Chalfonte and Johnson (1996) proposed that older adults' binding deficit may reflect, in part, poor encoding arising from deficits in one or more of the reflective component processes that are critical for establishing associations between elements of experience. Chalfonte and Johnson's original hypothesis focused mainly on the reflective component process of *reactivation* (i.e., a proposed cognitive process by which elements that have left consciousness are made active again; e.g., Johnson, 1992; Johnson &

Karen J. Mitchell, Marcia K. Johnson, Carol L. Raye, and Mara Mather, Department of Psychology, Princeton University; Mark D'Esposito, Department of Neurology, University of Pennsylvania Medical Center.

Marcia K. Johnson and Carol L. Raye are now at the Department of Psychology, Yale University. Mara Mather is now at the Department of Psychology, University of California, Santa Cruz. Mark D'Esposito is now at the Department of Psychology, University of California, Berkeley.

This research was supported by National Institute of Aging Grants AG09253 and AG15793.

Correspondence concerning this article should be addressed to Karen J. Mitchell, who is now at the Department of Psychology, Yale University, Box 208205, New Haven, Connecticut 06520-8205. Electronic mail may be sent to karen.mitchell@yale.edu.

[1] For clarity, throughout this article we reserve the term *conjunction* to refer to the perceptual experience of features "going together" and use the term *combination* to refer to the memorial representation that results from features being bound together. Discussion of the nature of this bound representation (e.g., a new "gestalt" versus "pointers" to the various independent features) is beyond the scope of this article. Although we use the term *binding* in this article to refer specifically to the binding of simple features (e.g., object + location), it is assumed that the same processes are involved in binding multiple complex "units" (themselves already well bound) together.

Chalfonte, 1994; Johnson & Hirst, 1993); however, deficits in other reflective processes would also be expected to affect encoding (for instance, *shifting* attention between features, cumulative *rehearsal*; e.g., Johnson, 1992). The present experiments are based on two assumptions: that these reflective component processes compose, at least in part, what is usually referred to as *working memory* (e.g., Baddeley, 1986, 1992; Baddeley & Hitch, 1974; Norman & Shallice, 1986), and that with respect to long-term memory, these reflective processes constitute encoding processes. This latter assumption is consistent with many models that assume that rehearsal increases the probability of long-term memory (e.g., Atkinson & Shiffrin, 1968).

Consistent with the idea that aging is associated with loss of efficiency of the reflective component processes required to encode episodic events in memory, older adults have been shown to exhibit deficits on many working memory tasks, especially when coordination or manipulation of information is necessary (e.g., Babcock & Salthouse, 1990; Bayles & Kasniak, 1987; Craik, 1986; Dobbs & Rule, 1989; Foos, 1989; Salthouse, 1990). For example, there is generally a greater age-related decrement in performance on backward than forward memory span tasks (see Babcock & Salthouse, 1990, for a meta-analysis; Craik, 1977, for a review; but see, e.g., Grégoire & Van der Linden, 1997). Likewise, Gick and Craik found an age-related deficit on a task in which participants were to repeat a short random series of words back in alphabetical order, but not on a simple forward digit-span task (as cited in Craik, 1986; but see Belleville, Rouleau, & Caza, 1998). Although there is still considerable debate regarding the interpretation of such findings (see, e.g., Belleville et al., 1998, for a discussion), it seems likely that age-related decrements in the efficacy of some of the reflective component processes that support the short-term maintenance or manipulation of information (e.g., Craik, 1977, 1986; Craik, Morris, & Gick, 1990; Salthouse, 1988, 1990) contribute to the binding deficits that underlie older adults' long-term episodic memory decrements.

In addition to the behavioral evidence regarding older adults' decrements on working memory tasks, there is accumulating evidence from the neuropsychological, neurophysiological, and neuroimaging domains that points to specific age-related deficits in brain areas known to be involved in working memory and binding processes. For example, the frontal lobes have long been known to play a critical role in reflective, or more executive, cognitive processes (e.g., planning, monitoring; see, e.g., Jonides et al., 1996; Petrides, 1994a, for reviews), and recent neuroimaging evidence from young adults points to the frontal region (especially prefrontal cortex) as a primary locus of working memory processes (e.g., J. D. Cohen et al., 1997; D'Esposito et al., 1995; Goldman-Rakic, 1995; Smith, Jonides, & Koeppe, 1996). There is evidence suggesting that normal aging affects the frontal lobes; in fact, a recent review of the neuroanatomical, neurochemical, and metabolic indicators of aging conducted by Raz (2000) suggests that the prefrontal cortex may be more affected by aging than any other cortical region (see also, Raz, Gunning-Dixon, Head, Dupuis, & Acker, 1998). Likewise, the hippocampal–medial temporal region has been implicated as a region critical in binding processes (e.g., N. J. Cohen & Eichenbaum, 1993; Eichenbaum & Bunsey, 1995; Henke, Weber, Kneifel, Wieser, & Buck, 1999; Kroll, Knight, Metcalfe, Wolf, & Tulving, 1996; Moscovitch, 1994; Squire, 1995), and in a recent neuroimaging study conducted by Grady et al. (1995) older, but not young, adults exhibited a lack of hippocampal activity on a task that may rely on the binding of features into complex representations (i.e., encoding of faces; Kroll et al., 1996). Thus, there is both behavioral and neurophysiological evidence to suggest that aging may be associated with deficits in the encoding processes thought to underlie memorial binding.

To explore this issue, in Experiment 1 we examined memory for individual features (i.e., object and location information) and memory for combinations of features (i.e., object plus its location) in a working memory task. Participants received alternating blocks of trials on which they were to remember either object information only, location information only, or the combination of object and location information. They were always given a "yes–no" recognition test of the information that they were told to study on that trial (Chalfonte & Johnson, 1996). Experiment 2 used a similar procedure to examine factors at both encoding and test.

Experiment 1

Method

Participants

Twenty-four undergraduates (mean age = 19.5 years, SD = 0.93 years, range = 18 to 21 years) received credit toward fulfillment of a course requirement in exchange for their participation. Twenty-four healthy older adults (mean age = 74.1 years, SD = 2.43 years, range = 63 to 85 years) were recruited from the community to serve as paid participants. All participants reported having normal or corrected-to-normal vision and also reported being able to see all of the stimuli well. The behavioral data regarding memory for the individual features (discussed below) support these reports. Older adults were also screened to ensure freedom from primary degenerative brain disorders (e.g., Alzheimer's disease, Parkinson's disease) and other conditions or medications that reasonably may be expected to affect memory (e.g., stroke, brain trauma, psychotropic medications).

Although the older adults had completed more years of formal education (M = 15.7, SD = 3.2) than the young adults (M = 13.2, SD = 0.83), $t(46) = 3.66$, $p < .001$, they were not atypical of people their age in terms of general memory abilities. Within a couple weeks of participating in this working memory experiment, 21 of the 24 older adults took part in another study that included subtests of the Wechsler Memory Scale–Revised (Wechsler, 1987) and the California Verbal Learning Test (Delis, Kramer, Kaplan, & Ober, 1987). Table 1 contains their mean scores on these tests, together with published normative data for adults ages 65 to 75 years. Examination of this table shows that our older adults were within established norms.

Design

In addition to the between-subjects variable, age, there were two within-subjects variables. We were primarily interested in young and older adults' ability to remember combined object + location information as compared with their memory for individual features. To this end, type of to-be-remembered information was manipulated within subjects. We were secondarily interested in the extent to which similarity of the objects across trials might differentially affect the performance of young and older adults (e.g., Kane & Hasher, 1996). Thus, for generality, half of all trials used the same eight objects across all trials (i.e., repeated objects trials) and half of the trials used different objects on each trial (i.e., nonrepeated object trials). This resulted in a 2 (age: young, old) × 3 (condition: object only, location only, or combination) × 2 (item type: repeated vs. nonrepeated objects) mixed design. Preliminary analyses showed that there were no significant effects of whether the objects were novel or

Table 1
Older Adults' Mean Scores and Standard Deviations on Select Standard Memory Tests and Published Norms for Adults Ages 65 to 75 Years

Test	Maximum score	Experiment 1[a]		Experiment 2[b]		Test norm	
		M	*SD*	*M*	*SD*	*M*	*SD*
Backward Digit Span	12[c]	7.4	2.1	7.1	2.0	7.6	2.0
Long Delay Cued Recall	16[d]	12.4	3.0	10.5	2.6	10.3	3.7
Verbal Paired Associates I	24[c]	19.7	3.4	14.8	4.4	16.5	4.0

Note. Test norms are from Glisky, Polster, and Routhieaux (1995).
[a] $n = 21$. [b] $n = 12$. [c] From the Wechsler Memory Scale–Revised (Wechsler, 1987). [d] From the California Verbal Learning Test (Delis, Kramer, Kaplan, & Ober, 1987).

repeated across trials, nor did this factor enter into any significant interactions. This factor, therefore, is not discussed further; the data were collapsed across this factor for all analyses.

Materials and Procedure

Stimuli were presented via computer, which also recorded responses and response latencies. Each session began with self-paced, computer-presented instructions for each task. The experimenter also went over the instructions verbally to ensure that participants understood the task. On average, slightly more explanation was required for the older participants. Participants were told that we were studying working memory and were particularly interested in young and older adults' ability to remember pictures of common objects, locations, and objects together with their locations, for brief periods of time. The entire procedure, including the test requirements, was fully explained to the participants at the beginning of the experiment.

A graphical representation of the procedure is presented in Figure 1. The start of each trial was signaled by a cue, presented in the middle of the screen

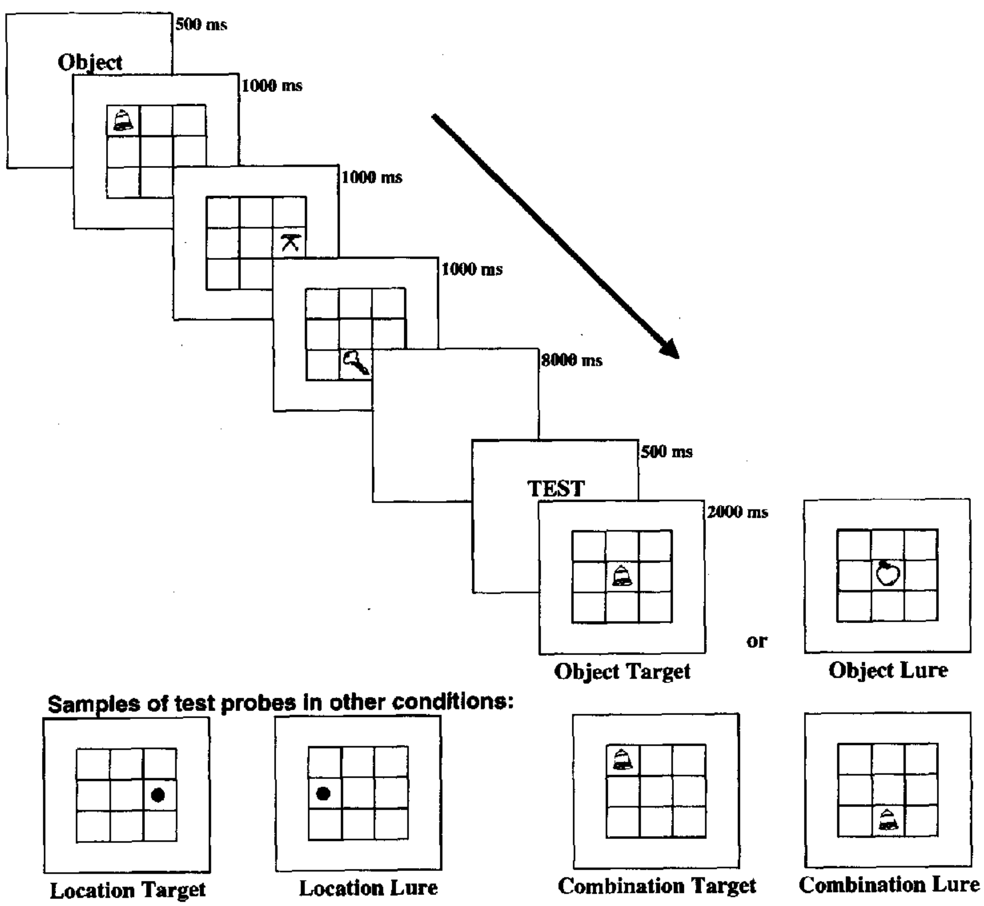


Figure 1. Procedure and timeline for Experiment 1 together with samples of test probes for each condition. In all conditions, each of the three study objects was presented in one of eight colors (each object on a given trial was in a different color). The test objects were always black and white.

for 500 ms (i.e., "object," "location," or "object + location"), intended to remind participants of what they were supposed to remember on that trial. This was followed by three 3 × 3 grids, approximately 7.9 × 7.9 in. (20 × 20 cm), each of which contained a drawing of a different object (e.g., bell, whistle, key; Snodgrass & Vanderwart, 1980) in a different location. Each appeared in a different one of eight colors (yellow, red, blue, brown, green, pink, purple, orange) and was sized so as to nearly fill the square in which it appeared. Mean familiarity and complexity of objects was roughly equated between conditions using the Snodgrass and Vanderwart norms, and objects, locations, and object–location combinations were counterbalanced between conditions. The grids were presented sequentially, in the center of the screen, for 1 s each. Presenting each study array separately encouraged participants of both age groups to attend equally to each stimulus.

In separate blocks of trials, participants were told to remember either (a) only what objects were presented (object trials), (b) only which locations were filled (location trials), or (c) each object and its location together as a unit (combination trials). They were told that color was irrelevant. Participants were tested (as described below) after an unfilled 8-s delay, during which the screen was blank. The instructions emphasized that each trial was independent and that participants would be tested on only the information they were told to study during the current trial. The first block of each type was preceded by more specific task instructions, including descriptions of the test probes, and two practice trials using abstract shapes as the objects. Briefer instructions were given before the subsequent blocks of each type as a reminder.

Participants received a total of 12 blocks of trials (4 object, 4 location, and 4 combination). Each block contained eight trials. The order of blocks was the same for all participants: object only, location only, combination, location, combination, object. After these 6 blocks, there was a brief rest period (approximately 2–3 min) followed by a second set of 6 blocks of trials in the same order. This order was used to allow participants to get comfortable with the general task (timing, pressing the keys, etc.) before they got to the more difficult task (i.e., the combination condition). In several subsequent experiments (e.g., Mitchell, Johnson, Raye, & D'Esposito, 1999, in press) in which order of conditions was counterbalanced, the same general pattern (to be discussed below) was obtained. Also, it should be noted that data recently reported by Uttl, Cosentino, and Graf (1999) suggest that older adults are not necessarily more affected by fatigue during cognitive tasks than young adults.

In all trials, 8 s after the third study array was presented participants were given a recognition test on a single object, location, or object + location combination. In all conditions, an equal number of targets and lures were randomly intermixed and targets were sampled equally often from each of the three study positions. Each test was cued by the word "TEST" presented in the middle of the screen for 500 ms. Targets were as follows: For object trials, a black-and-white line drawing of an object studied on that trial presented in the center square of the 3 × 3 grid (note that the center square of the grid was never used as a location at study; although this fact was not revealed to participants explicitly, it did not take long for participants to notice); for location trials, a filled black circle, just slightly smaller than the objects, presented in one of the squares that had been filled on that trial; and for combination trials, a black-and-white line drawing of an object presented in its correct, studied location. Lures were as follows: for object trials, a black-and-white line drawing of an object not studied on that trial presented in the center square of a grid (note that except for when a lure occurred on the first trial, the object used at test was one that had been studied on a previous trial); for location trials, a filled black circle presented in a location not filled on that trial; and for combination trials, a black-and-white line drawing of an object studied on that trial located in a square that had been filled by a different object on that trial (i.e., a studied object and studied location were re-paired). Test probes were presented for 2 s. Participants were instructed to respond as quickly as possible without sacrificing accuracy by pressing the button on the keyboard marked "Y" (the *S* key) for "yes" and the button marked "N" (the *L* key) for "no." Participants could respond any time after the test probe appeared. The computer continued to collect responses and response latencies during the 2-s intertrial interval, during which the computer screen was blank.

Results and Discussion

Table 2 shows the proportion of hits and false alarms (together with standard error of the mean) for each age group in each

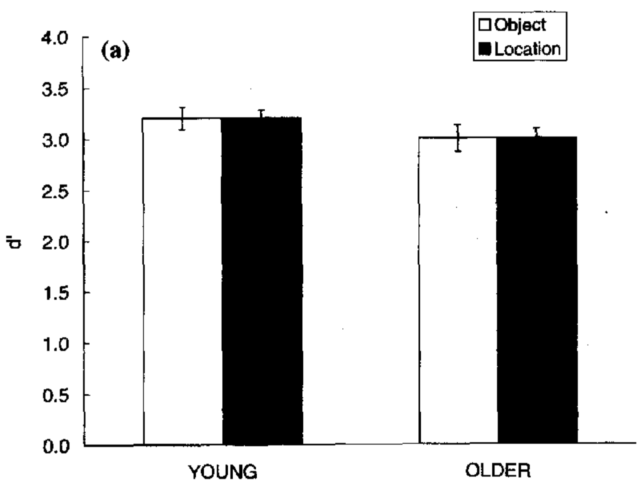
Table 2
Mean Proportion of Hits and False Alarms and Standard Errors of the Mean for Experiments 1 and 2 by Condition and Age

Age group	Condition	Hits		False alarms	
		M	SE_M	*M*	SE_M
Experiment 1					
Young	Combination	.93	.01	.05	.02
	Location	.95	.01	.04	.01
	Object	.93	.02	.03	.01
Older	Combination	.92	.01	.15	.03
	Location	.90	.01	.04	.01
	Object	.89	.02	.03	.01
Experiment 2					
Young	Study feature–test feature	.95	.01	.04	.01
	Study combination–test combination	.66	.02	.09	.03
	Study combination–test combination (old + new)	—		.01	.01
	Study combination–test single feature	.96	.01	.04	.01
	Study combination–test two features	.98	.01	.05	.02
Older	Study feature–test feature	.93	.02	.06	.02
	Study combination–test combination	.70	.02	.27	.06
	Study combination–test combination (old + new)	—		.05	.02
	Study combination–test single feature	.90	.02	.06	.02
	Study combination–test two features	.86	.03	.18	.04

Note. The means presented here are the means of the unadjusted scores.

condition; however, we focus our primary discussion on accuracy scores expressed as d'. Before calculating d', we adjusted scores as follows: $p(H) = 1$ was recalculated as $1 - 1/(2N)$; $p(FA) = 0$ was recalculated as $1/(2N)$, where N = the maximum number of hits or false alarms possible (Macmillan & Creelman, 1991). We indicate effect size in terms of η^2.

An initial 2 (age: young, older) × 2 (condition: object, location) contrast confirmed that there were no differences between the age groups in memory for objects or locations, $F(1, 46) = 2.53$, $MSE = 0.37$, $p > .10$, for the main effect of age, and $F(1, 46) < 1$ for the interaction (see Figure 2a). Thus, we collapsed across single feature conditions to examine the combination deficit. A 2 (age: young, older) × 2 (condition: single feature, combination) contrast revealed a marginal age by condition interaction, $F(1, 46) = 3.67$, $MSE = 0.21$, $p = .06$, $\eta^2 = .07$ (see Figure 2b). As predicted, planned comparisons confirmed that older adults performed significantly below young adults in the combination condition, $t(46) = 2.49$, $p < .05$, $\eta^2 = .12$, but not significantly below young adults in the single-feature condition, $t(46) = 1.59$, $p > .10$, $\eta^2 = .05$.



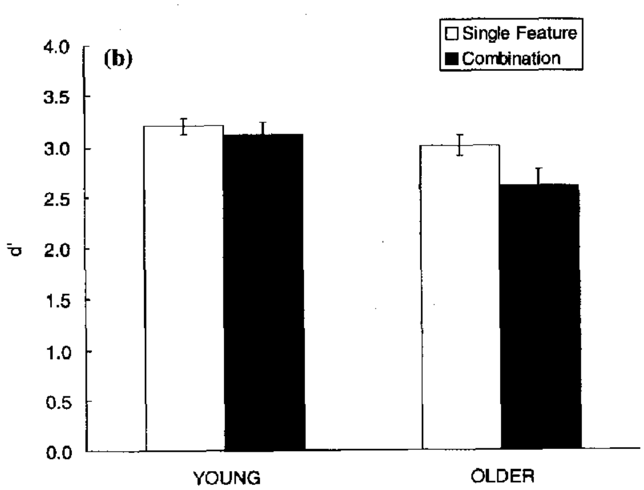


Figure 2. (a) d' scores for each age group in the object-only and location-only conditions (with standard errors) for Experiment 1. (b) d' scores for each age group on feature-only trials (collapsed across object-only and location-only trials) and combination trials (with standard errors) for Experiment 1.

Older adults' poor performance in the combination condition came about because they were more likely than young adults to produce false alarms in response to the re-paired lures, $t(46) = 3.20$, $p < .01$, $\eta^2 = .18$ (see Table 2). In contrast, hits were similar for the two age groups ($p > .10$, $\eta^2 = .003$; see Table 2).[2] This pattern is exactly what we would expect given our methodology. That is, older adults' deficit in the combination condition should manifest itself as difficulty in rejecting lures. To see why this is the case, consider first the feature-only conditions. Recall that in the feature-only conditions the lures were made up of novel features—features that, although experienced elsewhere in the experiment, were not seen on that trial. Such novel items should be relatively easy for both young and older adults to reject given their equally good memory for studied features. In other words, both accepting feature targets and rejecting lures can be done well if one has good memory for the features studied, and, as expected, our older adults had no problem remembering the features. But our prediction would be different in the combination condition, in which memory for bound features was most directly tested by the lures. In the combination condition, good memory for the two features alone could be enough to allow good performance on targets; even if participants said "yes" to a target only because they remembered studying both features, they would still be correct. And, in fact, our older adults did quite well accepting targets in the combination condition. Accurately rejecting lures, on the other hand, requires not only memory for the features but memory for how the features were paired during study. Thus, for the combination trials, performance on lures is most telling about memory for the unique association between the features studied on that trial because only by remembering how the features were originally combined can lures be rejected.

Note that the intent of this design (i.e., small set sizes and brief, unfilled delay intervals) was to examine memory for combinations of features under conditions in which memory for the features themselves was not at issue. Clearly, performance on the feature-only trials suggests we succeeded in that goal. Thus, it would

[2] Analysis of criterion scores c (Macmillan & Creelman, 1991, p. 33) shows that older adults were significantly biased to say "yes" in the combination condition ($c = -.17$) as compared with either the object ($c = .19$) or the location condition ($c = .16$), which did not differ from each other, $F(2, 46) = 18.19$, $MSE = 0.05$, $p < .001$. The young adults did not exhibit such a bias ($c = .05$, .03, and .08 for the combination, location, and object conditions, respectively; $F < 1$). One possibility is that, lacking memory for the associations between the features of the studied stimuli, older adults defaulted to a less stringent criterion, perhaps a familiarity-based strategy whereby they said "yes" to familiar features (e.g., see Verfaellie & Treadwell, 1993, for a similar suggestion about individuals with amnesia). However, it is hard to make too much of this given that the same pattern did not obtain in Experiment 2, in which older adults exhibited lower accuracy in the combination condition relative to features only, but no difference in criterion (in Experiment 2, c for older adults = .02, .01, .04, and $-.18$ for the study feature–test feature, test combination, test one feature, and test two feature conditions, respectively; $F < 1$).

appear that under conditions in which older and younger adults both have good memory for feature information, older adults exhibit a deficit in binding those features into more complex memories.

The pattern of response latencies supports the general idea that older adults experienced problems in remembering the combined information that they did not experience in remembering features only. It has been well-established that older adults are slower generally than young adults on most cognitive tasks (e.g., Cerella, 1990). However, there is also evidence that this difference may be exaggerated in more difficult conditions (e.g., Myerson, Hale, Wagstaff, Poon, & Smith, 1990). Therefore, we might expect that our older adults would be especially slowed in the combination condition compared with the feature-only conditions. And, again, because the lures most directly test memory for the association between the features, we would expect older adults' difficulties in the combination condition to be manifested as slower response times to reject the lures. Thus, if our hypothesis is correct, regardless of overall age differences in the baseline of response latencies, older adults should show a larger target–lure difference on combination trials than feature trials.

To examine this possibility, we calculated a difference score for each participant (median response time on lures minus median response time on targets) in each condition, which should control for individual differences within groups in base level of response latencies by anchoring participants' lure performance to their own target performance. We compared combination difference scores with a feature difference score calculated as the mean of the object and location difference scores. We included both correct and incorrect trials because we were interested here in time to make judgments about feature combinations, regardless of whether the judgment is correct. It seems reasonable that older adults' inability to bind features effectively will lead to slower responding on combination trials than single-feature trials because they must have encoded a less cohesive set of features.

Figure 3 shows the lure–target difference scores for combination and feature trials for each age group (with standard errors). Both young and older adults were faster on lures than targets in the feature-only conditions, as evidenced by the negative difference scores, and slower to respond to lures than targets in the combination condition, as evidenced by the positive difference scores. More important, however, a 2 (age) $\times$ 2 (condition) contrast revealed a significant interaction, $F(1, 46) = 11.14$, $MSE = 20{,}652.16$, $p < .01$, $\eta^2 = .19$, indicating that the age difference was greater on combination, $t(46) = 3.04$, $p < .01$, $\eta^2 = .11$, than feature trials, $t(46) = 1.57$, $p > .10$, $\eta^2 = .05$. Moreover, the difference between performance on the feature trials as compared with the combination trials was nearly twice as large for the older ($M = 419.89$) than the younger adults ($M = 224.08$), a reliable difference, $t(46) = 3.34$, $p < .01$, $\eta^2 = .19$, that remained marginally significant with log-transformed scores (mean differences between the feature and combination conditions = 0.09 and 0.12 for young and older adults, respectively), $t(46) = 1.57$, $p = .06$, one-tailed, $\eta^2 = .06$. Thus, the latencies are consistent with the idea that the older adults were having difficulty in rejecting the combination lures. They may have responded slowly in an attempt to improve their accuracy. However, as indicated by the accuracy data, their slower responding did not make up for their deficit in the processes that support binding during initial encoding. Together, the latency and accuracy data support the conclusion that older adults experienced difficulty binding object and location information in this working memory task.

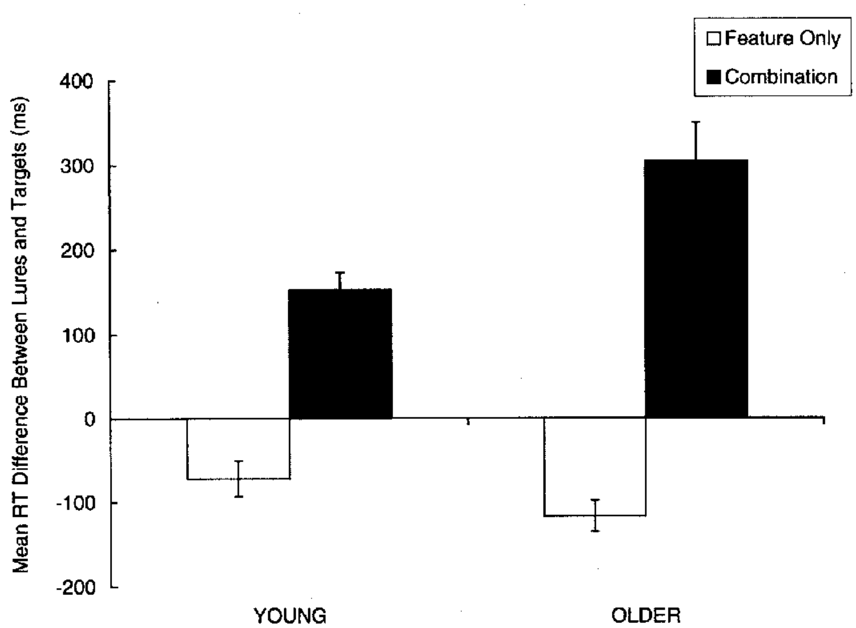


Figure 3. Response latency difference scores for each age group in Experiment 1, calculated as each participant's median latency on lure trials minus that participant's median latency on target trials (with standard errors). The feature-only scores represent the mean of the object and location conditions. RT = response time.

Experiment 2

It is possible that older adults' poor performance on the combination trials in Experiment 1 does not reflect only poor connections between features but also the fact that combination trials involve more information than feature trials. Generally speaking, remembering combinations of features requires processing more information than remembering individual features: (a) To the extent that binding does not necessarily reduce information through "chunking" (and even if it does, such chunking is not, presumably, instantaneous), our participants must maintain information about six features on combination trials but only three on feature trials, and (b) more information must be assessed at test on combination trials (e.g., information about two features and relational information) than on feature trials. Thus, it may be that older adults have more difficulty than young adults on combination trials because they are unable to maintain as much information over the delay (a *maintenance load* hypothesis), they have more difficulty evaluating multiple features at test (a *test load* hypothesis), or both.

Experiment 2 was designed to evaluate the possible contribution of these two factors to older adults' deficits on combination trials. With respect to the maintenance load hypothesis, if older adults have more difficulty maintaining larger amounts of information, they should also have difficulty on a single-feature test if it follows instructions to combine features at encoding. With respect to the test load hypothesis, if older adults have more difficulty assessing multiple features at test, they should show poor performance whenever two features are tested, even if memory for the relations between the features is not necessary.

We used the same stimuli and trial structure as in Experiment 1, with added conditions to test the maintenance load hypothesis and the test load hypothesis. Table 3 provides a summary of the conditions (described below). To replicate Experiment 1, we tested young and older adults' feature memory after instructions that encouraged them to remember a single feature (i.e., either object or location only, as in Experiment 1; Condition 1 in Table 3) and we tested their memory for combinations of features (i.e., object + location) after combination encoding instructions that encouraged them to try to remember the features combined (i.e., combination condition as in Experiment 1; Condition 2a in Table 3). One new condition tested participants' feature memory after combination encoding instructions (Condition 3 in Table 3). Good memory on the part of the older adults for the studied features would argue against a maintenance load explanation for older adults' performance decrements in the combination condition. A second new condition probed two values of a feature (i.e., two objects or two locations) following combination instructions at encoding (Condition 4 in Table 3). This condition tested the possibility that older adults' memory decrements in the combination condition were related to the need to respond to a test stimulus containing two features (i.e., an object and its location), as opposed to only one feature (i.e., either an object or a location)—the test load hypothesis. To the extent that increased test load can fully account for older adults' problems on combination test trials, their performance on the two-feature and combination tests should be equivalent.

Method

Participants

Sixteen undergraduates (mean age = 19.6 years, SD = 0.93 years, range = 18 to 22 years) and 16 healthy older adults (mean age = 75.1 years, SD = 2.43 years, range = 71 to 80 years) were recruited from the same pools as in Experiment 1. The young and older adults had completed equivalent levels of education (young: M = 13.1 years, SD = 0.56 years; older: M = 13.4 years, SD = 2.10 years), $t(30) < 1$. Scores from standard memory tests (obtained within a month of the present experiment) were available for 12 of the 16 older adults; they are provided in Table 1. The young adults participated in exchange for credit toward a course requirement; older adults were paid for their participation.

Design

Age was, obviously, a between-subjects variable. There were four conditions (described below), defined in terms of encoding instructions and test type, manipulated within subjects.

Materials and Procedure

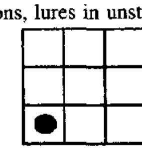
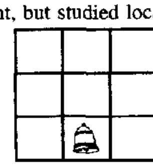
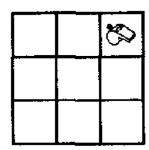
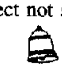
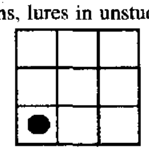
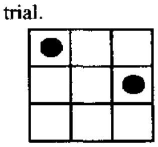
The stimuli and trial timeline were the same as those used in Experiment 1 (with repeated objects). However, in this experiment, participants were first given a single block of 16 trials of one of the single-feature conditions (i.e., either object only or location only), followed by a block of 16 trials of the alternate single-feature condition (starting order was counterbalanced). On these trials, which we refer to as *study feature–test feature* trials (Condition 1 in Table 3), participants were told to study only one feature (either object or location) and were then tested on that feature as in Experiment 1. Each block was preceded by verbal instructions with illustrations. The first block began with six practice trials using stimuli similar—but not identical—to the experimental stimuli. All trials, regardless of condition, began with the cue word "STUDY" (500 ms). The test probes (half targets, half lures, randomly intermixed) for these single-feature conditions were the same as in Experiment 1, with the exception that object memory was tested using a black-and-white line drawing in the center of the screen with no grid.

The study feature–test feature blocks were followed by 104 trials in which participants were given combination encoding instructions. That is, as on the combination trials of Experiment 1, participants were told that they should always try to remember each object plus its location combined as a unit. They were further informed that on these trials there would be several different types of test probes, which would vary randomly from trial to trial. It is important to emphasize that for all of the following conditions load and processing should have been identical until presentation of the test probe because the participants were told to remember object + location information on all trials—all that varied was the nature of the test probe. As summarized in Table 3, there were three test conditions on these study combination trials, described in the following sections.

Combination test (Condition 2a in Table 3). On 32 trials (16 targets, 16 lures), participants were tested for their memory for combinations of features (i.e., objects plus locations) as in Experiment 1. Test probes were as described in Experiment 1. We expected to replicate older adults' deficit in this condition, hereinafter referred to as the *study combination–test combination* condition.

We argued earlier that using lures composed of studied, but re-paired, features is the most sensitive way to test for age-related binding deficits because good performance under such conditions cannot be based on feature information alone. Information about the relationship between the two features is critical. We tested this assumption by including another eight study combination–test combination trials (all lures) in which the test probes contained one of the studied feature values and one feature value that was not studied on that trial (although it would be old from another

Table 3
Outline of Study and Test Conditions in Experiment 2

Study instruction	Test probe
	Study feature
Condition 1: Study feature–test feature Remember objects only[a]	Single black-and-white object in the center of the screen. Target = studied object; lure = object not studied on present trial. 
Remember locations only[a]	Black dot in one square of a grid. Targets appeared in studied locations, lures in unstudied locations. 
	Study combination
Condition 2: Study combinations–test combinations Remember objects + locations[a] (Condition 2a)	Black-and-white object in a location other than center. Targets = studied object + location combination; lures = re-paired features (i.e., studied object in a different, but studied location). 
Remember objects + locations (old + new; Condition 2b)	Black-and-white object in a location other than center. This involved only lures, which were composed of one studied feature plus one novel feature. 
Condition 3: Study combinations–test one feature Remember objects + locations	Single black-and-white object in the center of the screen. Targets = studied objects; lures = object not studied on present trial. 
Remember objects + locations	Black dot in one square of the grid. Targets appeared in studied locations, lures in unstudied locations. 
Condition 4: Study combinations–test two features Remember objects + locations	Two black-and-white objects presented side-by-side. Targets = two studied objects; lures = one studied object and one object not studied on present trial. 
Remember objects + locations	Two black dots presented in different squares of a single grid. Targets = both dots in studied locations; lures = one studied location and one location not studied on present trial. 

Note. At study, in all conditions, participants saw three 3 × 3 grids presented sequentially for 1 s each, with each grid containing a different colored object in a different location as in Experiment 1. Thus, all that varied between conditions were the encoding instructions and test probe (illustrated). For all conditions in the bottom panel of the table (Conditions 2–4), encoding instructions were identical (i.e., study object + location) and all that varied was the nature of the test probe. Thus, the load and the processes engaged in these "study combination" conditions should have been identical until presentation of the test probe.
[a] These conditions are exactly as in Experiment 1.

trial; e.g., a studied object in an unstudied location). For half of these trials the object was new, and for half the location was new. These trials are hereinafter referred to as *study combination–test combination (old + new)* trials (Condition 2b in Table 3). To the extent that these lures can be rejected out of hand if one remembers what features were studied on that trial, we expected older adults' inability to reject combination lures to be attenuated on these trials.

One-feature test (Condition 3 in Table 3). On 32 trials (16 targets, 16 lures), participants were tested for either object information only (i.e., a single black-and-white drawing in the middle of the screen without a grid; 16 trials) or location information only (i.e., a single black dot in one of the grid locations; 16 trials). These trials are hereinafter referred to in the aggregate as *study combination–test one feature.* Recall that at acquisition participants were told to try to remember the objects and their locations combined as single units. Thus, if the age-related decrement found in Experiment 1 is due solely to the increase in amount of to-be-remembered information (i.e., maintenance load) on the combination trials, we would expect a similar decrement in older adults' performance in this condition compared with the study feature–test feature trials.

Two-features test (Condition 4 in Table 3). We also wanted to include a feature test in which the test demands were equivalent, in terms of number of features, to those of the combination condition. Participants received 32 two-feature test trials. On the 16 object test trials (8 targets, 8 lures), two black-and-white objects appeared side by side without a grid. On the 16 location test trials (8 targets, 8 lures), two black dots appeared, each in a different location of the same grid. In this way, participants were tested for their memory for two values from a single feature category. Test targets were made up of two feature values studied on that trial (two studied objects or two dots each in a studied location); lures were composed of one studied value and one unstudied value (e.g., one object seen during that trial and one object seen on some previous trial). The order–position of the old and new feature were counterbalanced. Note that because both feature categories involved eight values that were repeated across trials, both values in a lure would have been previously studied, just not necessarily on that trial. This condition will hereinafter be referred to as the *study combination–test two features* condition. We reasoned that if older adults' decrements on the combination trials in Experiment 1 were due merely to the need to evaluate multiple features (i.e., the test load hypothesis), we would expect their performance on the two-feature test to be as low as their performance on the combination test. To the extent that a test load deficit contributes to, but cannot fully explain, older adults' poorer performance on combination trials, we would expect their performance in this condition to be poorer than in the single-feature test condition yet better than the combination test condition.

General procedures for study combination trials. The same pseudorandom order of trials was used for all participants. These 104 study combination trials were broken up into three blocks to allow participants brief rest periods (approximately 2–3 min). The first block was preceded by detailed instructions, including graphical presentation of the trial progression and a thorough discussion of all test probe types. For all test probe types, targets were sampled equally often from each of the three study positions. Participants were told that they should say "yes" if the test probe contained either features or feature combinations that were exactly as studied in the trial, and "no" otherwise. Participants were also given 10 practice trials to acquaint them with the test requirements. If necessary, older adults were given more practice. The experimental procedure began only after the experimenter was sure the participant understood the nature of the task requirements.

Results and Discussion

Hits and false alarms are shown in Table 2; our focus is on d'. Again, an initial 2 (age: young, old) × 2 (condition: object, location) contrast on d' scores shows that there were no significant differences between the two groups on their memory for objects only ($M = 2.85$, $SE = .09$, and $M = 2.82$, $SE = .07$, for young and older adults, respectively) or locations only ($M = 2.74$, $SE = .09$, and $M = 2.50$, $SE = .20$, for young and older adults, respectively), $F(1, 30) = 1.05$, $MSE = 0.28$, $p > .10$, for the main effect of age, and $F(1, 30) < 1$ for the interaction. Therefore, the remaining analyses were conducted collapsed across type of feature. To explore how well the pattern of data fit our specific hypotheses, we conducted a series of planned contrasts.

Older Adults Again Exhibit a Decrement in Their Memory for Combinations of Features

The study feature–test feature and study combination–test combination conditions of the present experiment correspond to the conditions of Experiment 1. Test combination (old + new) lures were not included in the following analyses; they are discussed separately. If we look at these conditions (see Figure 4), we see that the central finding of Experiment 1 was replicated: Although the age by condition interaction in a 2 (age) × 2 (condition) contrast failed to reach significance, $F(1, 30) = 2.38$, $MSE = 0.23$, $p > .10$, $\eta^2 = .07$, planned comparisons confirmed that the older adults again showed poorer performance than the young adults in the combination test condition, $t(30) = 2.24$, $p < .05$, $\eta^2 = .14$, but not the feature-only condition, $t(30) = 0.96$, $p > .10$, $\eta^2 = .03$ (see Figure 4). Unlike Experiment 1 (and a similar subsequent experiment; Mitchell et al., in press), the young adults also showed poorer performance in the combination test condition than in the study feature–test feature condition, $t(15) = 10.49$, and $t(15) = 9.07$, for young and older adults, respectively, $p < .0001$, $\eta^2 > .80$ for both (see Figure 4)—an interesting finding that we return to in the General Discussion section. Recall, however, that young adults' performance on combination test trials was still better than older adults'. Thus, as in Experiment 1, older adults showed a deficit on combination test trials although memory for individual features was quite good.

As in Experiment 1, we calculated a difference score for each participant in each condition that consisted of that participant's median response latency on lures minus that participant's median response latency on targets. Figure 5 shows these difference scores for each age group (with standard errors) for each condition. Again, both young and older adults were faster on lures than targets in the study feature–test feature condition, as evidenced by the negative difference scores, and slower to respond to lures than targets in the study combination–test combination condition, as evidenced by the positive difference scores. The difference between performance on the study feature–test feature trials as compared with the study combination–test combination trials was nearly twice as large for the older ($M = 283.63$) than the younger adults ($M = 154.34$), and a 2 (age) × 2 (condition) contrast showed a marginally significant interaction, $F(1, 30) = 3.58$, $MSE = 18{,}668.61$, $p < .07$, $\eta^2 = .10$ (see Figure 5). However, the interaction was not significant when log-transformed latencies were analyzed ($F < 1$, $\eta^2 = .02$).

As expected, both young, $t(15) = 3.27$, $p < .01$, $\eta^2 = .42$, and older adults, $t(15) = 4.58$, $p < .001$, $\eta^2 = .58$, were better able to reject combination lures when they contained one studied and one unstudied feature—that is, combination (old + new) test—than

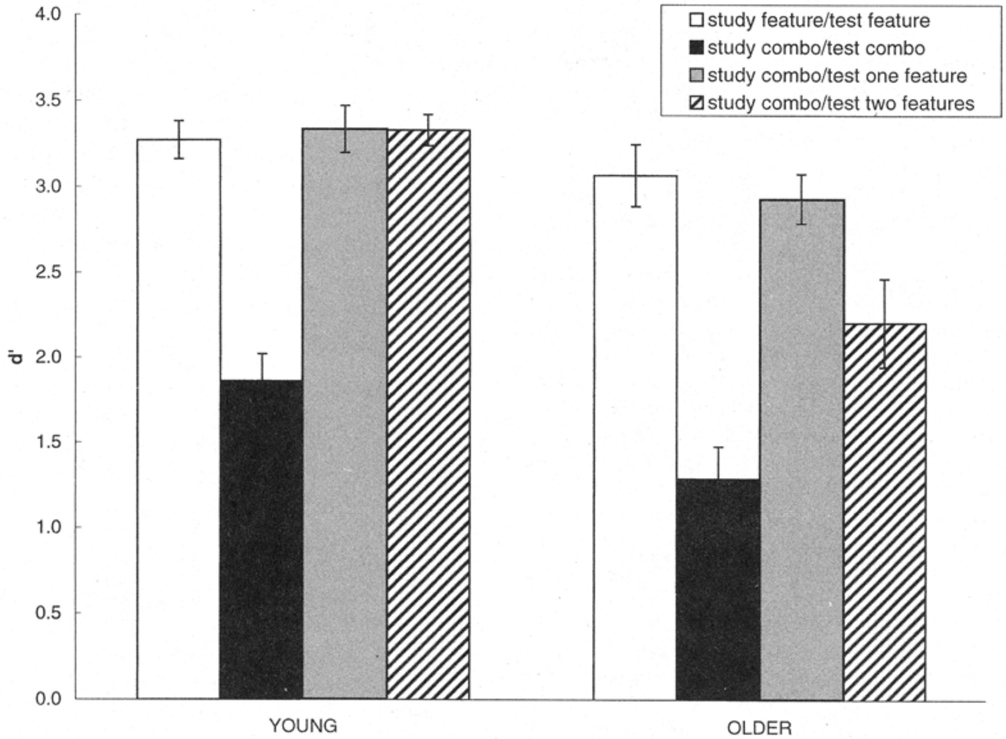


Figure 4. d' scores for each age group in each condition (with standard errors) for Experiment 2.

when the lures were composed of two studied, but re-paired, features—that is, the standard combination test. Also the Age × Condition interaction was reliable, $F(1, 30) = 7.14$, $MSE = 0.01$, $p \leq .01$, $\eta^2 = .18$ (see Table 2). The improvement in older adults' ability to reject the combination lures when they contained one new feature supports the conclusion that they had good feature memory. This finding also suggests that older adults' difficulties on the standard combination trials cannot be due solely to an

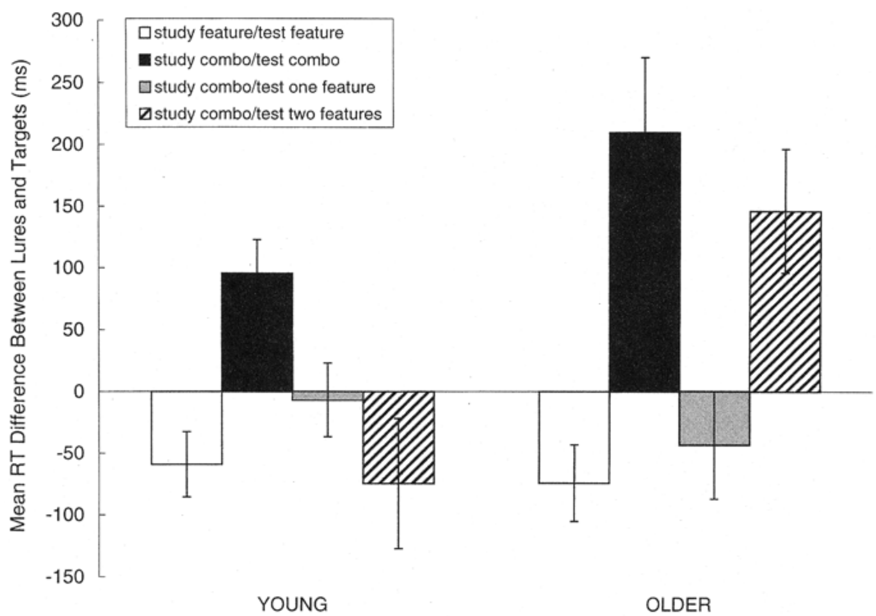


Figure 5. Response latency difference scores for each age group in Experiment 2, calculated as each participant's median latency on lure trials minus that participant's median latency on target trials (with standard errors). RT = response time.

inability to deal with multiple features at test—an issue we return to shortly.

To What Extent Does Increased Maintenance Load Contribute to Older Adults' Decrement in the Combination Test Condition?

If older adults' poor performance in the combination test condition is due solely to increased maintenance load, we should also expect to see an age-related decrement on a single-feature test following combination encoding instructions. However, direct comparison of the study feature–test feature and study combination–test one feature conditions with a 2 (age) $\times$ 2 (condition) contrast showed no reliable interaction, $F(1, 30) = 2.63$, $MSE = 0.06$, $p > .10$, $\eta^2 = .08$ (see Figure 4). The overall pattern of performance discussed thus far suggests two things. First, there is little evidence that decreased maintenance capacity on the part of older adults can explain age-related combination memory decrements in this task. Second, this pattern suggests that trying to bind the information at encoding—that is, trying to remember the combined information as units—did not hurt older adults' memory for the features themselves. They clearly have more information available about the features than about the associations between features.

To What Extent Does Increased Test Load Contribute to Older Adults' Decrement in the Combination Test Condition?

We previously suggested that the difference in the number of features involved in the test probes might contribute to older adults' difficulties on the combination test trials (i.e., a feature test probe requires a decision about only one feature value, whereas a combination test probe requires a joint decision about two feature values together with a decision about their relationship). Although older adults' relatively good performance on the combination (old + new) trials suggests that this cannot be the entire story behind their combination test deficits, the hypothesis that older adults may have difficulty dealing with multiple features at test gained some support in this experiment. A 2 (age) $\times$ 2 (condition: study combination–test one feature vs. study combination–test two features) contrast showed a reliable interaction, $F(1, 30) = 7.46$, $MSE = 0.27$, $p = .01$, $\eta^2 = .19$ (see Figure 4); the age decrement was greater on the two-feature test, $t(30) = 4.11$, $p < .001$, $\eta^2 = .35$, than the one-feature test, $t(30) = 2.01$, $p = .05$, $\eta^2 = .11$. Moreover, only older adults' performance was lower on the two-feature than the one-feature test, $t(15) = 0.02$, $p > .50$, $\eta^2 < .01$, and $t(15) = 3.11$, $p < .01$, $\eta^2 = .39$, for young and older adults, respectively. Thus, with encoding instructions and maintenance load held constant, older, but not younger, adults showed a deficit when they were required to make a memory judgment about two features compared with when they were required to make a decision about one. This decrement obtained even though they did not have to know the relationship between features to make the decision.

Response latencies are interesting with regard to this age-related test deficit. As is clear from Figure 5, on the one-feature test both young and older adults were slower on targets than lures (as indicated by the negative scores), but on the two-feature test the young were again slower on targets than lures and the older adults were slower on lures than targets (as indicated by the positive score). This produced a reliable interaction in the 2 (age) $\times$ 2 (condition) contrast, $F(1, 30) = 12.43$, $MSE = 21{,}250.6$, $p = .001$, $\eta^2 = .28$, indicating that the difference in the time it took young and older adults to respond to lures versus targets was greater for two-feature trials, $t(30) = 3.03$, $p < .01$, $\eta^2 = .23$, than one-feature trials, $t(30) = 0.69$, $p > .10$, $\eta^2 = .02$. Both the interaction, $F(1, 30) = 10.16$, $MSE = 0.002$, $p < .01$, and planned comparisons, $t(30) = 2.83$, $p < .01$, and $t(30) = .41$, $p > .05$, for the two-feature and one-feature trials, respectively, remained significant when log-transformed scores were used. Such a pattern supports the idea that the older adults were having trouble assessing multiple features at test, regardless of the need to remember the association between them.

If older adults' deficit on the combination trials were due solely to the need to respond to a probe containing two features as opposed to one feature, we would expect their performance on the two-feature test to be as low as their performance on the combination test—but, as is clear from Figure 4, it was not. Older adults had poorer performance in the combination test condition than in the two-feature test condition, $t(15) = 4.47$, $p < .001$, $\eta^2 = .57$, suggesting that test load alone cannot fully account for their decrement in the combination test condition. The young adults also exhibited a rather large deficit on combination trials compared with the two-feature trials, $t(15) = 9.61$, $p < .0001$, $\eta^2 = .85$, creating a reliable interaction in the 2 (age) $\times$ 2 (condition: study combination–test two features, study combination–test combination) contrast, $F(1, 30) = 4.69$, $MSE = 0.26$, $p < .05$, $\eta^2 = .13$. Together with the finding that young adults' performance on combination trials was significantly lower than their own one-feature test performance in this experiment, $t(15) = 9.21$, $p < .0001$, $\eta^2 = .85$, these results indicate that young adults had more difficulty on the combination trials in Experiment 2 than in Experiment 1—an issue we return to in the General Discussion section. Central to the present question is the fact that the age deficit was, nevertheless, reliable in both the combination test condition, $t(30) = 2.24$, $p < .05$, $\eta^2 = .24$, and the two-feature test condition, $t(30) = 4.11$, $p < .001$, $\eta^2 = .51$.

In sum, compared with their performance on the combination test, older adults' relatively good performance on the study combination–test one feature trials (which, it is important to note, did not differ from the study feature–test feature condition) suggests that trying to bind at study does not lead to worse memory overall on the part of older adults. That is, they do not simply exhibit a general working memory decrement for features under conditions of increased maintenance load. They clearly have a large proportion of the studied features in memory. Rather, it appears that older adults have two specific memory-related deficits. First, as in Experiment 1, older adults showed a performance decrement on the combination test compared with the single-feature test after combination encoding instructions, demonstrating that they have a binding problem at encoding. Second, older adults also showed a performance decrement on the two-feature test compared with the single-feature test after combination encoding instructions, suggesting they have an access–evaluation problem at test when a probe contains multiple features. Of importance, their accuracy in the combination condition was below that in the two-feature test condition, indicating that their binding deficit leads to difficulties above and beyond their difficulty in dealing with multiple features at test.

General Discussion

The present research addressed the question of whether older adults' long-term memory deficits for complex information reflect binding deficits that stem from deficits in reflective component processes used during initial encoding and consolidation of memories (Chalfonte & Johnson, 1996; Johnson & Chalfonte, 1994). We assume that working memory consists of reflective processes by which information is encoded for longer term memory (e.g., Johnson, 1992). Thus, we expected age-related feature binding deficits in a working memory task. Indeed, that is what we found. Both experiments show that after being presented with three successive arrays, each of which contained a different object in a different location, older adults had no problem recognizing individual features (either objects or locations). However, in both experiments older adults were significantly impaired in their ability to recognize combinations of features (an object and its location). These results suggest that the age-related source memory deficits seen in long-term memory have their origin, at least in part, at encoding, where disrupted or inefficient reflective processes result in less well bound complex memories.

We say "in part" because it appears that there also may be age-related differences in how information is accessed and evaluated at test. Experiment 2 demonstrates a new specific age-related test load deficit. Older adults showed an impairment when tested with two individual features, compared with a test that presented only one feature, under conditions that were equated on encoding demands (i.e., the two-feature test vs. the one-feature test after combination encoding instructions). This finding is provocative, and two hypotheses suggest themselves.

The first is that older adults' difficulties on combination trials were due entirely to their difficulty in dealing with multiple cues at test (i.e., a test load deficit). Older adults may experience difficulty whenever they are required to access multiple features, evaluate them, or both (e.g., Johnson, De Leonardis, Hashtroudi, & Ferguson, 1995; but see Bayen & Murnane, 1996). From the present data, we cannot determine whether the age-related test deficit demonstrated here represents age-related problems at the access or the evaluation phase of remembering. Nevertheless, to the extent that evaluation of complex (e.g., autobiographical) memories typically involves multiple features, such a deficit could help explain age-related source and reality monitoring deficits often found in longer term memory procedures (e.g., Ferguson et al., 1992; Hashtroudi, Johnson, & Chrosniak, 1990; Henkel, Johnson, & De Leonardis, 1998).

However, there are several reasons to suspect that older adults' difficulty in dealing with multiple cues at test probably is not the entire story behind their combination condition decrement in this working memory task. Older adults' performance improved significantly when the combination lures were composed of one studied feature and one new feature—the combination (old + new) test—compared with when they were tested with lures composed of two studied but re-paired features—the standard combination test. If their difficulty on combination trials was due to number of features, per se, this manipulation should have no effect. More important, older adults exhibited significantly poorer performance on combination test trials compared with two-feature test trials, suggesting that they experienced additional difficulty in the combination condition beyond their difficulty in dealing with two features.

A second hypothesis is that older adults' performance decrement on the two-feature test is another manifestation of a general binding deficit. One possibility is that because of effective interdimensional binding (i.e., object + location), young adults were more likely than older adults to successfully retrieve the corresponding "nontested" feature for each feature in the test probe (e.g., to retrieve the corresponding locations when presented with two objects) and use this information as evidence regarding the oldness of the presented features (see, e.g., Clark & Gronlund, 1996). It is also possible that, in addition to binding object + location information for each study stimulus, young adults were binding objects to objects and locations to locations within each trial (i.e., intradimensional binding). If so, then the two features of the probe could serve as potential cues for each other. To the extent that older adults have an intradimensional, as well as an interdimensional, binding deficit (which seems likely), they would derive less benefit from such cuing. In either case, the age-related deficit in the two-feature condition may be, like the age-related deficit in the combination condition, related to binding deficits and not to greater difficulty with complex tests, per se.

It is notable, in this regard, that the young adults showed a decrement on the combination test in Experiment 2 compared with their performance on the single-feature and two-feature tests. We have subsequently replicated the pattern of results from Experiment 1 using a similar procedure (e.g., Mitchell et al., in press). Thus, it seems reasonable that the relatively poor performance on combination trials shown by the young adults in Experiment 2 was brought about by using a mixed test format. One possibility is that although they were told to remember object + location information, young adults may have shifted some effort (i.e., reallocated reflective processes) from interdimensional to intradimensional binding because the majority of the trials in Experiment 2 tested features. We are currently exploring this possibility.

In any event, the present data raise the possibility that older adults experience two unique sources of working memory deficits—a binding deficit at encoding and a test deficit when presented with multiple cues. More work is necessary to tease apart the relative contribution of these two factors (e.g., whether older adults exhibit a test load deficit independent of an intradimensional binding deficit). Until then, it seems reasonable to tentatively conclude that both of these deficits contribute to older adults' long-term memory deficits for complex information.

Luck and Vogel (1997) recently presented findings suggesting that the binding of visual features may be a relatively automatic encoding process. Their conclusion was that visual working memory stores the complex multifeature results of this automatic binding process. On the one hand, we replicated their findings regarding young adults' feature binding in Experiment 1: As they found, our young adults showed equivalent performance on feature and combination trials. Thus, the present results generalize their findings to a more complex situation. Whereas Luck and Vogel used simple visual features (e.g., color, orientation), our materials involved both a visuospatial feature (i.e., location) and a nameable feature (i.e., objects). Moreover, our test for associations between features was more stringent than Luck and Vogel's because our combination lures were composed of studied, but re-paired, features, whereas their lures involved changes other than in feature combinations, per se.[3] On the other hand, our older adults showed a binding deficit in both experiments, suggesting by some criteria (e.g., Hasher & Zacks, 1979) that memorial binding is not an automatic, effortless process. Additional evidence that memorial binding is a resource-demanding process is the fact that our young adults showed a performance decrement on combination trials in Experiment 2 but not Experiment 1, suggesting that how participants distribute their reflection (i.e., effort or attentional resources) affects the probability of binding feature information.

The divergence of our Experiment 2 results and Luck and Vogel's (1997) findings could be due to differences in materials and procedure. However, it seems more likely that Luck and Vogel were actually testing perceptual binding (e.g., they presented the study arrays for only 100 ms, and their delay period was only 900 ms), whereas the present procedure tapped into a postperceptual process, namely, memorial binding. Alternatively, Luck and Vogel's procedure may tap into different reflective processes than our procedure. For example, it may be that the perceptual conjunction of features can be maintained for a brief period in working memory via *refreshing,* a proposed cognitive component process that extends the initial activation of information so that one can easily shift back to it for further rehearsal (e.g., Johnson, 1992; as suggested by Luck & Vogel's findings). However, this does not necessarily mean that this conjunction is being automatically bound for long-term memory (as suggested by the present findings). Memorial binding may also require reflective operations beyond refreshing—for example, *reactivation* (e.g., Johnson & Chalfonte, 1994; Johnson & Hirst, 1993), recursive *rehearsal,* and so on (e.g., Johnson, 1992). Although there is evidence to suggest that older adults may experience deficits in the refresh process (e.g., Johnson, Raye, & Reeder, 2000), little is known about the contribution of these individual component processes to memorial binding. Exploring these processes is a challenge for the future.

A related target for further investigation is potential age-related differences in explicit rehearsal strategies that may have an impact on binding. Recall that the 8-s delay in this study was unfilled; participants were free to engage in any strategy they developed. One possibility is that older adults had more difficulty than young adults in identifying or maintaining an effective rehearsal strategy. Postexperimental interviews with our participants suggested few obvious differences in the rehearsal strategies of young and older adults in the combination condition. However, the fact that self-reports did not identify strategy differences does not rule out the possibility that older adults developed their strategies later than young adults or that the same strategies simply are less effective in older adults (due to inefficiency of the underlying cognitive processes engaged by them). We are currently investigating these possibilities.

With regard to brain mechanisms, working memory is usually considered to be mediated by the prefrontal cortex (e.g., J. D. Cohen et al., 1997; D'Esposito et al., 1995; Goldman-Rakic, 1995; Smith et al., 1996), whereas binding is thought to be mediated by the hippocampal region (e.g., N. J. Cohen & Eichenbaum, 1993; Eichenbaum & Bunsey, 1995; Henke et al., 1999; Kroll et al., 1996; Moscovitch, 1994; Squire, 1995). The encoding of multiple features into complex long-term episodic memories likely relies on these two areas acting together in cooperation (e.g., Johnson, Hashtroudi, & Lindsay, 1993; Shimamura, 1994; Stuss, Eskes, & Foster, 1994). That is, the binding of multiple memorial attributes probably involves a frontal–hippocampal circuit in which the frontal regions mediate the reflective processes that maintain activation and monitor, control, and coordinate the agendas (e.g., task goals, strategies) necessary to direct the postperceptual, memorial binding of multiple attributes, and the hippocampal region mediates the binding of features that are coactive (e.g., Johnson, 1992). Thus, during encoding, frontal processes presumably both control what is coactive and extend the time (or number of occasions) it is coactive, presenting the hippocampal region with additional opportunities for binding.

Evidence of such a frontal–hippocampal circuit comes from several sources. There is neuroimaging evidence from young adults that prefrontal cortex maintains and manipulates information (e.g., D'Esposito et al., 1998; D'Esposito, Postle, Ballard, & Lease, 1999; Petrides, 1994a, 1994b). Moreover, the hypothesis that prefrontal cortex maintains feature combinations over time is consistent with recent findings from electrophysiological recordings in monkeys showing that some cells in prefrontal cortex code for both object and location (Rao, Rainer, & Miller, 1997) and functional magnetic resonance imaging (fMRI) data from young adults showing greater prefrontal cortex activation (right Brodmann's Area 10) when participants maintained integrated feature information versus a task in which they maintained the features separately (Prabhakaran, Narayanan, Zhao, & Gabrieli, 2000). A recent neuroimaging review suggests that anterior hippocampal activations may be particularly likely when tasks involve relational encoding (Schacter & Wagner, 1999). It is not surprising, then, that there is indirect evidence that age-related neuropathology in frontal and medial-temporal regions is correlated with age deficits in long-term source memory (e.g., Craik, Morris, Morris, & Loewen, 1990; Glisky, Polster, & Routhieaux, 1995; Henkel et al., 1998; Mather, Johnson, & De Leonardis, 1999; Shimamura, 1994). Such findings are consistent with the idea that memory for complex events is supported by a frontal–hippocampal circuit subserving specific, postperceptual, reflective processes (e.g., Johnson, 1992) and suggest that deficits in one or more of the components of this circuit underlie older adults' binding deficits.

We have recently found direct evidence for this proposal using fMRI (Mitchell et al., in press). The procedure was almost identical to that of Experiment 1. The behavioral data replicated that of Experiment 1, and the fMRI data showed areas of left anterior

[3] At test, Luck and Vogel (1997) re-presented the entire study array (which contained several multifeature stimuli, e.g., four colored bars in various orientations). In a lure array, one feature value would be changed (e.g., one of the four bars might be in the same orientation as at study but in a new color). This changed feature was sometimes a repetition of a value presented at study and sometimes a novel value, taken from a larger pool of feature values used to create their stimuli. As with our combination (old + new) test lures, participants could conceivably do well rejecting Luck and Vogel's lures without having bound the feature information at all (e.g., by noticing that the test array contained two instances of some feature value whereas the study array did not, or by noticing a completely novel feature in the test array). Thus, although participants might have been responding based on bound featural information, this procedure is not ideal for assessing binding. We thank Ed Vogel for supplying additional procedural details.

hippocampus and right prefrontal cortex (Brodmann's Area 10) that were differentially recruited in the combination condition relative to the feature conditions in young, but not older, adults. Such a pattern of activations suggests that age-related dysfunction of hippocampal regions, prefrontal regions, or both might help explain older adults' behavioral deficit in the combination condition in this working memory task.

In conclusion, age-related deficits in reflective processes that mediate feature binding are manifested in both working memory (the present experiments) and longer term episodic memory tasks (e.g., Chalfonte & Johnson, 1996). We also found evidence for an age-related working memory test load deficit, suggesting that older adults experience difficulty in accessing multiple features, in evaluating them, or in both. Taken together, such findings are consistent with the idea that there is overlap in the reflective processes that account for performance in working memory tasks and the reflective processes that encode, access, and evaluate information in long-term memory tasks. The degree, and nature, of this overlap remains to be clarified, for example, by using neuroimaging techniques to specify the extent to which common and different brain regions are activated during acquisition and test in working memory and long-term memory tasks using comparable materials.

References

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 89–195). New York: Academic Press.

Babcock, R. L., & Salthouse, T. A. (1990). Effects of increased processing demands on age differences in working memory. *Psychology and Aging, 5,* 421–428.

Baddeley, A. (1986). *Working memory.* Oxford, England: Oxford University Press.

Baddeley, A. (1992, January 31). Working memory. *Science, 255,* 556–559.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47–90). New York: Academic Press.

Bayen, U. J., & Murnane, K. (1996). Aging and the use of perceptual and temporal information in source memory tasks. *Psychology and Aging, 11,* 293–303.

Bayles, K. A., & Kasniak, A. W. (1987). *Communication and cognition in normal aging and dementia.* Boston: Little, Brown.

Belleville, S., Rouleau, N., & Caza, N. (1998). Effect of normal aging on the manipulation of information in working memory. *Memory & Cognition, 26,* 572–583.

Burke, D. M., & Light, L. (1981). Memory and aging: The role of retrieval processes. *Psychological Bulletin, 90,* 513–546.

Cerella, J. (1990). Aging and information-processing rate. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 201–221). San Diego, CA: Academic Press.

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

Clark, S. E., & Gronlund, S. D. (1996). Global matching models of recognition memory: How the models match the data. *Psychonomic Bulletin & Review, 3,* 37–60.

Cohen, J. D., Perlstein, W. M., Braver, T. S., Nystrom, L. E., Noll, D. C., Jonides, J., & Smith, E. E. (1997). Temporal dynamics of brain activation during a working memory task. *Nature, 386,* 604–608.

Cohen, N. J., & Eichenbaum, H. (1993). *Memory, amnesia, and the hippocampal system.* Cambridge, MA: MIT Press.

Craik, F. I. M. (1977). Age differences in human memory. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (pp. 384–420). New York: Van Nostrand Reinhold.

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

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

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

Delis, D. C., Kramer, J., Kaplan, E., & Ober, B. A. (1987). *The California verbal learning test.* San Antonio, TX: Psychological Corporation.

D'Esposito, M., Aguirre, G. K., Zarahn, E. K., Ballard, D., Shin, R. K., & Lease, J. (1998). Functional MRI studies of spatial and nonspatial working memory. *Cognitive Brain Research, 7,* 1–13.

D'Esposito, M., Detre, J. A., Alsop, D. C., Shin, R. K., Atlas, S., & Grossman, M. (1995). The neural basis of the central executive system of working memory. *Nature, 378,* 279–281.

D'Esposito, M., Postle, B. R., Ballard, D., & Lease, J. (1999). Maintenance versus manipulation of information held in working memory: An event-related fMRI study. *Brain and Cognition, 41,* 66–86.

Dobbs, A. R., & Rule, B. G. (1989). Adult age differences in working memory. *Psychology and Aging, 4,* 500–503.

Eichenbaum, H., & Bunsey, M. (1995). On the binding of associations in memory: Clues from studies on the role of the hippocampal region in paired-associate learning. *Current Directions in Psychological Science, 4,* 19–23.

Ferguson, S. A., Hashtroudi, S., & Johnson, M. K. (1992). Age differences in using source-relevant cues. *Psychology and Aging, 7,* 443–452.

Foos, P. W. (1989). Adult age differences in working memory. *Psychology and Aging, 4,* 269–275.

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

Goldman-Rakic, P. S. (1995). Architecture of the prefrontal cortex and the central executive. *Annals of the New York Academy of Sciences, 769,* 71–83.

Grady, C. L., McIntosh, A. R., Horwitz, B., Maisog, J. M., Ungerleider, L. G., Mentis, M. J., Pietrini, P., Schapiro, M. B., & Haxby, J. V. (1995, July 14). Age-related reductions in human recognition memory due to impaired encoding. *Science, 269,* 218–221.

Grégoire, J., & Van der Linden, M. (1997). Effect of age on forward and backward digit spans. *Aging, Neuropsychology, and Cognition, 4,* 140–149.

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General, 108,* 356–388.

Hashtroudi, S., Johnson, M. K., & Chrosniak, L. D. (1990). Aging and qualitative characteristics of memories for perceived and imagined complex events. *Psychology and Aging, 5,* 119–126.

Henke, K., Weber, B., Kneifel, S., Wieser, H. G., & Buck, A. (1999). Human hippocampus associates information in memory. *Proceedings of the National Academy of Sciences, USA, 96,* 5884–5889.

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

Johnson, M. K. (1992). MEM: Mechanisms of recollection. *Journal of Cognitive Neuroscience, 4,* 268–280.

Johnson, M. K., & Chalfonte, B. L. (1994). Binding complex memories: The role of reactivation and the hippocampus. In D. L. Schacter & E.

Tulving (Eds.), *Memory systems 1994* (pp. 311–350). Cambridge, MA: MIT Press.

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

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

Johnson, M. K., & Hirst, W. (1993). MEM: Memory subsystems as processes. In A. F. Collins, S. E. Gathercole, M. A. Conway, & P. E. Morris (Eds.), *Theories of memory* (pp. 241–286). East Sussex, England: Erlbaum.

Johnson, M. K., Raye, C. L., & Reeder, J. A. (2000). *Second thoughts versus second looks: An age-related deficit in reflectively refreshing activated information.* Manuscript in preparation, Princeton University.

Jonides, J., Reuter-Lorenz, P. A., Smith, E. E., Awh, E., Barnes, L. L., Drain, M., Glass, J., Lauber, E. J., Patalano, A. L., & Schumacher, E. H. (1996). Verbal and spatial working memory in humans. In D. L. Medin (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 35, pp. 43–88). San Diego, CA: Academic Press.

Kane, M. J., & Hasher, L. (1996). Interference. In G. Maddox (Ed.), *Encyclopedia of aging: A comprehensive resource in gerontology and geriatrics* (2nd ed., pp. 514–516). New York: Springer.

Kausler, D. H. (1994). *Learning and memory in normal aging.* San Diego, CA: Academic Press.

Kausler, D. H., & Puckett, J. M. (1981). Adult age differences in memory for sex of voice. *Journal of Gerontology, 36,* 44–50.

Kroll, N. E. A., Knight, R. T., Metcalfe, J., Wolf, E. S., & Tulving, E. (1996). Cohesion failure as a source of memory illusions. *Journal of Memory and Language, 35,* 176–196.

Light, L. L., & Zelinski, E. M. (1983). Memory for spatial information in young and older adults. *Developmental Psychology, 19,* 901–906.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390,* 279–281.

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide.* Cambridge, England: Cambridge University Press.

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

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

Mitchell, K. J., Johnson, M. K., Raye, C. L., & D'Esposito, M. (1999). [More on age-related binding deficits in working memory]. Unpublished raw data.

Mitchell, K. J., Johnson, M. K., Raye, C. L., & D'Esposito, M. (in press). *fMRI evidence of age-related hippocampal dysfunction in feature binding in working memory. Cognitive Brain Research.*

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

Myerson, J., Hale, S., Wagstaff, D., Poon, L. W., & Smith, G. A. (1990). The information-loss model: A mathematical theory of age-related cognitive slowing. *Psychological Review, 97,* 475–487.

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

Park, D. C., & Puglisi, J. T. (1985). Older adults' memory for the color of pictures and words. *Journal of Gerontology, 40,* 198–204.

Petrides, M. (1994a). Frontal lobes and behavior. *Current Opinion in Neurobiology, 4,* 207–211.

Petrides, M. (1994b). Frontal lobes and working memory: Evidence from investigations of the effects of cortical excisions in nonhuman primates. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (Vol. 9, pp. 59–82). Amsterdam: Elsevier Science.

Prabhakaran, V., Narayanan, K., Zhao, Z., & Gabrieli, J. D. E. (2000). Integration of diverse information in working memory within the frontal lobe. *Nature Neuroscience, 3,* 85–90.

Rabinowitz, J. C. (1989). Judgements of origin and generation effects: Comparisons between young and elderly adults. *Psychology and Aging, 4,* 259–268.

Rao, S. C., Rainer, G., & Miller, E. K. (1997, May 2). Integration of what and where in the primate prefrontal cortex. *Science, 276,* 821–824.

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.), *Handbook of aging and cognition* (Vol. 2, pp. 1–90). Mahwah, NJ: Erlbaum.

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

Salthouse, T. A. (1988). Initiating the formalization of theories of cognitive aging. *Psychology and Aging, 3,* 3–16.

Salthouse, T. A. (1990). Working memory as a processing resource in cognitive aging. *Developmental Review, 10,* 101–124.

Schacter, D. L., Kaszniak, A. W., Kihlstrom, J. F., & Valdiserri, M. (1991). The relation between source memory and aging. *Psychology and Aging, 6,* 559–568.

Schacter, D. L., & Wagner, A. D. (1999). Medial temporal lobe activations in fMRI and PET studies of episodic encoding and retrieval. *Hippocampus, 9,* 7–24.

Shimamura, A. P. (1994). Neuropsychological perspectives on memory and cognitive decline in normal human aging. *Seminars in Neurosciences, 6,* 387–394.

Smith, E. E., Jonides, J., & Koeppe, R. A. (1996). Dissociating verbal and spatial working memory using PET. *Cerebral Cortex, 6,* 11–20.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 174–215.

Spencer, W. D., & Raz, N. (1995). Differential effects of aging on memory for content and context: A meta-analysis. *Psychology and Aging, 10,* 527–539.

Squire, L. R. (1995). Biological foundations of accuracy and inaccuracy in memory. In D. L. Schacter (Ed.), *Memory distortion: How minds, brains, and societies reconstruct the past* (pp. 197–225). Cambridge, MA: Harvard University Press.

Stuss, D. T., Eskes, G. A., & Foster, J. K. (1994). Experimental neuropsychological studies of frontal lobe functions. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (Vol. 9, pp. 149–185). Amsterdam: Elsevier Science.

Uttl, B., Cosentino, S., & Graf, P. (1999, June). *Fatigue effects: Do older adults tire more quickly?* Paper presented at the meeting of the American Psychological Society, Denver, CO.

Verfaellie, M., & Treadwell, J. R. (1993). Status of recognition memory in amnesia. *Neuropsychology, 4,* 5–13.

Wechsler, D. (1987). *Wechsler Memory Scale—Revised manual.* New York: Psychological Corporation.

Received November 24, 1998
Revision received January 19, 2000
Accepted January 21, 2000 ■