Journal of Memory and Language **44,** 399–426 (2001)
doi:10.1006/jmla.2000.2734, available online at http://www.academicpress.com on IDEAL®

Dual Task Demands and Gist-Based False Recognition of Pictures in Younger and Older Adults

Wilma Koutstaal, Daniel L. Schacter, and Carolyn Brenner

Harvard University

In a yes/no recognition paradigm using categorized pictures (Koutstaal & Schacter, 1997), older adults show high rates of false recognition of category items where many related items are studied; they also show high levels of veridical recognition of targets from such categories (where gist-like memory representations might be sufficient) but impaired recognition of one-of-a-kind items (where item-specific memory may be required). Dual task demands at study were used to equate older and younger adults on veridical memory for one-of-a-kind items, but older adults still showed elevated false recognition. When we compared young adults under dual task conditions to a young control group, dual task performance at study, or at both study-and-test, substantially reduced veridical memory but did not reduce false recognition. Dual task demands at test also did not affect false recognition. Gist-based false recognition of pictures is robust to changes in encoding resources that exert substantial effects on veridical memory.

Key Words: memory illusions; recognition; retrieval; encoding; resources; division of attention; age differences.

The amount of attention that we devote to a given task admits of many degrees or gradations. We can perform a task with marked "singleness of purpose," devoting virtually all of the attentional, perceptual, and general cognitive–evaluative resources that we can muster to the task at hand. We can also perform a task under varying degrees of distraction or with competing purposes, attempting to negotiate between and accommodate the demands of more than one task or goal concurrently. Laboratory analogs of these two contrasting conditions—created by requiring participants to perform only a single task or to concurrently perform two tasks—have been used to explore many aspects of sensory, motor, and cognitive processing (Bourke, Duncan, & Nimmo-Smith, 1996; Pashler, 1994). Dual task methodology has been particularly fruitful in probing the role of attentional processing in long-term memory.

This research was supported by National Institute on Aging Grant AG08441 and the Human Frontiers Science Program (Grant RG0126). We thank Sara Greene for help with participant testing.

Address correspondence and reprint requests to Wilma Koutstaal, whose current address is Department of Psychology, University of Reading, Earley Gate, Whiteknights, Reading, RG6 6AL, United Kingdom. E-mail: w.koutstaal@reading.ac.uk.

Perhaps the most straightforward application of this methodology has been to examine the costs of limited attentional resources for *veridical* memory. What are the consequences for our later ability to recall or recognize an event, object, or other stimulus if—at the time of our initial experience of it—we can accord it only limited attention because we are (simultaneously) attempting to meet the requirements of another task? Are similar effects found if the additional task is imposed not at the time of encoding but at the time of attempted *retrieval*—during our efforts to recall or recognize the stimulus? Here, interest focuses on the effects of varying attentional demands on memory *accessibility*, the likelihood that target events that were experienced will be successfully recalled or recognized or, stated in terms of errors, the probability of negative errors or "errors of omission" ("misses"). A key finding from investigations of the effects of dual task performance on memory accessibility is that whereas dual task performance during initial encoding often leads to substantial decreases in memory (e.g., Baddeley, Lewis, Eldridge, & Thomson, 1984; Craik, Govoni, Naveh-Benjamin, & Anderson, 1996; Mulligan, 1998), dual task performance during retrieval typically yields no or comparatively

0749-596X/01 $35.00

smaller decreases in memory (e.g., Baddeley et al., 1984; Craik et al., 1996; Iidaka, Anderson, Kapur, Cabeza, Okamoto, & Craik, 1999; Naveh-Benjamin, Craik, Guez, & Dori, 1998; Nyberg, Nilsson, Olofsson, & Backman, 1997) although this "protection of retrieval" may be accompanied by decrements in performance on the secondary task (e.g., Anderson, Craik, & Naveh-Benjamin, 1998; Craik et al., 1996; Naveh-Benjamin et al., 1998), particularly for free recall and cued recall.

Dual task conditions also may be used to explore the effects of varying attentional demands on memory *accuracy* (Koriat & Goldsmith, 1996; Koutstaal & Schacter, 1997a) or the likelihood of positive errors or "errors of commission." Although human memory is often accurate or "true" to both the general outlines of the past and to more specific features of particular episodes, it is also—and equally importantly—prone to error, with errors including not only failures to access information about the past but also various forms of memory distortions and misattributions (Estes, 1997; Johnson, Hashtroudi, & Lindsay, 1993; Roediger, 1996; Schacter, 1995, 1999; Schacter, Norman, & Koutstaal, 1998). For example, the different characteristics of a complex event may be misconjoined or misaligned, as when we mistakenly ascribe a particular event that did occur to an incorrect location or time (Johnson et al., 1993), or we may confuse "similar-seeming" but not previously encountered objects (people, names, pictures, words) with objects that we did, in fact, experience. Are such "positive" errors *also* more likely under dual-task compared to single-task conditions? Here, dual task conditions during encoding and/or attempted retrieval have been thought to increase the likelihood that participants will rely on a general sense of familiarity, or comparatively broad undifferentiated assessments of the similarity of tested items to previously encountered items, rather than more detailed or specific forms of "recollection" (e.g., Jacoby, 1991; Jacoby, Toth, & Yonelinas, 1993), and generally these conditions have been found to increase "errors of commission." For example, higher levels of false recognition of semantically related lures have been found under dual task conditions at study than under a single task (Mandler & Worden, 1973; but also see Seamon, Luo, & Gallo, 1998), and performance of a secondary task at test has been associated with an increased likelihood of misattribution errors (Jacoby, Woloshyn, & Kelley, 1989; Jacoby, 1999b). The imposition of dual task requirements during encoding or retrieval among healthy young individuals also allows exploration of across-population differences: Does dual task performance lead younger adults to show patterns of memory performance (either in terms of accessibility, or in terms of accuracy) that are similar to those found in other populations, such as the memory deficits that may be associated with normal cognitive aging (e.g., Anderson et al., 1998; Craik, 1982; Craik & Byrd, 1982; Whiting & Smith, 1997)?

The present study, which examines both veridical and false recognition in older and younger adults under conditions where younger adults were exposed to dual task demands at encoding, retrieval, or both encoding and retrieval, relates to each of these issues. More specifically, the experiment reported here addresses three primary questions. The first focuses on false recognition in older and younger adults but uses dual task methodology to equate the two age groups on veridical recognition in certain conditions. The second and third also focus on false recognition but compare the performance of younger adults under specific dual-task versus single-task conditions. The motivation for these questions and relevant findings from prior studies are outlined below.

The Matching Question

Several studies have shown that older adults (individuals aged 60 years and above) may be more susceptible to false recall and false recognition than are younger adults (e.g., Rankin & Kausler, 1979; Smith 1975; for review see Schacter, Koutstaal, & Norman, 1998). For example, evidence from conditions where participants are asked to intentionally learn semantically related lists of words, where each of the words in the list (e.g., *bed*, *rest*, *awake*, *tired*, *dream*, etc.) is an associate of a nonpresented "critical" theme or lure word (e.g., *sleep;* Deese, 1959; Roediger & McDermott, 1995), suggests that older adults are more likely to incorrectly produce or "intrude" the semantically related theme words during free recall than are younger adults (Kensinger & Schacter, 1999; Norman & Schacter, 1997; Tun, Wingfield, Rosen, & Blanchard, 1998). Older adults may also show higher levels of false recognition of the semantically related lures (Norman & Schacter, 1997; Schacter, Israel, & Racine, 1999; Tun et al., 1998), including lures that are only weakly associated to the study list theme (Tun et al., 1998), and may produce their false recognition responses more rapidly (Tun et al., 1998) than do younger adults.

Age-related differences in false recognition have been found to be especially pronounced, however, for pictorial stimuli. Using a categorized pictures paradigm (Koutstaal & Schacter, 1997b) in which older and younger participants were first shown different numbers of objects from various object categories (e.g., 1, 9, or 18 different exemplars of cats, shelves, or teddy bears) and then later were asked if they recognized those items and new items that were from those same categories, we found a clear effect of category size on false recognition. Both older and younger adults were more likely to falsely claim to recognize lure items as an increasing number of categorically similar items were studied. (For similar effects of category size on false recognition of semantically related words, and abstract patterns, see, respectively, Arndt & Hirshman, 1998; Robinson & Roediger, 1997; Shiffrin, Huber, & Marinelli, 1995; and Homa, Cross, Cornell, Goldman, & Schwartz, 1973; Omohundro, 1981.) However, the rate of false recognition in the many-exemplar categories was strongly affected by the age of participants: Whereas for the largest categories, younger adults (between the ages of 18 and 35 years) showed false recognition rates of 25–35%, the false recognition rates of older adults (aged 60 to 75 years) were considerably greater, ranging between 60 and 70%. Moreover, these age differences in false recognition of categorically related objects proved to be remarkably consistent. Elevated false recognition among older compared to younger adults was observed not only across three experiments that used a simple incidental encoding task (participants were asked to indicate how much they liked each picture) and an old/new recognition test (Koutstaal & Schacter, 1997b) but also in follow-up experiments where the conditions at encoding and/or retrieval were modified in an attempt to reduce false recognition (Koutstaal, Schacter, Galluccio, & Stofer, 1999). Age-related increases in false recognition remained both when participants were explicitly encouraged to attend to differentiating or distinctive perceptual features of the objects during study and when careful "item-by-item" monitoring was required during testing. (In the latter condition, participants were asked to separately designate items as "old and identical" to items they had studied, "new but related" to items they had studied, or "new and unrelated" to the studied items.) In each case, although older adults clearly benefitted from additional encoding or retrieval support (as shown by comparison with older controls who were not given such support), their rate of false recognition remained elevated above that for younger adults. Elevated false recognition among older adults was also observed when encouragement to notice differentiating features was provided at both encoding and retrieval.

Critically, the finding of increased false recognition of categorized items in older adults was also accompanied by a marked, and repeatedly observed, divergent pattern in *veridical* recognition for the two age groups. Whereas older and younger adults showed essentially equivalent levels of veridical recognition for items from categories where many exemplars had been presented, older adults showed substantially reduced veridical recognition for items where only *one* or *no* categorically related exemplars had been presented. The combination of these three results—(a) equivalent hits by older and younger adults for studied items from large categories, (b) elevated false recognition among older adults for these same (large) categories, and (c) depressed correct recognition in older adults for one-of-a-kind items—suggested that older adults were relying on knowledge concerning the general kinds of items they had studied to a greater extent than were younger adults. All three findings would be explained if older adults were especially influenced by the general perceptual or conceptual similarities of the items they had encountered, what has been called “gist” (Brainerd, Reyna, & Kneer, 1995; Reyna & Brainerd, 1995) or general-similarity information (Curran, Schacter, Norman, & Galluccio, 1997; Hintzman, 1988; Hintzman & Curran, 1994, 1995). If, compared to their younger counterparts, older adults were especially relying on their knowledge of the general types or categories of objects they had studied to make recognition decisions, then their recognition of target items from categories where many related items were studied should be particularly aided because these categories were quite salient at study; further, if younger adults did not rely on categorical information to the same extent, this might allow the hit rates of older adults to approach or equal those of younger adults for these categories. However, reliance on such general similarities would not allow accurate differentiation between actually studied items (targets) and nonstudied but categorically related items (related lures or distractors), thus resulting in higher false recognition for older than younger adults for large category lures. Finally, memory for the categorical nature of the items would also be expected to be less strong in instances where only a single item from a category had been studied, so reliance on gist-based representations here might more often fail to support correct recognition for older adults.

The clear discrepancy in older adults’ veridical recognition for items from many-exemplar categories (not impaired relative to the young) versus one-of-a-kind items (substantially impaired) suggests that a key factor contributing to the false recognition of older adults may be decreased or degraded *item-specific memory* (cf. Rabinowitz & Ackerman, 1982; Rabinowitz, Craik, & Ackerman, 1982; also see Hay & Jacoby, 1999; Hess, 1985; Isingrini, Fontaine, Taconnat, & Duportal, 1995; Mäntylä & Bäckman, 1992; Rankin & Kausler, 1979; Schacter, Koutstaal, & Norman, 1997). If older adults possess less item-specific information, they may be less able to differentiate previously presented target items from categorically and perceptually similar lures, so they may show increased false recognition. This, however, raises the question: If older and younger adults had less radically divergent levels of memory for such “one-of-a-kind” items—suggesting they possessed similar levels of item-specific memory—would the difference in the false recognition rates of older and younger adults also be reduced, or possibly even eliminated? Or, if younger adults’ veridical recognition of one-of-a-kind items was somehow reduced to match that of older adults, would younger adults *also* show higher rates of similarity-based false recognition, rates more similar to those shown by older individuals?

We examined this question using a “dual-task” paradigm. In the experimental conditions, younger participants were presented categorized pictures and, simultaneously, were asked to perform a digit-shadowing and monitoring task. In this task, participants heard a continuous random series of digits and were asked both to orally repeat the digits immediately after they were presented and also to monitor the digits for sequences of “odd” digits, signaling each time there was a consecutive sequence of three or more odd digits in a row (cf. Craik, 1982; Jennings & Jacoby, 1993). Some younger participants performed the dual task only during the encoding phase, that is, during the initial presentation of the pictures (the “study-only” group); other participants performed the dual task only during retrieval, during yes/no recognition testing (the “test-only” group), and still others performed the dual task during both encoding and retrieval (the “study-and-test” group). In addition, to allow within-experiment comparisons of the effects of the dual task requirements, we included both an older control group and a younger control group; participants in these control conditions initially studied the pictures and were later tested under circumstances identical to those of the dual-task groups, but without the requirement to perform the digit shadowing and monitoring task at either study or test.

As noted above, previous studies using a dual-task methodology have found that veridical recall and recognition are often substantially impaired by the imposition of dual-task demands at the time of *study* or encoding (Baddeley et al., 1984; Craik et al., 1996; Park, Puglisi, Smith, & Dudley, 1987; Park, Smith, Dudley, &

Lafronza, 1989; Seamon et al., 1998; also cf. Mulligan, 1998; Schmitter-Edgecombe, 1996). Impairments in memory have, however, been less pronounced if the dual task is imposed at the time of retention testing or retrieval (Baddeley et al., 1984; Craik et al., 1996; Naveh-Benjamin et al., 1998; Nyberg et al., 1997), with either relatively smaller memory costs observed than for dual task at encoding (in the case of free or cued recall; cf. Park et al., 1989) or almost no memory costs observed (in the case of yes/no recognition). Given these findings, we anticipated that younger adults would most likely show decreased veridical recognition, including decreased correct recognition of the one-of-a-kind items, when dual task demands were present at study. The effects of such decreases in veridical recognition on *false recognition* could then be examined, with false recognition shown by younger adults in these conditions contrasted with that shown by older controls. If differences in false recognition of categorically related items in older vs younger adults are substantially driven by correlated differences in item-specific memory (as indexed by veridical memory for one-of-a-kind items) then, when item-specific memory in the two age groups is matched, the magnitude of the age difference in false recognition should be reduced. Alternatively, if age-related differences in false recognition are not primarily attributable to differences in item-specific memory (again, as indexed by veridical memory for one-of-a-kind items) substantial age-related differences should still be observed.

The "Preservation" of False Recognition Question

Although older adults show especially high levels of false recognition in the categorized pictures paradigm, the rates of false recognition shown by younger adults are also fairly substantial, with false recognition of categorized lures from the larger categories exceeding baseline rates of false alarms to novel categories by as much as 20 to 30%. What underlies these responses? One possibility is that false recognition responses for younger adults, like those for older adults, reflect general similarity responding—responding based on a general sense of familiarity with the categories or awareness of the general types of items that were studied. If the false recognition responses of younger adults derive from a more global sense of familiarity with the categorized items (e.g., "I saw many pictures of cats"), such familiarity might be acquired even under conditions of less complete or extensive encoding of specific details (cf. Mandler & Worden, 1973; also see Seamon et al., 1998). More specifically, although limiting processing resources at encoding for younger adults should depress veridical recognition (dependent on item-specific recollection and possibly also more generic familiarity), if false recognition depends on "gist-like" general similarity information, then false recognition should be largely "preserved" (or possibly somewhat increased) despite decrements in veridical recognition. Such an outcome would be broadly consistent with the findings of earlier studies, where an increased likelihood of memory misattributions such as "false fame" errors under conditions involving divided attention at study have been attributed to participants' decreased recollection but continued likelihood of responding on the basis of general familiarity (Jacoby et al., 1989; Jennings & Jacoby, 1993). This outcome would also cohere with evidence that dual task requirements at study primarily lower estimates of recollection (Jacoby et al., 1993) or subjective reports of "remembering" (Gardiner & Parkin, 1990) rather than estimates of "familiarity" or subjective reports of "knowing" in the absence of recalled episodic details.

However, there is another alternative (briefly outlined here and more fully evaluated in the Discussion). Given the comparatively high levels of item-specific memory shown by younger adults for the categorized pictures (including the one-of-a-kind items), it is possible that their false recognition responses are based on a form of "misplaced recollection" where highly specific features that are similar to features of the studied items lead to the mistaken recognition of the lure items. That is, it is possible that some of the categorized lure items share specific features with actually studied items and that these specific features provide a basis for the mistaken judgment that an item has been encountered previously (e.g., “I saw a large tawny-colored cat that looked very much like this particular cat”; cf. Koutstaal & Schacter, 1997b; also see Holmes, Waters, & Rajaram, 1998). On the one hand, to the extent that errors of this form depend on the initial encoding and retention of specific details about the studied items, decreasing the opportunity to encode item-specific information that supports veridical recognition (and particularly recollection, cf. Gardiner & Parkin, 1990; Jacoby et al., 1993) might also decrease *false* recognition. On the other hand, however, it is also possible that dual task conditions will continue to allow the extraction of relatively detailed feature information but will decrease the likelihood that those features will be successfully integrated or “bound” together with one another in a unified trace (e.g., Johnson & Chalfonte, 1994; Schacter, Norman, & Koutstaal, 1998). If so, dual task demands at study might *also* lead to an increased likelihood of “misplaced recollection” by increasing the likelihood of inappropriate binding of features and objects (cf. Reinitz, Morrissey, & Demb, 1994). On a gist-based account, then, false recognition should be maintained, or possibly increased, under dual task conditions at study. On a misplaced recollection account, depending on how the dual task requirements during encoding affect the extraction of features and the probability of successful binding, the dual task may decrease false recognition (if fewer features are extracted and retained, resulting in fewer misalignments of features and objects) *or* may maintain or increase false recognition (if specific features are still extracted, and the probability of misconjoining features remains constant or is increased by the dual task relative to single task requirements).

The extent to which the false recognition and veridical recognition responses of younger adults are *differentially affected* by reduced attentional and other processing at study can be examined by comparing the effects of dual task demands on veridical recognition versus false recognition, contrasting the rates of true versus false recognition observed when a secondary task is performed at study with the control condition. If dual task demands during study detract more from veridical recognition than from false recognition—as expected on a “gist-based” account of false recognition and on at least one version of a “misplaced recollection” account—then there should be an interaction between response type (true vs false recognition) and group (study-only vs control). If, however, dual task demands reduce the number of specific features that are retained without altering the probability of successful binding of those features, then according to the misplaced recollection account, both veridical and false recognition should be reduced.

The Test Monitoring Question

Although a number of studies have shown that, for veridical memory, dual task demands at study exert a larger detrimental effect on memory accessibility than do dual task demands at test, relatively fewer studies have examined the effects of a requirement to perform a dual task during retrieval on illusory or *false* memory. To the extent that deliberate or effortful processing is required to differentiate studied items from perceptually and categorically similar but nonstudied lures, decreasing the availability of cognitive or attentional resources at *test* may increase false recognition. Indeed, consistent with this, Jacoby, Woloshyn, and Kelley (1989, Exp. 3) found that college students showed greater susceptibility to familiarity-based errors (false fame errors) under conditions of divided attention at test than they did without such additional demands at test (also, cf. Moscovitch, 1994). Likewise consistent with this possibility, Merikle and Joordens (1997) found that younger adults placed under dual task requirements during recognition testing were more likely than those under single task conditions to mistakenly designate as “old” new items that were briefly “pre-exposed” prior to their full presentation in a yes/no recognition trial, possibly indicating misattribution of the test-derived familiarity to the earlier study phase. Other work has shown that younger adults may show greater errors in source memory judgments under dual task conditions at retrieval, with the source memory accuracy among younger adults under these conditions decreased to a level similar to that found for older adults (Dywan, Segalowitz, & Webster, 1998). There is also evidence from experiments using Jacoby's (1991) process-dissociation procedure that dividing attention at test may selectively reduce the "recollection" (as opposed to "familiarity") component of participants' responses, especially impairing the retrieval (Jacoby, Yonelinas, & Jennings, 1996) or controlled use (Gruppuso, Lindsay, & Kelley, 1997) of the necessary source-specifying information that would allow the required exclusion of "to-be-excluded" study items and inclusion of other study items.

In the present study, examination of the rates of false recognition when dual-task demands were present at test in comparison to when no demands were present during recognition testing thus allows assessment of the costs of "interference with test-monitoring"—particularly monitoring that may be dependent on the "recollection-like" retrieval and use of item-specific information. This question is also of particular interest because, as noted above, in previous work (Koutstaal et al., 1999a) we found that older adults especially benefited from the provision of retrieval monitoring support. Do younger adults, under conditions designed to make retrieval monitoring especially *difficult*, show higher rates of similarity-based false recognition errors?

We examined each of these three questions—the effect of first matching older and younger adults on veridical recognition of one-of-a-kind single items on age-related rates of false recognition of categorically related lures, the comparative "preservation" of false recognition relative to veridical recognition under dual task conditions at study vs control conditions in younger adults, and the effects of interference with test monitoring on false recognition of younger adults—using the dual-task procedure and categorized pictures paradigm described above.

METHOD

Experimental Design

The initial experiment included five between subjects groups, including an older control group, a younger control group, and three younger dual-task groups for whom the dual task was present at study-only, test-only, or both study-and-test. There was also a within subjects factor of category size. For studied items, category size had three levels: 1, 9, or 18 categorically related pictures were presented (also termed single, medium, and large categories, respectively). For nonstudied items, category size had four levels, the three just stated, plus novel categories, where 0 related items were presented during study. These "novel" category items provided an estimation of the baseline level of false alarms. In addition, several further "unrelated" items (miscellaneous, one-of-a-kind items that did not belong to any study categories) were also included (see Stimuli, below).

Because the initial experimental conditions resulted in near, but not exact, matching of younger and older adults' veridical recognition in the critical "one-of-a-kind" single condition, a further matching condition was included. In this condition, the dual task was again presented at study-only, but, in addition, the presentation duration of the stimuli during study was decreased from 2,000 to 500 ms per item (i.e., participants saw each picture for only 500 ms). To allow evaluation of the effects of the more rapid stimulus presentation, independent of the dual task, we also included a younger control group who were shown the stimulus items under the faster presentation rate, without any dual task requirements. The results for these conditions (referred to as fast study-only and fast control, respectively) are presented combined with those of the initial experiment.

Participants

The older control participants (n = 16, mean age = 67.2 years, range = 60–74) were recruited through newspaper advertisements and posted flyers, and they were screened for various medical and neuropsychological conditions, including a history of alcoholism or substance abuse, present or previous treatment for psychiatric illness, current treatment with psychoactive medication, drug toxicity, primary degenerative brain disorders, and brain damage sustained earlier from a known cause. Older controls had, on average, 15.6 years of formal education (range = 12–21 years). Younger participants in both the initial experiment (n = 64, mean age = 21.2 years, range = 17–34; 16 per condition) and the fast study-only and fast control conditions (n = 32, mean age = 21.1 years, range = 18–27; 16 per condition) were recruited through sign-up sheets posted at Harvard University and were screened for current use of psychoactive medications and depression. They had, on average, approximately 15 years of formal education (mean for the initial experiment = 14.6 years, mean for the fast study-only and fast control conditions = 15.1, overall range = 12–22 years; educational information for 6 individuals not available). Younger participants in the dual-task conditions were included only if they attained acceptable levels of accuracy on the digit-monitoring task; participants who did not meet a criterion level of 70% were excluded and replaced (see Results, Digit Monitoring Task). All participants were native speakers of English and had normal or corrected-to-normal vision. They were paid for their involvement in the experiment.

Stimuli

The stimuli were detailed, colored pictures of individual objects (or, in a few cases, coherent groupings of objects). The stimulus set was identical to that used in an earlier experiment (Koutstaal & Schacter, 1997b, Exp. 3) and consisted of categorized items from 20 different object categories. There were 21 exemplars of each category, with category size manipulated by systematically excluding some of the items from the category. The categories were assigned to four sets of five categories each (P, Q, R, and S), and these category sets were rotated through the four category size conditions (0, 1, 9, or 18 related items presented at study). In addition, to provide a further estimate of the level of veridical recognition of one-of-a-kind items—beyond that obtained for the single-item categories, where only 5 target items could be tested for each participant—we also included miscellaneous noncategorized "unrelated" items (e.g., a globe and a harp). Across participants, these unrelated items were counterbalanced across studied vs nonstudied status, with 15 unrelated items presented as targets and 15 as nonstudied lures.

In order to avoid confounding the number of categorized items presented at study with the number of items presented at test, only a subset of the presented target items from each category size was included on the recognition test. Specifically, 3 old items per category for the medium and large categories and 1 item for each of the single item categories were tested. In addition, 3 new related items per category were tested (or in the case of single-item categories, 1 new item per category), as well as 3 new items from each of the novel categories. The particular items within each category that served as targets or lures were initially determined randomly; these "critical" items were then assigned to two subsets (Sets A and B for the medium, large, and novel conditions; Sets A1 and B1 for the single condition) and were systematically counterbalanced across participants. The stimuli were assigned to the study and test lists so that items from the various categories and conditions were distributed throughout the lists, and so that no more than 2 items from any one category ever occurred consecutively. In addition, the test lists were constructed so that no more than 3 previously studied items or new items occurred consecutively.

In all, the study list comprised 215 items, including 5 single items (1 item from each of 5 categories), 45 medium-category items (9 items from each of 5 categories), 90 large-category items (18 items from each of 5 categories), 15 unrelated items, 54 filler items, and 6 buffer items. The filler items were included to increase the variety and length of the study list and were not scored. The test list comprised 115 items, including 50 targets and 65 lures (5 each for single old and single new, and 15 items each for medium old and new, large old and new, novel new, and unrelated old and new). A counterbalancing required 16 subjects, so each of the 16 participants in a given task condition received a different study and test list.

Procedure

All participants were tested individually, in two experimental sessions (the study and test phases) separated by a retention interval of 3 days.

Participants in all task conditions were told that they would be shown detailed colored pictures of objects, one at a time, on the computer screen, and that they would be asked to indicate if they did, or did not, like the pictures. Participants were told they would first be shown each picture (for 2 s in the initial experiment, or for 500 ms in the fast study-only and fast control conditions), followed by a 4-s prompt, and they were requested to enter their liking rating before the screen automatically advanced to the next picture. (To maintain a constant overall study phase duration for all conditions, the picture presentation in the fast study-only and fast control conditions was followed by a 1,500-ms "blank" interstimulus interval before the liking ratings prompt appeared.) Participants were instructed to make their liking ratings on the basis of the particular object that was shown, rather than the general class or type of object shown.

Participants in the conditions involving the dual task during the study phase (i.e., the study-only, study-and-test, and fast study-only groups) were then further instructed as to the nature of the digit-shadowing and monitoring task. They were told that they would also be asked to perform a second task, described as follows:

> In this task, you will be asked to listen to a series of digits. The digits will be presented one at a time, in a random order. Your task is to repeat each of the digits out loud. In addition, as each digit occurs, you should mentally determine if it is an odd number or an even number. Then, any time that three odd numbers occur consecutively (e.g., 1, 7, 9; 9, 9, 3; or 7, 7, 7) you should say the word "NOW."
>
> You will be asked to perform Task One [the picture liking rating task] and Task Two [the digit monitoring task] *at the same time*, performing each as *accurately* and *quickly* as you can. It is very important that you repeat each digit as it occurs, and also carefully monitor the digits for any occurrences of three odd numbers in a row while also performing the picture liking task as carefully and accurately as possible. I will record your responses.
>
> The digits will be played over a headset. I will first begin an audiotape recording of the digits, and you should immediately begin shadowing the numbers and monitoring for occurrences of "3-odds in a row." Very shortly afterwards, I will begin the presentation of the pictures. At the outset of the pictures, the same set of 3 pictures will appear several times. The purpose of this is to allow you time to become accustomed to performing the two tasks together. Then (without any break in between) new pictures will begin to be shown.

Largely similar instructions were given for participants who also, or instead, were given the dual task at test (i.e., the test-only and the study-and-test groups). In the test phase, the primary instructions for participants were as follows:

> I will again show you pictures on the computer screen, one at a time, just as we did on your earlier visit. Now, however, for each of the pictures we'd like you to indicate whether you think the picture is OLD, that is, was one of the pictures that you saw earlier, or is NEW, that is, was never previously presented on your earlier visit.
>
> The way the presentation will work is this. A picture will be shown for two seconds. After two seconds, the picture will disappear and a "New/Old?" prompt will appear. Press the key labeled "N" if you think the picture is "new"—it was not shown on your earlier visit. Press the key labeled "O" if you think the picture is old—you saw it on your earlier visit. You will have 4 seconds in which to make your response. After 4 seconds, the computer will automatically advance to the next picture. Please be sure to answer Old or New for every item.

The audiotape of random numbers was recorded from separate digitized recordings of the digits 1 through 9. The digits were recorded in a female voice and then were repeatedly and pseudo-randomly sampled according to a predetermined list, with the timing between items determined by a PsyScope script. Different portions of the list were approximately equated for the frequency of occurrence of 3 consecutive odd digits, with target trials (that is, trials where participants were required to respond “now”) occurring on just over 8% of the trials. Based on pilot data, the digits were presented with an interstimulus interval of 500 ms. This was a rate at which acceptable levels of accuracy were maintained, but that was sufficiently challenging to also be associated with some cost to the memory task. Both the participant and the experimenter listened to the digit recording via headphones; the experimenter recorded participants’ responses on a preprinted sheet that also listed the digits that were presented. Participants in the study-and-test condition monitored different sequences of random numbers at study and test.

RESULTS

Given the relatively large number of experimental conditions, we focused our analyses on the three primary questions of interest outlined in the Introduction, the matching question, the preservation of false recognition question, and the interference with test monitoring question. The average level of true and false recognition for each of the seven groups is shown in Fig. 1, separately as a function of category size. All analyses were conducted on true recognition and false recognition after correction for baseline levels of false alarms (i.e., false alarms to novel category items). False alarms to the novel category items were generally fairly infrequent, although somewhat higher for the older control group, and also the younger fast study-only group, than for the other conditions [means for older control, younger control, and younger study-only, test-only, and study-and-test conditions of 0.15, 0.09, 0.11, 0.10, and 0.08, respectively, $F < 1$ for an overall analysis of the five groups; $F(1, 78) = 3.01$, $MSe = 0.02$, $p = .09$ for the comparison of older controls versus the average of the four younger groups; means for younger fast study-only and fast control of 0.11 and 0.05, respectively, $F < 1$ for the comparison of older control and younger fast study-only; $F(1, 30) = 3.77$, $MSe = 0.02$, $p = .06$ for older control vs fast control]. Table 1 presents the true and false recognition proportions after correction for novel false alarms (designated as “novel-corrected” true and false recognition, respectively).

Digit-Monitoring Task

In order to ensure that participants were, indeed, attempting to conscientiously perform both the study or test task and the digit-monitoring task, accuracy on the digit-monitoring task was assessed, and participants who showed a target detection accuracy rate of less than 70% were excluded and replaced. A total of 9 participants in the initial experiment did not attain criterion on this task and were replaced (number of individuals excluded, by condition, for study-only, test-only, and study-and-test = 2, 4, and 3, respectively). The average target detection accuracy rates (*SD* in parentheses) for the included participants were, for study-only 89% (0.09), for test-only 91% (0.07), and, for study-and-test 88% (0.09) at study and 85% (0.08) at test. Only one participant in the fast study-only condition failed to meet the accuracy criterion. The average target-detection accuracy rate for the fast study-only condition was 86% ($SD = 0.08$).

The Matching Question

We first considered the patterns of performance shown by the older and younger *control* groups. Consistent with our earlier findings (Koutstaal & Schacter, 1997b), older and younger controls showed similar levels of veridical recognition for the categorized items when a large number of categorically related exemplars had been presented, novel-corrected recognition means of 0.66 and 0.72 respectively, $F < 1$ for the effect of age. Also as in previous experiments, in the absence of any additional task requirements, older controls showed significantly *impaired* correct recognition of single “one-of-a-kind” items, novel-corrected recognition of 0.53 and 0.77 respectively, $F(1, 30) = 6.98$, $MSe = 0.07$, $p = .01$. A similar age-related deficit in veridical recognition of one-of-a-kind items was found for the unrelated items, corrected recognition (i.e., unrelated hits–unrelated false alarms) of 0.43 and 0.81, respectively, $F(1, 30) = 24.91$, $MSe = 0.05$, $p < .0001$. Performance for the medium (9-item) categories was intermediate between that for the large and one-of-a-kind categories, with older

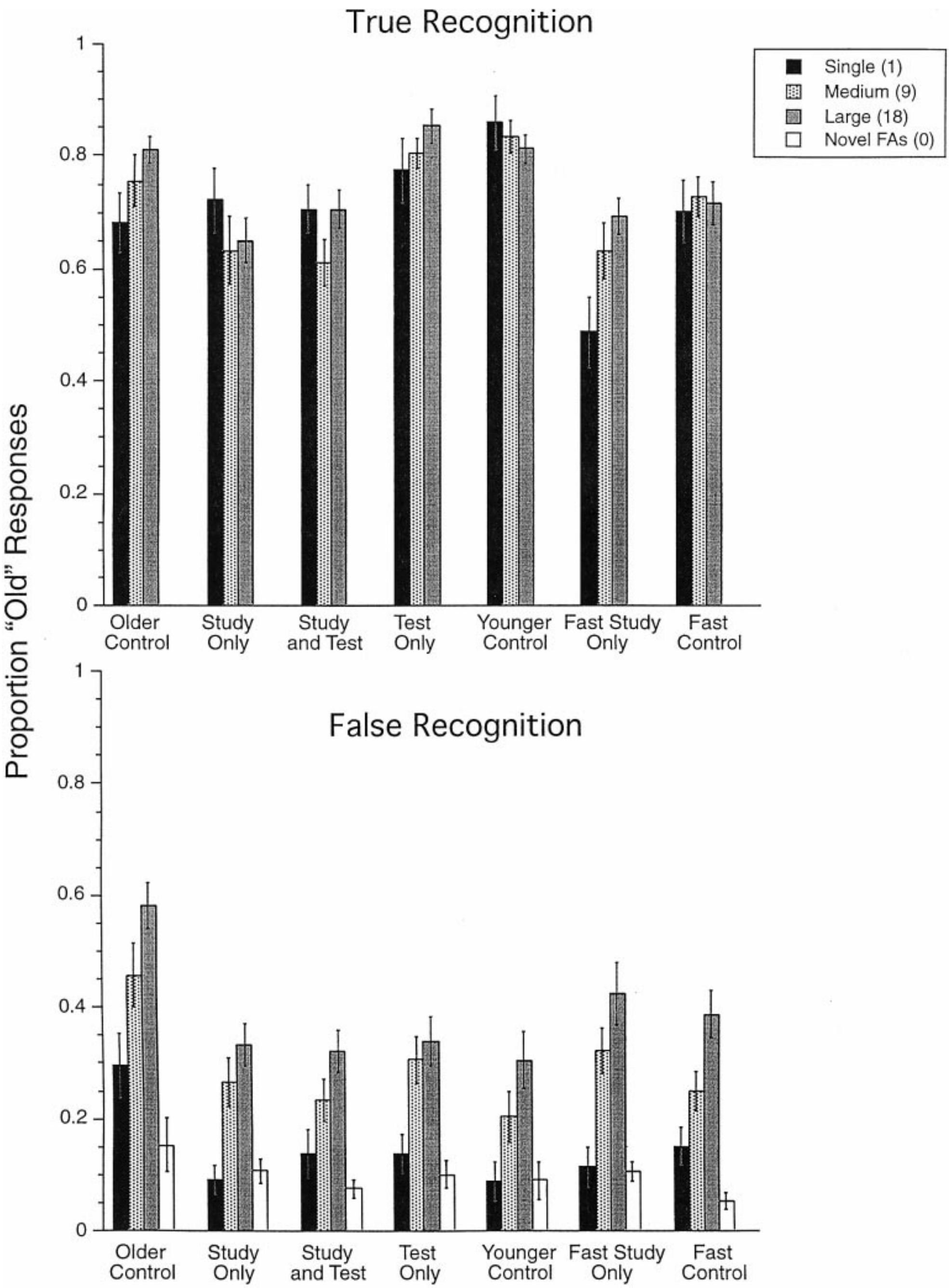

FIG. 1. Mean proportion of "old" responses for studied items ("True Recognition") and nonstudied items ("False Recognition") as a function of category size and group. Category size, or the number of categorically related items presented during study, was 0, 1, 9, or 18 items; the category size of 0 provides the baseline false alarm rate. Study-Only, younger group given dual task at study only; Test-Only, younger group given dual task at test only; Study-and-Test, younger group given dual task at study and test; Fast Study-Only, younger group given stimulus presentation rate of 500 ms at study with dual task at study only; Fast Control, younger group given stimulus presentation rate of 500 ms at study. Error bars show the standard error of the mean.

TABLE 1

True and False Recognition, after Correction for Novel-Category False Alarms, as a Function of Category Size and Group

	True recognition				False recognition			
	Single (1)	Medium (9)	Large (18)	*M*	Single (1)	Medium (9)	Large (18)	*M*
Initial experiment								
Older control	0.53	0.60	0.66	**0.60**	0.14	0.30	0.43	**0.29**
Younger control	0.77	0.74	0.72	**0.75**	–0.002	0.12	0.22	**0.11**
Younger study-only	0.61	0.53	0.54	**0.56**	–0.02	0.16	0.23	**0.12**
Younger test-only	0.68	0.70	0.75	**0.71**	0.04	0.21	0.24	**0.16**
Younger study-and-test	0.63	0.54	0.63	**0.60**	0.06	0.16	0.25	**0.16**
Further conditions								
Younger fast control	0.65	0.68	0.66	**0.66**	0.10	0.20	0.33	**0.21**
Younger fast study-only	0.38	0.53	0.59	**0.50**	0.008	0.22	0.32	**0.18**

Note. Results are shown separately for categories where 1, 9, or 18 categorically related items were presented at study: Younger study-only, dual task at study only; Younger test-only, dual task at test only; Younger study-and test, dual task at study and test; Younger fast control, presentation rate of 500 ms at study; Younger fast study-only, presentation rate of 500 ms at study with dual task at study only.

controls showing a trend toward depressed recognition relative to the younger group, novel-corrected recognition of 0.60 and 0.74, respectively, $F(1, 30) = 3.06$, $MSe = 0.05$, $p = .09$.

We next assessed how well the requirement to perform the dual task at study had matched the younger study-only and older control groups on veridical recognition for one-of-a-kind items. As shown in Fig. 1, the level of veridical recognition of single items in the study-only group was noticeably lower than for the test-only or younger control groups and was relatively similar to that of the older controls. However, single item veridical recognition by older controls was nonetheless still slightly lower than in the study-only group, with novel-corrected scores (see Table 1) of 0.53 versus 0.61, respectively. It was for this reason that we also tested the additional "fast study-only" group, where the stimuli at study were each presented for 500 ms rather than 2000 ms and participants were required to perform the dual task at study-only. From Table 1, it is clear that the combination of a dual task at study together with the decreased stimulus exposure time at encoding markedly reduced recognition of one-of-a-kind single items (novel-corrected recognition of 0.38)—now, though, somewhat "undershooting" the level of older controls for these items. However, *averaging* the initial younger study-only and the fast study-only groups achieved the hoped-for outcome of matched veridical memory for the single items, means of 0.53 for older controls and 0.50 for the combined younger groups, $F < 1$ for the effect of group. Recognition of the unrelated items (which, as noted, also constituted a type of "one-of-a-kind" item) also yielded a nearly perfect match for older adults and the combined study-only plus fast study-only younger groups, corrected recognition of 0.43 and 0.42, respectively, $F < 1$. (Recognition rates for the unrelated items, for all five groups separately, including older control, study-only, study-and-test, test-only, and younger control, respectively, were 0.59, 0.52, 0.55, 0.84, and 0.88; the corresponding false alarm rates for unrelated items were 0.15, 0.12, 0.06, 0.09, and 0.06. For the fast study-only and fast control groups, recognition rates for unrelated items were 0.52 and 0.64, respectively; the corresponding false alarm rates were 0.08 and 0.04, respectively).

Given these successfully matched rates of one-of-a-kind veridical recognition for older adults and the combined study-only plus fast study-only younger groups, we next addressed the critical question regarding the rate of false recognition for categorically related lures in these groups. The outcome was clear: Despite matched levels of one-of-a-kind veridical recognition, older adults still showed significantly elevated false recognition for large (18-exemplar) categories, novel-corrected false recognition of 0.43 versus 0.27, $F(1, 46) = 5.81$, $MSe = 0.05$, $p = .02$ for the effect of group. A similar pattern was found for the medium (9-exemplar) categories, novel-corrected false recognition of 0.30 versus 0.19, $F(1, 46) = 3.23$, $MSe = 0.05$, $p = .08$, and for the combined medium and large categories, means of 0.37 versus 0.23, $F(1, 46) = 5.15$, $MSe = 0.04$, $p = .03$.

As is clear from Fig. 1, this pattern of elevated false recognition in the older control group held for each of the two younger groups (i.e., study-only and fast study-only) and also for each of the remaining younger groups (study-and-test, test-only, younger control, and fast control). For purposes of comparison with previous studies, we also conducted an overall ANOVA on the corrected false recognition scores of the initial five groups given the 2-s stimulus presentation at study, treating group as a between subjects factor (5 levels) and category size (3 levels: single, medium, large) as a within subjects factor. This analysis showed a main effect of group, $F(4, 75) = 3.31$, $MSe = 0.08$, $p = .01$, and of category size, $F(2, 150) = 50.91$, $MSe = 0.02$, $p < .0001$, with no interaction, $F < 1$. Consistent with our earlier studies (Koutstaal & Schacter, 1997b; Koutstaal, Schacter, Galluccio et al., 1999a), and with many other studies using verbal (e.g., Robinson & Roediger, 1997; Shiffrin et al., 1995) and abstract visual materials (e.g., Homa et al., 1973), false recognition increased with increasing category size, novel-corrected means for single, medium, and large of 0.04, 0.19, and 0.27, respectively, pairwise comparisons, large > medium > single, smallest $F = 14.44$. The overall effect of group was nearly entirely attributable to the performance of the older control group, whose average level of novel-corrected false recognition (0.29) well exceeded that of younger participants in any of the other conditions (young control = 0.11, study-only = 0.12, test-only = 0.16, study-and-test = 0.16); for the group comparison of older controls versus all others, $F(1, 78) = 12.37$, $MSe = 0.08$, $p = .0007$.

The Preservation of False Recognition Question

The second question we addressed concerned the relative costs to veridical versus false recognition associated with dual task demands at study. The outcome regarding this question, too, was clear. Despite marked decreases in *veridical* memory shown in the younger study-only and study-and-test groups relative to the younger control group, there was no indication of a decrement in *false* recognition in the younger groups who were given dual task demands at study: novel-corrected false recognition means for the medium and large categories combined were 19 and 20% for the study-only and study-and-test groups, respectively, compared to 17% for the younger control group. This points to the robustness of false recognition in the face of manipulations that reduced veridical recognition.

A 2 × 2 ANOVA, contrasting novel-corrected veridical vs false recognition in the younger control and study-only conditions showed no overall effect of group ($F = 1.90$), a main effect of response type (reflecting greater novel-corrected veridical than false recognition), and, most importantly, a group × response type interaction, $F(1, 30) = 12.56$, $MSe = 0.02$, $p = .001$. Whereas, compared to the younger control group, the dual task study manipulation resulted in a decrease of 20% in veridical recognition (means of 73% vs 53% for the medium and large categories combined), it had very little effect on false recognition (increase of 3%). A similar analysis contrasting novel-corrected veridical vs false recognition of the younger control and study-and-test conditions yielded an essentially parallel pattern of results—no effect of group ($F < 1.2$), an effect of response type (indicating greater novel-corrected veridical than false recognition), and a group × response type interaction, $F(1, 30) = 9.93$, $MSe = 0.01$,

$p = .004$.[1] Relative to the younger control group, dual-task demands at both study and test were associated with a cost to veridical recognition of 15% and an increase of 4% for false recognition.

The two faster presentation conditions also offered a further opportunity to examine the relative robustness of false recognition among younger adults in the face of two manipulations—faster stimulus presentation alone and faster stimulus presentation in combination with a secondary task at study—each of which could be expected to impair veridical memory. Accordingly, we contrasted veridical recognition for the medium and large categories against false recognition for those categories, separately contrasting the fast study-only group and the younger fast control group with the younger control group. For the first of these comparisons, a 2 (response type: novel-corrected veridical or false recognition) × 2 (group: younger control or younger fast study-only) ANOVA revealed no effect of group, $F < 1$, an effect of response type (indicating greater novel-corrected veridical than false recognition), and, most importantly, a group × response type interaction, $F(1, 30) = 23.86$, $MSe = 0.01$, $p < .0001$. Whereas fast stimulus presentation in combination with the dual task at study was associated with a decrease of 18% in veridical recognition (means of 73% vs 56% for younger control and younger fast study-only groups, respectively), it was associated with an increase of 10% in false recognition. A similar analysis contrasting the fast control and young control groups also revealed no effect of group, $F < 1$, and a main effect of response type (again reflecting greater veridical than false recognition), together with a group × response type interaction, $F(1, 30) = 7.85$, $MSe = 0.01$, $p = .009$. However, in this case, the costs to veridical recognition (7%) were less than the *increment* in false recognition (10%).

From these analyses and from Table 1, it appears that younger adults showed somewhat higher levels of false recognition of the categorized lures under the faster study presentation rate, with novel-corrected false recognition for the large categories in the two faster presentation groups (0.33 for fast control and 0.32 for fast study-only) falling intermediate between that for older controls (0.43) and the other younger groups (0.23, 0.25, 0.24, and 0.22 for study-only, study-test, test-only, and young control respectively). Consistent with this, a focused comparison contrasting the two younger groups who studied under the faster presentation rate with those who studied under the slower presentation rate showed that, for the large (18-item) categories, the average rate of novel-corrected false recognition under the faster stimulus presentation rate (33%) was significantly greater than that for the slower rate (23%), $F(1, 94) = 5.11$, $MSe = 0.04$, $p = .03$. A similar, but smaller and nonsignificant, difference was observed for categories comprised of 9 related exemplars (means of 21% compared to 16%, $F < 1.5$). Taken together with the evidence for veridical recognition—which showed a marked decrease in true recognition with the faster presentation, particularly for the single and unrelated items in the fast study-only group (see above), but also for the fast control group and for the medium and large category items as well (see Fig. 1 and Table 1)—these findings suggest that, with a faster presentation rate, younger individuals may extract less item-specific information than with the slower presentation rate, but they may continue to extract sufficient gist-information to support false recognition. Alternatively, or in addition, the faster presentation may also have: (a) partially prevented the extraction of detailed information that otherwise would have allowed younger adults to *correctly reject* some of the categori-

[1] This test of the interaction of veridical vs false memory for the study-and-test group vs younger controls is not independent of our planned test of the same interaction, reported above for the study-only group vs younger controls, nor is it independent of the two further interaction tests, reported later, for the fast study-only and fast control conditions, each of which also involved comparisons with the younger control group. Under these conditions, the Bonferroni procedure provides a conservative correction (Rosenthal & Rosnow, 1991). The younger control group is involved in four interaction tests and also enters into the comparisons of the effects of interference with test monitoring for the test-only and the study-and-test groups, yielding a total of six comparisons. The Bonferroni adjusted alpha level is then $0.05 / 6 = 0.008$. According to this criterion, all three “nonplanned” interaction tests would remain significant.

cally similar lures, or (b) decreased the extent to which participants could form integrated and closely "bound" representations of the multiple features of the objects, thereby increasing the likelihood of errors based on "misplaced recollection," where particular features of the target items lead to the mistaken "recognition" of new items.

The Interference with Test Monitoring Question

Did the requirement to perform the dual task at the time of retrieval lead to increased false recognition in the younger test-only group relative to that observed in the younger control group? Consideration of Fig. 1 and Table 1 suggests that there was little evidence for such an effect. For the large (18-item) categories, neither the test-only nor the study-and-test condition showed appreciably increased false recognition relative to the younger control group (novel-corrected false recognition rates of 0.24, 0.25, and 0.22, respectively, *F*s < 1 for the two pairwise comparisons). For the medium (9-item) categories, there was a numerical tendency toward increased false recognition when the dual task was imposed during recognition testing (novel-corrected false recognition rates of 0.21 and 0.16 for test-only and study-and-test conditions, respectively, versus 0.12 for the young control group); however, these differences were not reliable ($F = 1.49$ and $F < 1$ for the two pairwise comparisons), and the most salient outcome yielded by comparison of these conditions is the considerable similarity (and relatively low level) of false recognition across the three younger groups.

DISCUSSION

The results of this experiment provide evidence for several conclusions relating to (1) the origin of age-related differences in false recognition in the categorized pictures paradigm, particularly with regard to differences in item-specific memory (the matching question), (2) the relative "preservation" of false recognition compared to veridical recognition in younger adults, and (3) the effects of interference with test monitoring on false recognition in younger adults. We discuss each of these in turn.

Age-Related Differences in False Recognition (The Matching Question)

First, age-related differences in gist-based false recognition cannot simply, or entirely, be attributed to accompanying age-related differences in veridical memory, particularly veridical memory for "one-of-a-kind" items. Combining across the two younger groups given the dual task at study-only and those given the dual task at study-only together with a faster stimulus presentation rate at study (i.e., the study-only and fast study-only groups), we achieved an essentially perfect match on one-of-a-kind veridical recognition for younger versus older adults, both as indexed by recognition of categorized items where only a single item per category was presented and by recognition of "miscellaneous" unrelated items. Yet, despite these matched levels of veridical recognition, and with matching achieved for precisely those items that probably most strongly depended on the retention and use of item-specific memory, we still observed a clear elevation in false recognition among older individuals, with the rate of false recognition in the older group (0.37 for the large and medium categories combined) some 1.6 times greater than that found for the matched younger groups (0.23).[2]

This suggests that something in addition to a simple deficit in item-specific memory must contribute to older adults' elevated levels of false recognition in this paradigm. One possibility is that, compared to younger adults, older adults are less likely to actively use or query any item-specific information that they do retain unless provided explicit guidance and support

[2] An examination of the correlations between veridical recognition of the "one-of-a-kind items" and false recognition of the "many-exemplar" category lures both for our earlier experiments and for the current experiment also largely pointed to the comparative "independence" of these measures. For older adults, the average correlation between one-of-a-kind veridical recognition and false recognition across the three earlier experiments of Koutstaal and Schacter (1997) was −.03; for older controls in the present experiment it was .31. For younger adults in the earlier experiments, the average correlation was .31; for the various groups of the current experiment the correlations were younger controls, .47; study-only, .16; test-only, .21; study-and-test, .03; fast study-only, .27; and fast-controls, .19. (Correlations are based on unrelated items for Experiments 1 and 2 of Koutstaal &

during retrieval. The outcomes of our previous study, noted in the Introduction, where older adults showed reduced false recognition when we provided additional retrieval monitoring support during testing by requiring participants to differentiate "old-and-identical" items from "new-but-related" or "new-and-unrelated" items, are consistent with this suggestion (Koutstaal, Schacter, Galluccio et al., 1999a; also cf. Multhaup, 1995). Signal detection analyses on the data from the present experiment (see the Appendix) are likewise consistent with this suggestion. These analyses showed that, for the medium and especially the large categories, older adults were generally more lenient in their responding than were younger adults. Relatively lax criteria among older adults were most consistently observed when comparing hits to false recognition of categorically related items, but a similar tendency toward greater leniency was also found for a measure of "gist sensitivity," where false alarms to related items are treated as "hits" and as indicating memory for the categorical nature of the stimuli. This suggests that older adults may find it especially difficult to avoid "familiarity-based" or gist-based responding.

On the one hand, findings from previous studies demonstrating that older adults may show greater susceptibility to "errors of commission" that may derive from general familiarity, such as "false fame" errors (Dywan & Jacoby, 1990; also see Bartlett, Strater, & Fulton, 1991) and the mistaken recognition of repeated lure items (Jennings & Jacoby, 1997), are also congruent with this interpretation. For example, in the latter situation, Jennings and Jacoby (1997) found that older adults were especially likely to mistakenly endorse as "old" lure items that had occurred previously during the test session itself even though they were instructed that if they noticed that a word had occurred previously during the test, then the word must be a new item. Thus, despite the fact that recollection of the prior within-test presentation of the item would be sufficient to definitively classify it as a lure, older participants were misled by the familiarity of the item.

On the other hand, however, other findings with the false fame paradigm and ironic effects of repetition (Jacoby, 1999b, Exp. 3), where a dual task was imposed during encoding, diverge from the outcomes reported here in that—under these conditions—dual task requirements lead younger adults to respond *similarly* to older adults. For example, Jennings and Jacoby (1993; also cf. Jacoby et al., 1989) found that older adults (under full attention) and younger adults under divided attention at study showed similar rates of false fame errors (the false fame errors of younger adults under full attention were lower than those of the other two groups). These results clearly diverge from the pattern found here, where, despite the effectiveness of the dual task in equating the two age groups on veridical recognition, age differences in false recognition were still found.

Perhaps the strongest parallel to the pattern of findings reported here—including both the component of "matched" veridical recognition in younger and older adults (achieved through the imposition of a secondary task at study) *and* a persistently elevated level of positive errors (errors of commission) in older adults—involved a somewhat different "test-priming" paradigm used by Jacoby (1999a). In this paradigm, older and younger participants initially studied lists of related word pairs (e.g., *bed sheet*, *eagle bird*, *knee bone*). Thereafter, their recall for the target words was tested by presenting the first word of each pair, together with a fragment of the second word (e.g., *knee b_n_*). Additionally, however, during recall some of the words were pre-

Schacter, 1997, and the unweighted average of unrelated items plus single items in Experiment 3 of Koutstaal & Schacter, 1997, and the present experiment; the correlations were obtained separately for the medium and large categories and then averaged). Thus, younger adults did not show a negative correlation between veridical memory for the one-of-a-kind items and false recognition of category lures but, rather, modestly positive correlations. Median split analyses in which we examined the levels of false recognition for the many-exemplar categories separately for individuals within each age group who scored above versus below the median level of veridical recognition for one-of-a-kind items likewise did not show a strong association between these factors. Individuals scoring above versus below the median showed similar levels of false recognition with, if anything, slightly higher false recognition rates shown by those achieving above-median scores.

ceded by the presentation of a "prime" word that was either *congruent* with the studied word pair (e.g., *sheet*; *bed s_ee_*) or *incongruent* with the studied word pair (e.g., *sleep*; *bed s_ee_*). Baseline test items in which no prime word was presented were also included (e.g., *&&&&*; *eagle b__d*). Participants were explicitly informed that the prime word would often be misleading and were told that they should be sure to recall the earlier presented target item rather than be misled. Compared to younger adults, older adults much more often incorrectly intruded the incongruent item. Furthermore, this pattern was *also* found when younger adults were required to perform a dual task during the study phase, thereby depressing their level of correct recall for the baseline test trials to a level equivalent to that of the older group (who performed only the memory encoding task at study).[3]

Taken together, the findings from the present study and those of Jacoby (1999a) point to a limit on the extent to which dual task requirements may act to "mimic" the effects of aging on memory (Craik, 1982), and thus they represent exceptions to the typical pattern in the literature, where placing younger individuals under dual task demands has most often yielded outcomes paralleling those found for older adults (e.g., Dywan et al., 1998; Jacoby, 1999b, Exp. 3; Jennings & Jacoby, 1993; Mäntylä & Bäckman, 1992, Exp. 3; for discussion see Craik, 1982; Craik & Byrd, 1982). Although the imposition of dual task requirements on younger adults may result in a pattern of performance similar to that shown by older adults, these are now two instances where, although the age groups were equated for veridical memory (i.e., probability of errors of omission), age differences were still found in false memory (i.e., probability of errors of commission). One possibility is that these age-related differences in errors of commission were found because the retrieval conditions in these two experiments were such that they placed particularly stringent demands on the need to strategically oppose *not only* a very broad "undifferentiated" form of familiarity-based responding *but also* relatively specific "target-consistent" features of the lure items. In both the Jacoby (1999a) experiment and the present experiment, multiple features of the "lures" may have been consistent with participants' memory representations of the studied item. In the Jacoby experiment, the semantic context and some of the orthographic and phonetic information of the lures matched those of the target items. Likewise, in the present experiment, in the case of items from large categories, in addition to global sources of familiarity for the categorized lures (e.g., semantic information about the categories), specific perceptual features of the lures, such as color or overall shape, may have more or less closely "echoed" features of presented target items. Rejection of such "similar-seeming lures" may have proved especially difficult, with younger adults more often noticing, or actively seeking out, "target-inconsistent" features that would allow correct rejection of the lures. This proposal is also consistent with recent evidence reported by Henkel, Johnson, and De Leonardis (1998) from a source monitoring experiment, where older adults showed particular difficulties in discriminating between items that they had imagined versus items that they had actually perceived when the imagined and perceived items were either physically or conceptually similar to one another (e.g., a lollipop and a magnifying glass, or a banana and an apple). These age differences in source monitoring errors were found even though older and

[3] A further parallel between the results of the Jacoby (1999a) experiments and the current study concerns the patterns of correct performance on congruent trials. In the Jacoby (1999a) experiments, older adults also tended to show more *correct recalls* than did younger adults under dual task conditions for the congruent trials because—unlike younger participants—older adults did not as often try to strategically avoid the effects of accessibility bias, making them both more often correct (when bias would lead to a correct outcome) and more often incorrect (when bias would lead to an incorrect outcome) than was true for the younger divided-attention participants. Examination of the pattern of veridical recognition for the large category items in the present experiment indicates that a similar pattern was found here: In each of the conditions where younger adults were exposed to either a dual task manipulation or a faster study-presentation during encoding, the older controls showed numerically higher levels of (uncorrected) true recognition for the large category items than did the younger groups, with this holding true for the study-only, study-and-test, fast study-only, and fast control groups.

younger adults had been *equated* on their level of old–new recognition (through using a longer retention interval for the younger adults) and were "stimulus specific" in that a similar age-related source monitoring deficit was not found for control items (i.e., items where the perceived and imagined items had no particular conceptual or physical similarity with one another).

Although these outcomes underscore the possible particular importance of more stimulus-specific similarities in the lures in contributing to the false recognition performance of older adults, two considerations suggest that age-related differences under these conditions may not entirely or exclusively relate to *strategic* retrieval-monitoring differences in older and younger adults. First, this account might also lead to the expectation that—with dual task demands imposed at the time of testing, or at both study-and-test—younger adults should *also* show a parallel increased difficulty in successfully rejecting lures. Yet—as the results reported above demonstrated—this outcome was not found. This apparently "stronger-than-expected" resistance to false recognition shown by younger adults is discussed further below in regard to the interference with test monitoring question. Second, neuropsychological evidence suggests that frontal regions are particularly involved in strategic retrieval and evaluation (e.g., Johnson et al., 1993; Moscovitch, 1994; West, 1996) and older adults' deficits in source monitoring and difficult forms of retrieval have been associated with aging-related declines in frontal lobe functioning (e.g., Johnson et al., 1993; Parkin & Walter, 1992; Schacter, Savage, Alpert, Rauch, & Albert, 1996; Spencer & Raz, 1995). However, although correlations between measures of frontal functioning and source memory performance have sometimes been found (e.g., Craik, Morris, Morris, & Loewen, 1990; Glisky, Polster, & Routhieaux, 1995; Spencer & Raz, 1994), such correlations have not always emerged (e.g., Schacter, Kaszniak, Kihlstrom, & Valdiserri, 1991; Spencer & Raz, 1994). Furthermore, in the condition of Henkel et al. (1998), where older and younger adults were matched on old/new recognition but older adults showed item-specific deficits in source monitoring, there was no correlation between overall source memory accuracy and a battery of neuropsychological tests measuring frontal function, although there *was* a correlation with measures of temporal lobe–hippocampal function (and both frontal and temporal batteries showed correlations with source accuracy at a longer retention interval).

On the basis of this finding, and also given evidence indicating that (a) in addition to changes in frontal lobe functioning, older adults may also show age-related neuropathology in temporal and hippocampal regions (e.g., Grady et al., 1995; Raz, Millman, & Sarpel, 1990) and (b) older adults may have particular deficits in binding multiple features of an episode with one another (e.g., Chalfonte & Johnson, 1996), Henkel et al. (1998) argued that attributing source memory errors among older adults entirely to frontal lobe functioning would be an oversimplification. Rather, particularly under conditions where there is a high level of similarity between studied items, efficient binding processes in medial-temporal regions may be particularly required to prevent confusion. In a review of aging-related patterns of false recall and recognition, we (Schacter, Norman, & Koutstaal, 1998) addressed somewhat similar issues, explicitly noting that, apart from possible age-related differences in specificity of encoding or criterion-setting at retrieval, the high levels of false recognition among older adults for categorically related items might derive from a specific impairment of the hippocampal mechanisms involved in pattern separation and binding. To the extent that (in addition to less stringent retrieval monitoring) older adults *also* show relatively weaker "binding" of the many features of episodically encountered stimuli, two further outcomes might follow. First, older individuals might be more likely to incorrectly accept particular lure items because those lures have "isolated" features that are consistent with the studied items; second, older individuals might also be less likely to successfully reject lures based on the detection of features in the new items that are inconsistent with the studied items. Additionally, to the extent that such age differences in the effectiveness of binding and medial-temporal functions might be largely independent of attentional factors relevant during *retrieval* but possibly more closely "simulated" by a curtailed processing period at study, two of the outcomes that we observed for younger adults—no increase in false recognition among younger adults under dual task conditions at retrieval (see also below) but a modest and significant increase in false recognition given a faster presentation rate at study—might also follow.

The Relative "Preservation" of False Recognition Compared to Veridical Recognition in Younger Adults

Perhaps the most important outcome yielded by the comparison of the younger groups was the demonstration of differential effects of dual task demands during *study* on veridical versus false recognition. Whereas, consistent with the findings of many previous studies (e.g., Baddeley et al., 1984; Craik et al., 1996; Mulligan, 1998; Naveh-Benjamin et al., 1998; Park et al., 1989), we found that dual task demands at study produced large decrements in veridical recognition, these same manipulations entirely spared (or slightly augmented) false recognition responses. Thus we observed substantial decrements in veridical recognition in several conditions—dual task demands at study-only, dual task at both study and test, dual task at study in combination with the rapid stimulus presentation rate, and the fast control condition with only a rapid presentation rate during study (without a secondary task)—yet we observed no decrements in false recognition but, rather, slight-to-modest increases. Reliable interactions between response type (veridical or false recognition) and study condition were obtained in comparisons contrasting young controls against each of these four conditions.

These outcomes are consistent with an interpretation of false recognition among younger adults as deriving from general similarity information or a sense of familiarity with the studied items and the perceptual/conceptual categories that they represented. Whereas access to information about the general categories or types of items that were studied may not be impaired by manipulations of the form used here, allowing relative "preservation" of false recognition, access to more detailed and item-specific information may be decreased, resulting in impaired veridical recognition, particularly if the task manipulation did not lead to a shift to more general similarity-based responding (in the case of targets from many-exemplar categories) or if the opportunities for such responding were minimized (in the case of one-of-a-kind items).

The observation of such interactions is also broadly consistent with a global matching perspective of true and false recognition, such as that recently outlined and tested by Arndt and Hirshman (1998) in relation to semantically associated words. In this account, the features of each item are probabilistically encoded into a unique vector in memory. During recognition testing, the test item is matched to all of the items stored in memory, with the consequent "familiarity" or "strength of memory" signal greater wherever features "match" than when they do not. However, particularly where many of the studied items are related to one another, a high amount of familiarity can derive from one of two sources, a large number of small or partial matches derived from many studied items or a highly "specific" match, where a given trace strongly matches the test probe of a particular studied exemplar. Applying this conception to the present experiments, whereas performance of the secondary digit-monitoring task during encoding might be expected (under at least some accounts) to diminish the number of features that are included in the specific traces of individual items—and therefore decrease correct recognition responses based on a high degree of "match" between the test probe and those specific traces—the secondary task might exert much less influence on the smaller amounts of similarity derived from features common to all or multiple instances within a category. Common features, such as the global shape of objects within a category or their category names, might be comparatively easily or automatically encoded for most or all of the items within a category, thereby leaving unaffected (or possibly increasing) false recognition responses based on the summed similarity across these features.

In the present case, emphasis on these two, quite different, sources of high "familiarity" or memory "strength" that can support affirmative recognition responses may provide an account of why the dual task manipulations differentially affected veridical versus false recognition. However, it is also important to note that there are other manipulations or subject conditions that exert *parallel* effects on true and false recognition (e.g., Toglia, Neuschatz, & Goodwin, 1999). For example, using the Deese (1959)/Roediger-McDermott (1995) verbal converging associates paradigm, Seamon et al. (1998, Exp. 2) found that a very rapid presentation rate during study decreased both veridical recognition of list items *and* false recognition of critical lures (though veridical memory was more adversely affected than false memory). Similarly, Arndt and Hirshman (1998) found that both veridical and false recognition (d') of verbal associates were lower for shorter lists of associates (comprised of 4 associates) than longer lists (comprised of 16 associates), though, in this case, false recognition was affected more than veridical memory. Likewise, investigations of false recognition in amnesics, i.e., persons who, as a consequence of damage to medial temporal and diencephalic brain regions, suffer from impaired veridical memory, have generally shown parallel decreases in false memory: Under conditions where control participants show high rates of false recognition, as in the case of multiple semantically associated words (Schacter, Verfaellie, & Anes, 1997; Schacter, Verfaellie, Anes, & Racine, 1998; Schacter, Verfaellie, & Pradere, 1996) or false recognition of the prototypes of perceptually similar novel objects (Koutstaal, Schacter, Verfaellie, Brenner, & Jackson, 1999b), amnesics show both impaired veridical recognition *and* impaired false recognition.

These outcomes may partially reflect differential degrees of reliance on "high quality" matches to a single trace, versus the summed similarity derived from multiple partial matches, under different encoding and retrieval conditions (and encoding/retrieval *combinations*; cf. Schacter, Israel, & Racine, 1999; Schacter, Koutstaal, & Norman, 1997). Circumstances under which veridical memory and false memory are impaired in parallel may be those in which veridical memory, too, is more heavily dependent on a form of general similarity or familiarity: for example, where no specific encoding task is used or where the stimulus items have relatively few differentiating perceptual features that can be encoded. To the extent that both true and false recognition are supported by a similar form of "summed similarity" across multiple items, we might expect to see parallel increments and decrements on measures of veridical and illusory recognition. From this vantage point, what is notable in the present study is the resistance shown by younger adults to falling back on this form of responding, with the consequence that, although we observed constant or slightly incremented levels of false recognition in younger adults, the levels of false recognition remained below that shown by older adults. Despite conditions that depleted the amount and quality of the attentional and perceptual resources that they could devote to the encoding of pictures, younger adults did not opt to primarily respond on the basis of categorical information: indeed, as noted above, for categories where numerous related items had been presented, younger adults in the conditions involving dual task at study showed slightly lower recognition, combined with trends toward more stringent response criteria, than did older adults. This resistance to relying largely or entirely on generic information thus bore some cost to veridical recognition, but it was accompanied by little increment in false recognition.

One alternative account of all these findings was briefly noted in the Introduction and also in the Discussion above: Might false recognition responses derive from "misplaced recollection"—the miscombining or misconjoining (cf. Reinitz et al., 1994) of specific recollected features—rather than (or in addition to) mistaken familiarity or general similarity responding? It is possible that at least some false recognition responses of the categorized picture stimuli of these (and earlier) experiments for both older and younger adults are of this form, broadly characterized as relatively more specific errors ("I saw a large tawny colored cat that looked very much like this particular cat") than entirely generic errors ("I saw many cats"). Indeed, stated in these terms, the two forms of errors need not be entirely disjunct from one another as it is quite probable that even “categorically based” errors are partially based on some specific perceptual features of the lure items. The existence of at least some perceptual specificity in false recognition is implied by the observation that not *all* lures within a given category are falsely recognized (suggesting that some lures seem more similar to studied items than others). Relatedly, some perceptual specificity is also implied by the levels of within-category discriminability (that is, hits–related false alarms) that were observed. Averaging across the medium (9-exemplar) and large (18-exemplar) categories, all groups continued to show some such within-category discriminability. Ordered from the group showing the least to the greatest within-category discriminability, the mean hit–related false alarm rates were older controls (0.26), fast study-only (0.29), study-only (0.34), study-and-test (0.38), fast control (0.40), test-only (0.51), and young controls (0.57). Pairwise comparisons relative to older controls revealed no difference between older controls and the fast study-only group, $F < 1$, a weak trend toward more accurate discrimination in the study-only group than in older controls, $F(1, 30) = 2.10$, $MSe = 0.02$, $p = .16$, and significantly more accurate discrimination in each of the remaining groups relative to older controls, smallest $F = 5.96$.

Contrasting these two accounts may thus underscore a key point: Responding on the basis of “gist” or “general similarity” may to some extent occur at multiple different levels of grain—I saw cats, I saw black cats, I saw black cats with short hair. And, as more and more features are specified (or required) before individuals are willing to provide a positive recognition response, responses become closer to “recollection”—and also potentially to “misplaced recollection.” Additionally, with exposure to a large number of different exemplars, keeping representations of the studied exemplars separate (“pattern separation”; cf. Schacter et al., 1998) may become increasingly difficult. Thus, rather than two entirely separate bases of false recognition, the gist-based false recognition and misplaced recollection accounts may represent points on a continuum, with the responses of older adults situated closer to the sparsely specified end (primarily category information but with some additional relatively specific information given that, for older adults, too, not all categorically related lures are mistakenly identified as old) and the responses of younger adults possibly situated at the relatively more detailed pole (indicated by increased within-category discriminability and possibly involving multiple or conjunctive features rather than more isolated features). Some evidence for this may be provided by a consideration of the *correlations between veridical and false recognition* for categorized items. Although consistently positive and generally quite strong, particularly for the large (18-item) categories for all of the groups, this correlation was stronger for older adults than for any of the younger groups (ordered by strength of correlation, from most to least, the correlations for large category items were, for older controls, $r = .83$, for study-only, $r = .68$, for fast study-only, $r = .67$, for fast control, $r = .52$, for test-only, $r = .50$, for young control, $r = .44$, and for study-and-test, $r = .36$; for the medium categories, the veridical–false recognition correlations, also ordered from the strongest to least strong were, for older controls, $r = .79$, for fast study-only, $r = .54$, for young control, $r = .54$, for study-only, $r = .43$, for study-and-test, $r = .42$, for test-only, $r = .09$, and for fast control, $r = .08$).

The Effects of Interference with Test Monitoring on False Recognition in Younger Adults

We found that, in younger adults, dual task demands at test carried only slight and unreliable attendant costs in the form of an increased likelihood of false recognition. Averaging across the conditions where either 9 or 18 categorically related items had been presented, the rate of false recognition in the test-only condition was only 3% greater than that shown by younger adults under conditions with no secondary task during testing. Moreover, as can be seen from Fig. 1, there was no indication that this apparent “resistance” to increased false recognition was achieved at a cost to veridical recognition (i.e., by more often incorrectly rejecting studied items): the level of novel-corrected veridical recognition for the many-exemplar conditions was nearly identical in the two groups (means of 0.729 and 0.734 for the test-only group and younger control group, respectively, $F < 1$ for the effect of group).

This finding is at least consistent with the possibility that there are relatively pervasive—and robust—differences in the way information is queried during retrieval by younger versus older adults (cf. Bartlett, Leslie, Tubbs, & Fulton, 1989; Reder, Wible, & Martin, 1986; also see earlier discussion of a global matching perspective). Although, in our previous studies (Koutstaal et al., 1999a), we found that false recognition in older adults could be considerably reduced by providing support for careful monitoring at test, it does not appear that simply increasing the demands during testing is sufficient to induce younger adults to alter their recognition decision-making and to behave more similarly to older adults in the absence of such retrieval support. Indeed, even when younger adults were asked to negotiate the extra demands imposed by the digit-monitoring task *both* during study and during testing—and despite clear attendant costs in veridical recognition from such demands—novel-corrected false recognition of younger adults for the medium and large categories (20%) was, on average, still nearly only half that shown by older adults (37%) for those items, $F(1, 30) = 5.96$, $MSe = 0.04$, $p = .02$ for the effect of age.

This quite marked pattern of virtually *no* interference with test monitoring in younger adults contrasts with previous findings from other paradigms where the requirement to perform a secondary task at the time of attempted retrieval has been found to increase the likelihood that younger adults will show a number of types of "errors of commission," including false fame errors (Jacoby et al., 1989, Exp. 3), false recognition of repeated or "test-primed" lures (Merikle & Joordens, 1997), and increased likelihood of source misattributions (Dodson, Holland, & Shimamura, 1998; Dywan et al., 1998). For example, Dodson et al. (1998) found that dividing attention at retrieval impaired younger participants' memory for specific source information (which of two female speakers or two male speakers had read the items) although it did not impair access to "partial-source information" (a more "coarse-grained" judgment of whether the item was presented by a male or a female speaker, i.e., gender information regarding the source). Dodson et al. (1998) suggested that attention at retrieval might support a "recollective focusing mechanism," with divided attention during retrieval "blurring" this focusing mechanism, and thus making it more difficult to recover relatively precise information than less precise information.

There are several possible accounts of why, in the present experiment, the imposition of a dual task at retrieval did not lead to increased false recognition (as might be expected if relatively specific information was required to reject the nonstudied but conceptually and perceptually related lures). One possibility is that the general attentional and cognitive demands of the secondary task were insufficiently strong and participants could readily perform the conjoint tasks of yes/no recognition and digit monitoring. Although the digit-shadowing and monitoring task was sufficiently difficult to substantially decrease veridical recognition when it occurred at *study*, the task might not have been sufficiently demanding to exact costs during *recognition testing*, particularly given the timing parameters that we used (each picture was shown for 2 s, followed by 4 s to provide an old/new response), and also the known relative robustness of retrieval to disruption by additional attentional demands (Craik et al., 1996; Naveh-Benjamin et al., 1998; also cf. Moscovitch, 1994). Thus, increased false recognition might be observed under more difficult conditions, for instance, if the pictures were presented for a shorter period of time at test or if less time was provided for the recognition decisions. (Note that we found that a faster stimulus presentation time at *study* reliably increased false recognition in younger adults whereas the imposition of a dual task alone did not do so.) Another possibility is that younger adults were highly adept at detecting or noticing "target-incongruent" information in the lures, perhaps doing so relatively "effortlessly" or "automatically," thereby allowing correct rejection of the lures even under dual task conditions. (This connects with the earlier proposal of the importance of considering not only age-related differences in strategic and frontally guided factors but also a possible contribution from medial-temporal binding processes.) The comparatively rich and detailed nature of the pictures used in this experiment might provide more opportunities for the detection of such "give-away" (target-inconsistent) information than were present in the previous studies where increased errors of commission were observed under divided attention at test (e.g., errors involving false fame and within-test repetition of lures), and where participants themselves might need to more *actively retrieve* or *seek out* contextual information that was not provided in the stimulus itself and that would allow correct rejection of the lure. A systematic manipulation of the amount and type of similarity between the target and lure items would allow examination of this question (how much "give-away" or target-inconsistent information needs to be present for younger participants to continue to show resistance to false recognition under dual task conditions?) and would also permit examination of the possibility that younger adults are more adept at noticing features that do not "match" with studied items than are older adults. Nonetheless, under circumstances of the sort used here, it is clear that—although younger adults in the test-only and study-and-test conditions did sometimes incorrectly claim to recognize categorized lure items—the likelihood of such false recognition was no greater than that shown by younger controls.

CONCLUSIONS

Taken together, the outcomes regarding each of these three questions underscore the value of using the dual task methodology to examine both errors of omission and errors of commission in long-term memory. They point to three primary conclusions and interpretations. First, the observation that older adults continue to show elevated false recognition even when their level of veridical recognition of "one-of-a-kind" items is equated with that of younger adults may be partially attributable to age-related differences in (frontally mediated) strategic monitoring or decision criteria but possibly *also* to age-related differences in (temporally mediated) "binding" processes. Under conditions involving high levels of feature similarity across items, more efficient feature-binding may allow younger adults to more often successfully reject similar-seeming lures (even in the presence of a dual task at test). By contrast, under these same conditions, older adults may find that lures with some features that (nearly) match features of the target items are especially difficult to resist—thus accounting for the persistence of higher levels of false recognition among older adults in this paradigm even when younger adults are operating under dual task demands, but not in other paradigms where the lures are "attractive" primarily on the basis of broad familiarity.

Second, the maintenance or "preservation" of false recognition in younger adults under dual task conditions that markedly depressed veridical recognition (dual task at study-only, dual task at study-and-test) may indicate that false recognition in younger adults is based on a combination of familiarity and some more specific similarity information. In younger adults, we found that only conditions that may have involved degraded encoding and binding processes (a faster presentation rate at study or a faster presentation rate at study in combination with dual task requirements at study) produced a significant increase in false recognition.

Third, among younger adults, the imposition of dual task demands at the time of retrieval was not sufficient to impair veridical recognition, nor was it enough to increase false recognition. This "null" finding for false recognition is possibly attributable to a low level of difficulty for the combined tasks, such that younger adults could readily accommodate the demands of both the recognition and the digit-monitoring tasks. Alternatively, it is also possible that secondary task performance at the time of testing did not greatly interfere with younger participants' ability to make detailed within-category discriminations because—provided that the items were initially adequately encoded—"incongruent" information was relatively readily and "effortlessly" noticed by younger adults at the time of test. More detailed specification of the conditions under which the requirement to negotiate between the attentional requirements of two tasks concurrently yields patterns of errors of commission in younger adults that *parallel* those found in older adults (e.g., Jennings & Jacoby, 1993; also cf. Jacoby et al., 1989), versus the circumstances under which such requirements, while matching age groups on some measures of "memory accessibility," nonetheless yield age-related differences in memory accuracy (Jacoby, 1999a; Henkel et al., 1998; the present experiment), should further illuminate how dual task demands influence encoding and retrieval processing and should help to more clearly delineate the factors contributing to age differences in veridical and illusory memory.

APPENDIX

In order to more fully characterize the nature of the dual task effects on performance, we also computed signal detection measures for all of the groups (the five groups from the initial experiment and the two additional fast stimulus presentation groups), providing measures of sensitivity and response bias under varying levels of concurrent task demands at study and/or test. The measures we used were A′, a measure of sensitivity, and $B_{D''}$, a measure of response bias (Grier, 1971; also see Donaldson, 1993; as recommended by Snodgrass & Corwin, 1988, all data were first transformed by computing $p(x)$ as $(x + .5) / n + 1$, rather than x/n; in addition, modified formulas for below chance performance, from Aaronson & Watts, 1987, were used). The measure A′ can vary between 0.00 and 1.00, with values of 0.50 indicating a chance level of performance. Measure $B_{D''}$ can vary between –1.00, indicating extremely lenient responding, and +1.00, indicating extremely conservative responding; $B_{D''}$ values near zero thus indicate unbiased responding. We obtained measures of sensitivity and response bias for two forms of item-specific memory, comparing Hits to Novel False Alarms (designated as A′-Novel and $B_{D''}$-Novel, respectively) and comparing Hits to Related False Alarms (designated as A′-Related and $B_{D''}$-Related, respectively). We also computed sensitivity and response bias measures for "gist-memory," where false alarms to categorically related lures are treated as "hits," indicating memory for the general class or type of items studied (i.e., conceptual or perceptual "gist") and are compared to false alarms to novel category items (designated as A′-Gist and $B_{D''}$-Gist, respectively; cf. Koutstaal & Schacter, 1997b; Tussing & Greene, 1997). These three measures and their corresponding measures of response bias are shown in Table A1, separately as a function of group and category size (single, medium, and large); the means for each group are also shown.

TABLE A1

Measures of Sensitivity and Response Bias as a Function of Category Size and Group, Initial Experiment and Fast Stimulus Presentation Conditions Combined

	A′-Novel				BD″-Novel			
Group	Single (1)	Medium (9)	Large (18)	*M*	Single (1)	Medium (9)	Large (18)	*M*
Older control	0.82	0.86	0.87	**0.85**	0.47	0.21	0.22	**0.30**
Younger control	0.90	0.90	0.90	**0.90**	0.28	0.34	0.42	**0.34**
Younger study-only	0.86	0.83	0.84	**0.84**	0.38	0.45	0.54	**0.46**
Younger test-only	0.87	0.90	0.91	**0.90**	0.37	0.31	0.13	**0.27**
Younger study-and-test	0.87	0.85	0.88	**0.87**	0.54	0.63	0.51	**0.56**
Younger fast control	0.88	0.89	0.89	**0.89**	0.56	0.53	0.59	**0.56**
Younger fast study-only	0.77	0.83	0.86	**0.82**	0.64	0.49	0.48	**0.54**
M	**0.85**	**0.87**	**0.88**		**0.46**	**0.42**	**0.41**	

	A′-Related				BD″-Related			
Group	Single (1)	Medium (9)	Large (18)	*M*	Single (1)	Medium (9)	Large (18)	*M*
Older control	0.73	0.72	0.69	**0.71**	0.04	–.33	–.58	**–0.29**
Younger control	0.88	0.87	0.82	**0.86**	0.11	–.02	–.19	**–0.03**
Younger study-only	0.85	0.75	0.73	**0.78**	0.28	.14	.01	**0.14**
Younger test-only	0.85	0.80	0.83	**0.83**	0.10	–.26	–.39	**–0.18**
Younger study-and-test	0.82	0.76	0.76	**0.78**	0.32	.30	–.03	**0.19**
Younger fast control	0.81	0.81	0.74	**0.78**	0.25	.03	–.20	**0.03**
Younger fast study-only	0.74	0.72	0.70	**0.72**	0.57	.05	–.12	**0.17**
M	**0.81**	**0.78**	**0.75**		**0.24**	**–0.01**	**–0.21**	

Group	A′-Gist				BD″-Gist			
	Single (1)	Medium (9)	Large (18)	*M*	Single (1)	Medium (9)	Large (18)	*M*
Older control	0.63	0.73	0.79	**0.72**	0.70	0.66	0.54	**0.63**
Younger control	0.59	0.63	0.70	**0.64**	0.72	0.76	0.79	**0.76**
Younger study-only	0.54	0.64	0.69	**0.63**	0.62	0.81	0.78	**0.74**
Younger test-only	0.59	0.69	0.70	**0.66**	0.73	0.78	0.82	**0.78**
Younger study-and-test	0.61	0.65	0.72	**0.66**	0.67	0.82	0.86	**0.78**
Younger fast control	0.67	0.70	0.77	**0.71**	0.83	0.92	0.87	**0.87**
Younger fast study-only	0.54	0.69	0.74	**0.66**	0.48	0.81	0.74	**0.68**
M	**0.60**	**0.67**	**0.73**		**0.68**	**0.79**	**0.77**	

Note. Results are shown separately for categories where 1, 9, or 18 categorically related items were presented at study: Younger study-only, dual task at study only; Younger test-only, dual task at test only; Younger study-and-test, dual task at study and test; Younger fast control, stimulus presentation rate of 500 ms at study; Younger fast study-only, stimulus presentation rate of 500 ms at study with dual task at study only.

REFERENCES

Aaronson, D., & Watts, B. (1987). Extensions of Grier's computational formulas for A′ and B″ to below-chance performance. *Psychological Bulletin*, **102,** 439–442.

Anderson, N. D., Craik, F. I. M., & Naveh-Benjamin, M. (1998). The attentional demands of encoding and retrieval in younger and older adults: 1. Evidence from divided attention costs. *Psychology and Aging*, **13,** 405–423.

Arndt, J., & Hirshman, E. (1998). True and false recognition in MINERVA2: Explanations from a global matching perspective. *Journal of Memory and Language*, **39,** 371–391.

Baddeley, A. D., Lewis, V., Eldridge, M., & Thomson, N. (1984). Attention and retrieval from long-term memory. *Journal of Experimental Psychology: General*, **13,** 518–540.

Bartlett, J. C., Leslie, J. E., Tubbs, A., & Fulton, A. (1989). Aging and memory for pictures of faces. *Psychology and Aging*, **4,** 276–283.

Bartlett, J. C., Strater, L., & Fulton, A. (1991). False recency and false fame of faces in young adulthood and old age. *Memory & Cognition*, **19,** 177–188.

Bourke, P. A., Duncan, J., & Nimmo-Smith, I. (1996). A general factor involved in dual-task performance decrement. *Quarterly Journal of Experimental Psychology*, **49A,** 525–545.

Brainerd, C. J., Reyna, V. F., & Kneer, R. (1995). False-recognition reversal: When similarity is distinctive. *Journal of Memory and Language*, **34,** 157–185.

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

Craik, F. I. M. (1982). Selective changes in encoding as a function of reduced processing capacity. In F. Klix, J. Hoffman, & E. van der Meer (Eds.), *Cognitive research in psychology* (pp. 152–161). Berlin: Deutscher Verlag der Wissenschaffen.

Craik, F. I. M., & Byrd, M. (1982). Aging and cognitive deficits: The role of attentional resources. In F. I. M. Craik & S. Trehub (Eds.), *Aging and cognitive processes* (pp. 191–211). New York: Plenum.

Craik, F. I. M., Govoni, R., Naveh-Benjamin, M., & Anderson, N. D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology: General*, **125,** 159–180.

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.

Curran, T., Schacter, D. L., Norman, K. A., & Galluccio, L. (1997). False recognition after a right frontal lobe infarction: Memory for general and specific information. *Neuropsychologia*, **35,** 1035–1049.

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

Dodson, C. S., Holland, P. W., & Shimamura, A. P. (1998). On the recollection of specific- and partial-source information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **24,** 1121–1136.

Donaldson, W. (1993). Accuracy of *d*′ and *A*′ as estimates of sensitivity. *Bulletin of the Psychonomic Society*, **31,** 271–274.

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

Dywan, J., Segalowitz, S. J., & Webster, L. (1998). Source monitoring: ERP evidence for greater reactivity to nontarget information in older adults. *Brain and Cognition*, **36,** 390–430.

Estes, W. K. (1997). Processes of memory loss, recovery, and distortion. *Psychological Review*, **104,** 148–169.

Gardiner, J. M., & Parkin, A. J. (1990). Attention and recollective experience in recognition memory. *Memory & Cognition*, **18,** 579–583.

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

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

Grier, J. B. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. *Psychological Bulletin*, **75,** 424–429.

Gruppuso, V., Lindsay, S. D., & Kelley, C. M. (1997). The process-dissociation procedure and similarity: Defining and estimating recollection and familiarity in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **23,** 259–278.

Hay, J. F., & Jacoby, L. L. (1999). Separating habit and recollection in young and older adults: Effects of elaborative processing and distinctiveness. *Psychology and Aging*, **14,** 122–134.

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

Hess, T. M. (1985). Aging and context influences on recognition memory for typical and atypical script actions. *Developmental Psychology*, **21,** 1139–1151.

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace theory of memory. *Psychological Review*, **95,** 528–551.

Hintzman, D. L., & Curran, T. (1994). Retrieval dynamics of recognition and frequency judgments: Evidence for separate processes of familiarity and recall. *Journal of Memory and Language*, **33,** 1–18.

Hintzman, D. L., & Curran, T. (1995). When encoding fails: Instructions, feedback, and registration without learning. *Memory & Cognition*, **23,** 213–226.

Holmes, J. B., Waters, H. S., & Rajaram, S. (1998). The phenomenology of false memories: Episodic content and confidence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **24,** 1026–1040.

Homa, D., Cross, J., Cornell, D., Goldman, D., & Shwartz, S. (1973). Prototype abstraction and classification of new instances as a function of number of instances defining the prototype. *Journal of Experimental Psychology*, **101,** 116–122.

Iidaka, T., Anderson, N., Kapur, S., Cabeza, R., Okamoto, C., & Craik, F. I. M. (1999). Age-related differences in brain activation during encoding and retrieval under divided attention: A positron emission tomography (PET) study. *Brain and Cognition*, **39,** 53–55.

Isingrini, M., Fontaine, R., Taconnat, L., & Duportal, A. (1995). Aging and encoding in memory: False alarms and decision criteria in a word-pair recognition task. *International Journal of Aging and Human Development*, **41,** 79–88.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, **30,** 513–541.

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

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

Jacoby, L. L., Toth, J. P., & Yonelinas, A. P. (1993). Separating conscious and unconscious influences of memory: Measuring recollection. *Journal of Experimental Psychology: General*, **122,** 139–154.

Jacoby, L. L., Woloshyn, V., & Kelley, C. (1989). Becoming famous without being recognized: Unconscious influences of memory produced by dividing attention. *Journal of Experimental Psychology: General*, **118,** 115–125.

Jacoby, L. L., Yonelinas, A. P., & Jennings, J. M. (1996). The relation between conscious and unconscious (automatic) influences: A declaration of independence. In J. Cohen & J. W. Schooler (Eds.), *Scientific approaches to the question of consciousness* (pp. 13–47). Hillsdale, NJ: Erlbaum.

Jennings, J. M., & Jacoby, L. L. (1993). Automatic versus intentional uses of memory: Aging, attention, and control. *Psychology and Aging*, **8,** 283–293.

Jennings, J. M., & Jacoby, L. L. (1997). An opposition procedure for detecting age-related deficits in recollection: Telling effects of repetition. *Psychology and Aging*, **12,** 352–361.

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

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

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

Koutstaal, W., & Schacter, D. L. (1997a). Inaccuracy and inaccessibility in memory retrieval: Contributions from cognitive psychology and neuropsychology. In P. S. Appelbaum, L. A. Uyehara, & M. R. Elin (Eds.), *Trauma and memory: Clinical and legal controversies* (pp. 93–137). New York: Oxford Univ. Press.

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

Koutstaal, W., Schacter, D. L., Galluccio, L., & Stofer, K. A. (1999a). Reducing gist-based false recognition in older adults: Encoding and retrieval manipulations. *Psychology and Aging*, **14,** 220–237.

Koutstaal, W., Schacter, D. L., Verfaellie, M., Brenner, C., & Jackson, E. M. (1999b). Perceptually based false recognition of novel objects in amnesia: Effects of category size and similarity to category prototypes. *Cognitive Neuropsychology*, **16,** 317–341.

Mandler, G., & Worden, P. E. (1973). Semantic processing without permanent storage. *Journal of Experimental Psychology*, **100,** 277–283.

Mäntylä, T., & Bäckman, L. (1992). Aging and memory for expected and unexpected objects in real-world settings. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **18,** 1298–1309.

Merikle, P. M., & Joordens, S. (1997). Parallels between perception without attention and perception without awareness. *Consciousness and Cognition*, **6**, 219–236.

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.

Mulligan, N. W. (1998). The role of attention during encoding in implicit and explicit memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **24,** 27–47.

Multhaup, K. S. (1995). Aging, source, and decision criteria: When false fame errors do and do not occur. *Psychology and Aging*, **10,** 492–497.

Naveh-Benjamin, M., Craik, F. I. M., Guez, J., & Dori, H. (1998). Effects of divided attention on encoding and retrieval processes in human memory: Further support for an asymmetry. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **24,** 1091–1104.

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

Nyberg, L., Nilsson, L-G., Olofsson, U., & Backman, L. (1997). Effects of division of attention during encoding and retrieval on age differences in episodic memory. *Experimental Aging Research*, **23,** 137–143.

Omohundro, J. (1981). Recognition vs. classification of ill-defined category exemplars. *Memory & Cognition*, **9,** 324–331.

Park, D. C., Puglisi, J. T., Smith, A. D., & Dudley, W. N. (1987). Cue utilization and encoding specificity in picture recognition by older adults. *Journal of Gerontology*, **42,** 423–425.

Park, D. C., Smith, A. D., Dudley, W. N., & Lafronza, V. N. (1989). Effects of age and a divided attention task presented during encoding and retrieval on memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **15,** 1185–1191.

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

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin*, **116,** 220–244.

Rabinowitz, J. C., & Ackerman, B. P. (1982). General encoding of episodic events by elderly adults. In F. I. M. Craik & S. Trehub, *Aging and cognitive processes* (pp. 145–154). New York: Plenum.

Rabinowitz, J. C., Craik, F. I. M., & Ackerman, B. P. (1982). A processing resource account of age differences in recall. *Canadian Journal of Psychology*, **36,** 325–344.

Rankin, J. L., & Kausler, D. H. (1979). Adult age differences in false recognitions. *Journal of Gerontology*, **34,** 58–65.

Raz, N., Millman, D., & Sarpel, G. (1990). Cerebral correlates of cognitive aging: Gray–white matter differentiation in the medial temporal lobes, and fluid versus crystalized abilities. *Psychobiology*, **18,** 475–481.

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

Reinitz, M. T., Morrissey, J., & Demb, J. (1994). Role of attention in face encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **20,** 161–168.

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

Robinson, K. J., & Roediger, H. L., III (1997). Associative processes in false recall and false recognition. *Psychological Science*, **8,** 231–237.

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

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

Rosenthal, R., & Rosnow, R. L. (1991). *Essentials of behavioral research: Methods and data analysis* (2nd ed.). New York: McGraw–Hill.

Schacter, D. L. (1995). Memory distortion: History and current status. In D. L. Schacter, J. T. Coyle, G. D. Fischbach, M. M. Mesulam, & L. E. Sullivan (Eds.), *Memory distortion: How minds, brains, and societies reconstruct the past* (pp. 1–43). Cambridge, MA: Harvard Univ. Press.

Schacter, D. L. (1999). The seven sins of memory: Insights from psychology and cognitive neuroscience. *American Psychologist*, **54,** 182–203.

Schacter, D. L., Israel, L., & Racine, C. (1999). Suppressing false recognition in younger and older adults: The distinctiveness heuristic. *Journal of Memory and Language*, **40,** 1–24.

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., Koutstaal, W., & Norman, K. A. (1997). False memories and aging. *Trends in Cognitive Sciences*, **1,** 229–236.

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

Schacter, D. L., Savage, C. R., Alpert, N. M., Rauch, S. L., & Albert, M. S. (1996). The role of hippocampus and frontal cortex in age-related memory changes: A PET study. *NeuroReport*, **7,** 1165–1169.

Schacter, D. L., Verfaellie, M., & Anes, M. D. (1997). Illusory memories in amnesic patients: Conceptual and perceptual false recognition. *Neuropsychology*, **11,** 331–342.

Schacter, D. L., Verfaellie, M., Anes, M. D., & Racine, C. (1998). When true recognition suppresses false recog-

nition: Evidence from amnesic patients. *Journal of Cognitive Neuroscience*, **10,** 668–679.
Schacter, D. L., Verfaellie, M., & Pradere, D. (1996). The neuropsychology of memory illusions: False recall and recognition in amnesic patients. *Journal of Memory and Language*, **35,** 319–334.
Schmitter-Edgecombe, M. (1996). The effects of divided attention on implicit and explicit memory performance. *Journal of the International Neuropsychological Society*, **2,** 111–125.
Seamon, J. G., Luo, C. R., & Gallo, D. A. (1998). Creating false memories of words with or without recognition of list items: Evidence for nonconscious processes. *Psychological Science*, **9,** 20–26.
Shiffrin, R. M., Huber, D. E., & Marinelli, K. (1995). Effects of category length and strength on familiarity in recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **21,** 267–287.
Smith, A. D. (1975). Partial learning and recognition memory in the aged. *International Journal of Aging and Human Development*, **6,** 359–365.
Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, **117,** 34–50.
Spencer, W. D., & Raz, N. (1994). Memory for facts, source, and context: Can frontal lobe dysfunction explain age-related differences? *Psychology and Aging*, **9,** 149–159.
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.
Toglia, M. P., Neuschatz, J. S., & Goodwin, K. A. (1999). Recall accuracy and illusory memories: When more is less. *Memory*, **7,** 233–256.
Tun, P. A., Wingfield, A., Rosen, M. J., & Blanchard, L. (1998). Response latencies for false memories: Gist-based processes in normal aging. *Psychology and Aging*, **13,** 230–241.
Tussing, A. A., & Greene, R. L. (1997). False recognition of associates: How robust is the effect? *Psychonomic Bulletin & Review*, **4,** 572–576.
West, R. L. (1996). An application of prefrontal cortex function theory to cognitive aging. *Psychological Bulletin*, **120,** 272–292.
Whiting, W. L., & Smith, A. D. (1997). Differential age-related processing limitations in recall and recognition tasks. *Psychology and Aging*, **12,** 216–224.

(Received October 29, 1998)
(Revision received June 18, 1999)