Aging and Strategic Retrieval Processes: Reducing False Memories
With a Distinctiveness Heuristic

Chad S. Dodson
University of Virginia

Daniel L. Schacter
Harvard University

The authors show that a strategic retrieval process—the distinctiveness heuristic—is a powerful mechanism for reducing false memories in the elderly. Individuals studied words, pictures, or both types of items and then completed a recognition test on which the studied items appeared once, whereas the new words appeared twice. After studying either pictures only or a mixture of pictures and words, both younger and older adults falsely recognized fewer repeated new words than did participants who studied words. Studying pictures provided a basis for using a distinctiveness heuristic during the recognition test: Individuals inferred that the absence of memory for picture information indicates that an item is “new.”

During the past several years, evidence has mounted that elderly adults are particularly vulnerable to experiencing false memories, illusory recollections in which people believe that they have earlier encountered an item that is actually novel (Johnson, Hashtroudi, & Lindsay, 1993; Koutstaal & Schacter, 1997; Rankin & Kausler, 1979; Schacter, Koutstaal, & Norman, 1997). Initial investigations of aging and false recognition used a continuous recognition paradigm introduced by Underwood (1965) in which individuals indicate for each of a series of words whether they had encountered it previously in the series. A. D. Smith (1975) and Rankin and Kausler found that older adults were more susceptible than younger adults to falsely recognizing new words that were acoustically or semantically related to earlier studied words, such as mistakenly recognizing the new word gem after having studied hem.

More recently, similar age-related effects have been found with a paradigm developed by Deese (1959) and modified by Roediger and McDermott (1995; see also Read, 1996). In the Deese/Roediger-McDermott (DRM) paradigm, individuals study lists of semantically related words (e.g., bed, tired, dream) and then complete a recognition test containing studied words (e.g., bed), new related words (e.g., sleep), and new unrelated words (e.g., crown). Older adults are more likely than younger adults to falsely recognize related lures, even though both age groups recognize studied words at comparable rates or older adults exhibit reduced true recognition (Balota et al., 1999; Norman & Schacter, 1997; Tun, Wingfield, Rosen, & Blankard, 1998). Koutstaal and Schacter (1997) reported even larger age differences in false recognition using a paradigm in which older and younger adults study a series of pictures from various categories (e.g., cars, shoes). After studying large categories (e.g., 9 or 18 pictures from the same category), older adults made nearly twice as many false alarms to new pictures from studied categories as did younger adults. In summary, there is converging evidence from a variety of paradigms that one consequence of aging is an increase in false recognition responses (Dywan & Jacoby, 1990; Jennings & Jacoby, 1997; Koriat, Ben-Zur, & Sheffer, 1988).

The increased vulnerability of older adults to false memories is attributable, at least in part, to their difficulties remembering source information—specific information that identifies the origin of a memory, such as when and where it occurred (e.g., Johnson et al., 1993). Older adults are generally less accurate than younger adults at remembering which of two speakers presented a novel fact (Bayen & Murnane, 1996; McIntyre & Craik, 1987; Schacter, Kaszniak, Kihlstrom, & Valdiserri, 1991; cf. Ferguson, Hashtroudi, & Johnson, 1992; Hashtroudi, Johnson, Vnek, & Ferguson, 1994) or whether an event was witnessed in a videotape or only seen in a photograph (Schacter, Koutstaal, Johnson, Gross, & Angell, 1997). The elderly also have more difficulty than younger adults in identifying whether an item was earlier seen or imagined (Henkel, Johnson, & DeLeonardis, 1998), was thought about or spoken aloud (Hashtroudi, Johnson, & Chrosniak, 1989), or was presented in uppercase or lowercase letters (Kausler & Puckett, 1980). Moreover, older adults exhibit impaired source memory even when they are able to distinguish study items from new items as well as younger adults (e.g., Henkel et al., 1998; Schacter et al., 1991). Thus, the age-related source memory deficit does not appear to be a by-product of generally worse memory for events overall (see Spencer & Raz, 1995, for a review and meta-analysis).

Remembering source information can be a valuable tool for preventing false recognition errors because it enables individuals to determine why an item seems familiar. For instance, when confronted with the new word gem in the Underwood (1965) paradigm, younger adults may recollect having seen hem earlier and thus conclude that, although gem is familiar because it sounds like hem, it does not match the initially studied word. By contrast,
older adults’ deficiency in recollecting source information would hinder them from identifying why a test item (e.g., gem) is familiar. Consequently, the elderly may be more prone than young adults to basing recognition decisions primarily on an item’s overall familiarity (or plausibility), thereby contributing to their high false recognition rate (e.g., Jacoby, 1999b; Reder, Wible, & Martin, 1986).

Jacoby and colleagues have shown with a number of different paradigms that age-related changes in recollecting source information leave older adults vulnerable to accepting familiar but incorrect events. For example, Jennings and Jacoby (1997; see Jacoby, 1999a, 1999b, for related work) modified a paradigm used earlier by Underwood and Freund (1970; see also Fischler & Juola, 1971, and Koriat et al., 1988) and presented older and younger adults with a list of words to study. All participants then completed an old–new recognition test in which each old word appeared once, but each new word appeared twice with a varying lag between its first and second occurrence. Participants were instructed to respond “old” only when a word had appeared on the study list. In addition, they were told that if a word occurred twice on the test, it was safe to conclude that it could not be a study word because studied words would appear only once. Thus, source recollection was required to oppose any familiarity produced by the repeated presentation of a new word on the recognition test. Older adults responded “old” to repeated new words significantly more often than did younger adults, even when the lag between repetitions was as short as four intervening items, and even though older and younger adults responded “old” equally often to nonrepeated new words. Dywan, Segalowitz, and Webster (1998), using the same test as Jennings and Jacoby (1997), and Koriat et al. (1988), using a slightly different test in which both old and new items repeated, have also found that older adults are more likely than younger adults to make false alarms to repeated new words.

Given their vulnerability to experiencing false memories, an important theoretical and practical question arises: Can the elderly reduce or suppress their false recognition responses? Kensinger and Schacter (1999) addressed this question in a study using the DRM paradigm. Lists of semantic associates were presented and tested five times. Kensinger and Schacter reasoned that repetition would produce more detailed memory for the specific items that appeared on the study lists, which in turn could be used to reduce false recognition for nonstudied associates. Younger adults indeed showed reduced false recognition of semantic associates across repetitions: They made significantly fewer false alarms to related lures on the final trial than on the first trial. By contrast, older adults showed no reduction in false recognition across trials (see also Budson, Daffner, Desikan, & Schacter, 2000). These results suggest that older adults have little ability to use recollection of specific items from a previously studied list to reduce false recognition.

In a separate series of studies, we have identified a strategic retrieval process for reducing misattribution errors that we call the distinctiveness heuristic (Dodson & Schacter, 2001, 2002b; Schacter, Israel, & Racine, 1999): a mode of responding when people expect to remember vivid details of an experience and make recognition decisions on the basis of this metacognitive expectation. When a novel event or item lacks the expected distinctive information, individuals use this absence of critical evidence to reject the item. The distinctiveness heuristic in some respects resembles Einstein, Hunt, McDaniel and Smith’s notion of distinctiveness within their theory of organizational and distinctive processing (e.g., Hunt & Einstein, 1981; Hunt & McDaniel, 1993; R. E. Smith & Hunt, 2001). Although both perspectives emphasize the importance of distinctive information for memory performance, there is also a critical difference: The distinctiveness heuristic is a retrieval mechanism, based on individuals’ beliefs about what they expect to remember. By contrast, Einstein, Hunt, McDaniel, and Smith referred to distinctive processing as an encoding orientation that focuses on specific item or featural information. This orientation allows the item to be distinguished from other items occurring in the same episode. Distinctive processing during encoding sets the stage for the use of the distinctiveness heuristic at retrieval.

We demonstrated the operation of the distinctiveness heuristic in three sets of experiments using the DRM paradigm (Dodson & Schacter, 2001; Schacter, Cendan, Dodson, & Clifford, 2001; Schacter et al., 1999). For instance, Schacter et al. (1999; see also Israel & Schacter, 1997) modified the DRM procedure by auditorily presenting each word in a semantically related list along with a picture of the item. Compared with a condition where participants studied only words (in both visual and auditory modalities), false recognition of related lures was reduced dramatically following pictorial encoding. Schacter et al. (1999) argued that the reduction in false recognition was attributable to participants’ expectation that they should be able to remember the distinctive pictorial information. Thus, the absence of memory for this distinctive information indicated that the test item is new (cf. Rotello, 1999; Strack & Bless, 1994). By contrast, participants who studied words would not expect detailed recollections of studied items, and hence, would not base recognition decisions on the absence of memory for such information.

It is important to note that Schacter et al. (1999) found that older adults were also able to reduce their false recognition rate to related lure words after picture encoding as compared to word encoding. Thus, the distinctiveness heuristic appears to be a tool that older adults can use to reduce false memories. Because this evidence stems entirely from the DRM false recognition paradigm, however, the generality of older adults’ use of the distinctiveness heuristic is unknown. In particular, the DRM paradigm has the notable feature that the study items share much conceptual similarity with each other (see also Koutstaal, Schacter, Galluccio, & Stofer, 1999). However, if the distinctiveness heuristic is a mechanism that can be used broadly across a range of situations, it should be possible to provide convincing evidence for its operation in paradigms where test items are unrelated to previously studied items.

In the DRM paradigm studied by Schacter et al. (1999), age differences in false recognition are relatively modest and sometimes are not observed at all (cf. Norman & Schacter, 1997; Tun et al., 1998). To determine whether older adults can still effectively use the distinctiveness heuristic when age differences in false recognition are more substantial and when new items are unrelated to previously studied ones, we used a repetition lag procedure (e.g., Jennings & Jacoby, 1997; see also Fischler & Juola, 1971; Underwood & Freund, 1970). In this paradigm, individuals study a list of unrelated words and then make old–new recognition judgments about previously studied words and new words. In addition, new words on the recognition test are repeated after
Obtained mean scores of 62.0 (range 59.8–70) on the Vocabulary subtest of the Wechsler Adult Intelligence Scale—Revised (WAIS–R; Wechsler, 1981). The older adults obtained mean scores of 16.7 years of formal education. All but one of the older adults had a history of alcoholism or drug toxicity, primary degenerative disorders (e.g., Alzheimer’s disease, Parkinson’s disease, or Huntington’s disease), and brain damage from a known cause (e.g., hypoxia). Older adults would invoke a distinctiveness heuristic on the recognition test after studying pictures and thus avoid making false alarms to repeated new words.

**Experiment 1**

We began by asking whether older adults would be able to reduce their false recognition rate to repeated new words after picture encoding compared with word encoding. The general idea is that after studying pictures, participants would expect to remember pictorial information during the recognition test. Because repeated new words lack this distinctive information, participants should be able to reduce false alarms to repeated new words by relying on the distinctiveness heuristic. If such an outcome were observed, it would provide evidence for the generality of the distinctiveness heuristic by demonstrating its operation in a novel paradigm where new items are unrelated to previously studied items. Specifically, we expected to replicate Jennings and Jacoby’s (1997) results in the word encoding condition, namely, that older adults would be more likely than younger adults to falsely recognize repeated new words. The central question was whether older adults would invoke a distinctiveness heuristic on the recognition test after studying pictures and thus avoid making false alarms to repeated new words.

**Method**

**Participants.** Thirty-two younger adults and 32 older adults participated in this experiment. The younger adults were recruited from the student population at Harvard University, with a mean age of 19.4 years (range = 17–27) and had, on average, 13.3 years of formal education. The older adults were recruited through flyers and were interviewed to exclude those with any of the following conditions: a history of alcoholism or substance abuse, cerebrovascular accident, recent myocardial infarction, present or previous treatment for psychiatric illness, current treatment with psychoactive medication, metabolic or drug toxicity, primary degenerative disorders (e.g., Alzheimer’s disease, Parkinson’s disease, or Huntington’s disease), and brain damage from a known cause (e.g., hypoxia). Older participants’ mean age was 67.5 years (range = 62–75), and they had, on average, 16.7 years of formal education. All but one of the older adults completed the Vocabulary and Information subtests of the Wechsler Adult Intelligence Scale—Revised (WAIS–R; Wechsler, 1981). The older adults obtained mean scores of 62.0 (range = 50–70) on the Vocabulary subtest and 25.1 (range = 13–29) on the Information subtest. Although these scores were not collected for the younger participants, this sample was drawn from the same population of undergraduates that participated in the study by Schacter, Koutstaal, Johnson, et al. (1997), which provides subtest scores for both Vocabulary (Experiment 1, M = 60.3, range = 45–69; Experiment 2, M = 59.8, range = 45–68) and Information (Experiment 1, M = 24.2, range = 12–29; Experiment 2, M = 24.3, range = 15–28).

Thus, the performance of our older adults on these measures closely resembled that of undergraduates drawn from the same population as those in our study. Sixteen younger and 16 older participants were included in the word encoding and picture encoding conditions. Data concerning the younger participants are reported separately by Dodson and Schacter (2002b).

**Design and materials.** The stimuli consisted of 120 pictures and their corresponding verbal labels, used by Snodgrass and Vanderwart (1980). Sixty items were studied and also served as the old items on the test. The remaining 60 items were the new items on the test. Each new item was repeated at either Lag 24 or Lag 48. The stimuli were divided into four lists of 30 items. The lists were balanced so that they had similar mean ratings for picture familiarity (range = 3.5–3.6), picture complexity (range = 2.6–2.8), and word frequency (33). Two lists were presented at study, and two lists were presented as new items on the recognition test at the two lag intervals. Four different counterbalancing formats rotated the lists of items so that across participants each list appeared at study and also was presented as a new word in each of the lag conditions at test.

An Apple G3 (Cupertino, CA) computer presented all the stimuli in the center of the screen. The pictures were approximately the same size and fit within a 6 in × 6 in (15.2 cm × 15.2 cm) area of the screen. Each picture was also accompanied with the auditory presentation of its name. The words appeared in lowercase, 48-point letters in the Geneva font. For each study item, the phrase “How many syllables?” appeared at the bottom of the screen. After a response, the screen cleared and was followed by a 1-s delay before the presentation of the next study item. The study items were randomly intermixed with the restriction that each third of the study list contained an equivalent number of list items. Similarly, the order of the test items was random with the restriction that no more than three old or new items could occur consecutively. The test was identical for all participants and contained either the names of earlier studied pictures or words, mixed together with new words; no pictures were presented on the test. The test words were presented in lowercase, in 48-point letters in the Geneva font. All new words were repeated after either 24 or 48 intervening items.

**Procedure.** We assigned participants to either the word encoding condition or the picture encoding condition. In contrast to the intentional study instructions used by Jennings and Jacoby (1997), we gave everyone incidental instructions and told them to enter as fast as possible the number of syllables in the study item. No mention was made of a later memory test. We informed participants in the picture encoding condition that they would see a series of pictures, one at a time, and that each would be accompanied with the auditory presentation of its name.

Immediately after the study phase, everyone received the test instructions. We told the participants that the test was based on their memory for all of the studied items. Those who studied pictures were informed that the test would contain names of the picture study items. In addition, we told all of them that the test would contain new words and that these new words would be repeated so that they would appear on two different occasions. We instructed them to respond old to the studied items only and new to the new items by pressing the “a” and the “z” keys, respectively. We emphasized to participants that they should respond new to the repeated new words. We cautioned them against mistaking a repeated new word for a studied item.

**Results and Discussion**

Table 1 displays the probabilities of responding “old” in the two different encoding conditions to studied items, new words, and repeated new words at the two lag intervals. The results reveal four notable patterns. First, the hit rates to studied items and false alarm rates to new words on their first occurrence (hereafter referred to as baseline false-alarm rates) are generally lower for older than younger adults. Second, signal detection measures of d’ (discrimination) and C (bias) indicate that these hit and false-alarm-rate
significant effect of study condition, ANOVA of the baseline false-alarm rates revealed a significant effect of study condition, \(F(1, 60) = 6.23, MSE = .027, p < .05, \omega^2 = .08, \) and no other significant effects. As is evident in Table 1, younger adults exhibit higher hit rates than do older adults. A 2 (age) \( \times \) 2 (study condition) ANOVA of the baseline false-alarm rates revealed a significant effect of study condition, \(F(1, 60) = 6.23, MSE = .011, p < .05, \omega^2 = .08, \) indicating that both younger and older adults tended to give false alarms more to new words (on their first occurrence) after having studied words than after studying pictures.

For item recognition, we examined hit rates to studied items and baseline false-alarm rates to new words. A 2 (age: young vs. elderly) \( \times \) 2 (study condition: picture vs. word) analysis of variance (ANOVA) of the hit rates yielded a significant effect of age, \(F(1, 60) = 6.37, MSE = .027, p < .05, \omega^2 = .08, \) and no other significant effects. As is evident in Table 1, younger adults exhibit higher hit rates than do older adults. A 2 (age) \( \times \) 2 (study condition) ANOVA of the baseline false-alarm rates revealed a significant effect of study condition, \(F(1, 60) = 6.23, MSE = .011, p < .05, \omega^2 = .08, \) indicating that both younger and older adults tended to give false alarms more to new words (on their first occurrence) after having studied words than pictures. There were no other significant effects in the analysis. Overall, then, older adults responded more conservatively than did younger adults, but there were no age-related differences in distinguishing studied items from once-presented new items.

Figure 1 presents the corrected false recognition rates to the repeated new words in the different conditions. These scores represent the false recognition rates to repeated new words after subtracting out the false recognition rate to new words on their first occurrence. The line at zero, for example, is the point at which there is no difference between the false-alarm rate for repeated new words and the baseline false-alarm rate. As expected, older adults in the word encoding condition were much more vulnerable than younger adults in this condition to falsely recognizing repeated new words, replicating similar results found by Jennings.

<table>
<thead>
<tr>
<th>Age and study condition</th>
<th>Studied</th>
<th>New</th>
<th>(d')</th>
<th>C</th>
<th>Lag 24</th>
<th>Lag 48</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Younger Word</td>
<td>0.62</td>
<td>0.13</td>
<td>0.18</td>
<td>0.11</td>
<td>1.30</td>
<td>0.47</td>
</tr>
<tr>
<td>Picture</td>
<td>0.62</td>
<td>0.16</td>
<td>0.12</td>
<td>0.09</td>
<td>1.64</td>
<td>0.75</td>
</tr>
<tr>
<td>Older</td>
<td>0.50</td>
<td>0.13</td>
<td>0.15</td>
<td>0.12</td>
<td>1.14</td>
<td>0.47</td>
</tr>
<tr>
<td>Picture</td>
<td>0.53</td>
<td>0.23</td>
<td>0.09</td>
<td>0.10</td>
<td>1.58</td>
<td>0.46</td>
</tr>
</tbody>
</table>

*Note.* The signal detection measures, \(d'\) (discrimination) and C (bias), are derived from the “old” responses to studied and new items.

---

**Table 1**

*Probabilities of Responding “Old” to Studied Items, New Words, and Repeated New Words at Each of the Lag Intervals in Experiment 1*

![Figure 1](image-url)
and Jacoby (1997). Likewise, we replicated previous findings that younger adults can successfully reject repeated new words after picture encoding (Dodson & Schacter, 2002b). Picture encoding also allowed older adults to greatly improve their performance, relative to the word encoding group, by reducing their corrected false recognition rate to repeated new words.

We examined the corrected false recognition rates with a 2 (age) × 2 (study condition) × 2 (lag: 24 vs. 48) ANOVA, which yielded significant main effects of age, $F(1, 60) = 10.15, MSE = .024, p < .01, \omega^2 = .10$, and study condition, $F(1, 60) = 16.58, MSE = .024, p < .001, \omega^2 = .17$. Younger adults showed lower corrected false recognition rates than did older adults. Both groups of participants, however, falsely recognized fewer repeated new words after picture than word encoding. In addition to the above significant main effects, there was a marginally significant effect of lag, $F(1, 60) = 2.95, MSE = .006, p < .10, \omega^2 = .03$, and a marginally significant Lag × Age interaction, $F(1, 60) = 3.04, MSE = .006, p < .10, \omega^2 = .03$. As seen in Figure 1, younger adults showed a greater probability of falsely recognizing new words that repeated at Lag 48 than Lag 24, $F(1, 60) = 11.98, MSE = .006, p < .01$. Presumably at the longer lag intervals there is an increasing likelihood that younger adults forget that the repeated new word was encountered before and mistake its familiarity (derived from its earlier exposure on the test) for prior presentation in the study phase. By contrast, older adults did not exhibit a tendency to falsely recognize more Lag 48 than Lag 24 repeated new words, $F(1, 60) < 1$. This finding is likely attributable to older adults reaching their forgetting asymptote, and thus, their maximum corrected false recognition rate, well before Lag 24. Indeed, Jennings and Jacoby (1997) found that older adults’ false recognition rate of repeated new words increased dramatically from Lag 0 to Lag 4 and remained constant thereafter to Lag 48. Thus, our results are consistent with previous findings.

In addition, Figure 1 shows a nonsignificant Age × Study Condition interaction, $F(1, 60) = 1.26, \omega^2 = 0$, which reflects the pattern that the elderly, as compared to younger adults, are more error-prone after word than picture encoding.

As we discuss in the General Discussion, this pattern in the two encoding conditions reflects a combination of relatively impaired recollective abilities on the part of the elderly and relatively spared use of the distinctiveness heuristic. Specifically, after word encoding the elderly are particularly vulnerable to falsely recognizing repeated new words because (a) they have difficulty recollecting, and thus, rejecting earlier seen new words, and (b) after word encoding there is no basis for using the distinctiveness heuristic. By contrast, after studying pictures, the young and elderly have comparable corrected false recognition rates because both groups of individuals can use the distinctiveness heuristic to reject repeated new words.

In summary, the key result from this experiment is that older adults who studied pictures were able to use the distinctiveness heuristic to successfully reject repeated new words. In fact, older and younger adults in the picture encoding condition showed very similar corrected false recognition rates of the new words that repeat at Lag 48 (6% vs. 8%, respectively). These results indicate that the distinctiveness heuristic is a powerful mechanism for reducing age-related false memory effects.

Experiment 2

Experiment 1 demonstrated that older adults can use the distinctiveness heuristic to reduce their false recognition responses to repeated new words. One central question about this heuristic concerns the conditions that elicit it: How diagnostic—predictive of a test item’s oldness—must the distinctive study information be for participants to use the distinctiveness heuristic? In the picture encoding condition in Experiment 1, for example, memory for pictorial information was perfectly diagnostic of an item’s oldness. Failing to remember pictorial information about a test item was thus a highly reliable indicator that the item was novel. Do older and younger adults invoke the distinctiveness heuristic when pictorial information is less than perfectly diagnostic?

Although earlier evidence from the DRM paradigm by Schacter et al. (1999) indicated that older and younger adults abandon the distinctiveness heuristic when it is not perfectly diagnostic, this conclusion is questionable and needs reevaluation because of a confound in the DRM paradigm. Schacter et al. used a between-groups design and showed that older and younger adults falsely recognized fewer related lure words when they studied all of the lists of semantically related items as pictures than as words. However, when the diagnosticity of the pictorial information was decreased by presenting only half of the study lists of related items as pictures and the remaining lists as words (i.e., a within-groups design), Schacter et al. observed no difference in the false recognition rates of lures that were related to lists encoded as pictures as compared to those that were related to lists encoded as words. Notice that the distinctive information is less than perfectly diagnostic in this within-groups design because failing to remember pictorial information about a test item does not mean that it is novel. Instead, the test item may have been studied in one of the lists presented as words. Thus, these results suggest that older and younger adults abandon the distinctiveness heuristic when it is less than perfectly diagnostic.

However, there is a confounding feature in the within-groups version of the DRM paradigm that questions the interpretation of the Schacter et al. (1999) results. Because half of the lists were studied as pictures in this within-groups design, participants may try to apply the distinctiveness heuristic selectively to the test items related to the lists studied as pictures (e.g., Did I see this test item in one of the pictured lists?). If individuals forget which test items were studied as pictures and which were studied as words, then instead of abandoning the distinctiveness heuristic because of its lack of diagnosticity, participants may use the heuristic globally for all test items. Such a global use of the distinctiveness heuristic would suppress the false recognition rate to all of the lures and thus produce the Schacter et al. (1999) finding of no difference in the false recognition rate to lures related to lists presented as pictures or words. Indeed, Dodson and Schacter (2001) and Schacter et al. (2001) reported evidence from their experiments, and those of Schacter et al. (1999), indicating that participants used the distinctiveness heuristic in the within-groups version of the DRM paradigm, even though the distinctive information is not diagnostic of prior study. They noted that if participants had abandoned the distinctiveness heuristic, then one would expect false recognition
rates for the related lures to be at such a high level that they would be comparable to false recognition rates from those participants who had studied all of the lists as words (i.e., the word-only condition in the between-groups design that typically yields a robust false recognition rate). This outcome was not observed. Participants who studied half of the lists as pictures and half as words later falsely recognized fewer critical lures that were related to either picture or word lists, as compared to the false recognition rate of participants who studied all of the lists as words. However, because the apparent use of the distinctiveness heuristic in the within-groups condition may be attributable to the aforementioned memory limitation (i.e., failure to remember list presentation mode), data from the DRM paradigm cannot speak conclusively to the importance of diagnosticity for use of the distinctiveness heuristic.

By contrast, the repetition lag paradigm provides a more direct test of the role of diagnostic information in eliciting the distinctiveness heuristic than does the DRM paradigm, because the former procedure is not confounded by the memory-for-list-presentation-mode issue that blurs interpretation of results from the latter procedure. That is, in the repetition lag procedure repeated new words are not associated to any specific items from the study list, in the same sense that related lure words are linked to specific study lists in the DRM procedure.

In Experiment 2, we examined the role of diagnosticity in eliciting the distinctiveness heuristic in older and younger adults by comparing performance in two different conditions. In the word encoding condition, participants studied all of the items as words. In a second condition, they studied half of the items as pictures and half as words. In this latter 50% picture encoding condition, failing to remember pictorial information about an item is not conclusive evidence of novelty, because the item may be one that was studied as a word. However, if participants do reduce their false recognition rates to repeated new words in the 50% picture condition, compared with the word-only encoding condition, then this result would suggest that diagnosticity is not an important factor for activating the distinctiveness heuristic.

A further change in the procedure from Experiment 1 is that in the word encoding condition, all of the words were presented at study in both the visual and auditory modalities (i.e., participants both heard the name of the word and saw its visual form). This change means that the auditory presentation is constant for all the stimuli in both the word and 50% picture conditions.

**Method**

Participants. Thirty-two younger and 32 older adults participated in this experiment; they had not participated in Experiment 1. The younger adults were recruited from the student population at Harvard University, with a mean age of 19.5 (range = 17–25) and had, on average, 13.3 years of formal education. The older adults were recruited through flyers and were interviewed to exclude those who fit the criteria detailed in Experiment 1. Older participants’ mean age was 68.5 (range = 62–75), and they had, on average, 16.6 years of formal education. All of the older adults completed the Vocabulary and Information subtests of the WAIS–R and received mean scores of 61.3 (range = 50–70) and 24.7 (range = 18–29) on the two subtests, respectively. These data were not collected from the younger participants but, as noted earlier, Schacter, Koutstaal, Johnson, Gross, & Angell (1997) provided this information for a sample of young participants that was drawn from the same population as our sample. Sixteen younger and 16 older participants were included in the word encoding and 50% picture encoding conditions. Data from younger participants in the word encoding condition are reported separately by Dodson and Schacter (in press-b).

Design and Materials. The stimuli were identical to those used in Experiment 1. The 120 stimuli were divided into 4 different lists of 30 items. Two of the lists were presented at study and two lists served as new items that repeated at either Lag 24 or Lag 48. The word-encoding condition was the same as the corresponding condition in Experiment 1 in all aspects except that participants heard the name of the word in addition to seeing its visual form. In the 50% picture condition, one of the study lists was presented as pictures and the other as words (i.e., 30 pictures and 30 words). As in Experiment 1, the picture was accompanied with the auditory presentation of its name. The study lists and test lists were constructed in the same manner as in Experiment 1.

Procedure. We used the same study and test instructions that we used in Experiment 1. The participants in the 50% picture condition were told at study that they would see a mixture of pictures and words.

### Results and Discussion

Table 2 presents the probabilities of responding “old” to the various types of test items (i.e., studied, new and repeated new) in the two different encoding conditions. Item recognition (i.e., hit rates to studied items and baseline false-alarm rates) was similar

<table>
<thead>
<tr>
<th>Age and study condition</th>
<th>Studied</th>
<th>New</th>
<th>d’</th>
<th>C</th>
<th>Lag 24</th>
<th>Lag 48</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M/SD</td>
<td>M/SD</td>
<td>M/SD</td>
<td>M/SD</td>
<td>M/SD</td>
<td>M/SD</td>
</tr>
<tr>
<td>Younger Word</td>
<td>0.61/0.17</td>
<td>0.17/0.10</td>
<td>1.31/0.59</td>
<td>0.34/0.35</td>
<td>0.24/0.19</td>
<td>0.31/0.18</td>
</tr>
<tr>
<td>50% picture</td>
<td>0.53/0.15</td>
<td>0.12/0.11</td>
<td>1.37/0.61</td>
<td>0.59/0.34</td>
<td>0.11/0.08</td>
<td>0.13/0.11</td>
</tr>
<tr>
<td>Older Word</td>
<td>0.60/0.12</td>
<td>0.21/0.13</td>
<td>1.18/0.48</td>
<td>0.32/0.37</td>
<td>0.48/0.20</td>
<td>0.48/0.20</td>
</tr>
<tr>
<td>50% picture</td>
<td>0.58/0.11</td>
<td>0.17/0.15</td>
<td>1.34/0.65</td>
<td>0.47/0.44</td>
<td>0.31/0.20</td>
<td>0.30/0.22</td>
</tr>
</tbody>
</table>

**Note.** The signal detection measures, d’ (discrimination) and C (bias), are derived from the “old” responses to studied and new items.
across both age groups and in both study conditions. As in Experiment 1, in the word encoding condition we observed a large increase (especially by the older adults) in false recognition responses to the repeated new words, relative to the baseline false-alarm rate. After studying 50% of the items as pictures, both younger and older adults suppressed their false recognition rate to repeated new words.

We assessed item recognition with 2 (age) \times 2 (study condition) ANOVAs of the hit and false-alarm rates. There were no significant effects; all $F(1, 60) < 2.40$, $p > .12$. In addition, a 2 (age) \times 2 (study condition) ANOVA of the $d'$ (discrimination) scores yielded no significant effects, all $F(1, 60) < 1$. However, a 2 (age) \times 2 (study condition) ANOVA of the bias, C, scores revealed a significant effect of study condition, $F(1, 60) = 4.55$, $MSE = .142$, $p < .05$, $\omega^2 = .05$, and no other significant effects. Participants responded more conservatively to study and new items in the 50% picture encoding condition (.53) than in the word encoding condition (.33). Overall, as in Experiment 1, there were no age-related differences in distinguishing between studied and new words that were seen for the first time.

Figure 2 presents the corrected false recognition rates of the repeated new words at each of the lag intervals. A 2 (age) \times 2 (study condition) \times 2 (lag) ANOVA of the corrected false recognition rates yielded significant effects of age, $F(1, 60) = 22.90$, $MSE = .03$, $p < .0001$, $\omega^2 = .22$, and study condition, $F(1, 60) = 15.72$, $MSE = .03$, $p < .001$, $\omega^2 = .15$. Overall, older adults were more vulnerable than younger adults to falsely recognizing repeated new words. However, both older and younger adults successfully suppressed their false recognition responses to the repeated new words after studying 50% of the items as pictures. In addition to the above main effects, there was a marginally significant Age \times Lag interaction, $F(1, 60) = 3.23$, $MSE = .006$, $p < .10$, $\omega^2 = .03$; there were no other significant effects in this analysis. As in Experiment 1, younger adults were more likely to falsely recognize Lag 48 than Lag 24 new words, $F(1, 60) = 9.66$, $MSE = .006$, $p < .01$, whereas older adults showed no difference in their corrected false recognition rates of the new words across the two lag intervals, $F(1, 60) < 1$. As we noted in Experiment 1, younger adults, in contrast to older adults, were more likely to recollect and thus reject new words that repeated at Lag 24 than those that repeated at Lag 48. Having already reached their forgetting asymptote, older adults were no more likely to recollect Lag 24 new words than Lag 48 new words. Figure 2 shows a nonsignificant trend toward an Age \times Study Condition interaction, $F(1, 60) < 1$, similar to that shown in Experiment 1. As we discuss in the next section, this pattern reflects a combination of two factors: (a) elderly adults’ relatively spared use of the distinctiveness heuristic after picture encoding, which minimizes false recognition errors between the young and the old, and (b) the elderly’s impaired ability to recollect and reject the repeated new words, which contributes to their high corrected false recognition rate in the word encoding condition.

General Discussion

The central finding of our experiments is that older adults are able to use a retrieval strategy that we call the distinctiveness heuristic to reduce false recognition errors under conditions in which new items are unrelated to previously studied items, and older adults are highly vulnerable to false recognition. In two experiments, younger and older adults studied words, pictures, or a mixture of both and then completed a recognition test in which studied items appeared once and new items appeared twice. Both experiments demonstrated that in a word encoding condition, older adults were more likely than younger adults to falsely recognize repeated new words, replicating the findings of Dywan et al. (1998), and Jennings and Jacoby (1997), as well as those of Koriat et al. (1988) who used a slightly modified version of the repetition lag paradigm (i.e., both old and new words repeated at test). Studying pictures, however, produced a large reduction in the false recognition rates to repeated new words for both older and younger adults. In fact, in the picture encoding condition of Experiment 1, corrected false recognition rates of new words that repeated at Lag 48 were no different for older and younger adults. We interpret this false recognition suppression effect in the picture encoding condition as the by-product of a metacognitive heuristic in which individuals infer that a test item is novel when they do not remember expected pictorial information. Our finding that the elderly can reduce their false alarms to repeated new words after picture encoding conceptually replicates the results of Schacter et al. (1999) in the DRM paradigm, who found that older adults falsely recognized fewer related lures after picture than word encoding.

These experiments also examined the role of diagnosticity for eliciting the distinctiveness heuristic by varying the proportion of studied items that appeared as pictures. Compared with a word encoding condition, in Experiment 2 both younger and older participants successfully rejected repeated new words after studying 50% of the items as pictures. In a related study using the same paradigm, we have shown that studying as few as 25% of the items as pictures is sufficient for younger adults to activate the distincti-
tiveness heuristic (Dodson & Schacter, 2002b). Thus, the distinctive information need not be completely diagnostic (i.e., perfectly predictive of an item’s oldness) for either younger or older adults to use the distinctiveness heuristic. An issue that requires investigation in future research concerns the reasons why encoding a particular source (e.g., Anderson, 1984; Foley, Johnson, & Raye, 1983; Hashtroudi et al., 1989; Hicks & Marsh, 1999; Johnson, Raye, Foley, & Foley, 1981; Kelley, Jacoby, & Hollingshead, 1989). For instance, Johnson et al.’s (1981) “It had to be you” effect refers to a test bias in which individuals who have earlier heard some words and generated others are more likely to respond that falsely recognized new words were earlier heard than generated. This bias presumably reflects the metamnemonic belief that self-generated information is more memorable than heard information (Johnson & Raye, 1981). Consequently, when in doubt, a familiar test item is judged to have been heard earlier because of the absence of memory for having generated the item. The distinctiveness heuristic is also compatible with the monitoring processes discussed by several investigators, including Schacter, Norman, and Koutstaal (1998) in their constructive memory framework, and Roediger, McDermott and colleagues’ activation/monitoring account of performance in the DRM paradigm (e.g., McDermott & Watson, 2001; Roediger, Watson, McDermott, & Gallo, 2001). Specifically, drawing on Johnson et al.’s (1993) source monitoring framework as well as Jacoby, Kelley, and Dywan’s (1989) attribution approach to remembering, the activation/monitoring account posits that a variety of processes, such as those that allow memories for internally generated events to be distinguished from memories for externally derived events, help to suppress the occurrence of false recall and recognition. Hicks and Marsh (1999) have shown that a decision strategy, based on the absence of memory for expected source information, allows participants to reduce the rate of false recall of related lures in the DRM paradigm. Similar findings from other paradigms suggest that the lack of expected memory or knowledge about a particular event is used as a guide when making a response (e.g., Brewer & Treyens, 1981; Brown, Lewis, & Monk, 1977; Collins, Warrnc, Aiello, & Miller, 1975; Gentner & Collins, 1981; Strack & Bless, 1994; cf. Rotello, 1999; Wixted, 1992; see Dodson & Schacter, 2002a, 2002b, for further discussion of the distinctiveness heuristic in relation to similar retrieval strategies). In summary, the distinctiveness heuristic is an instance of a general class of metacognitive strategies in which the failure to remember expected information is a signal for the event’s nonoccurrence.

The repetition lag paradigm is a useful tool for analyzing the contributions of three potentially separable processes: familiarity, memory for source specifying information, and the distinctiveness heuristic. Familiarity contributes to the false recognition of repeated new words when participants do not remember encountering the new word earlier on the test and do not invoke the distinctiveness heuristic. Recollecting source, or item-specific, information about seeing the new word earlier on the test serves as a “recall-to-reject” mechanism and thereby can contribute to reducing the occurrence of false recognition responses (for discussion of recall-to-reject processes, see Clark & Gronlund, 1996; Rotello & Heit, 1999; Rotello, Macmillan, & Van Tassel, 2000). In younger adults, the recall-to-reject process often occurs when the new words repeat after a short interval, such as Lag 4 or Lag 12 (e.g., Dodson & Schacter, in press-b; Jennings & Jacoby, 1997). This recollection process can help to explain the finding that the false recognition rate to the new words that repeat at short lags of 4 or 12 is no different from the baseline false recognition rate. However, new words that repeat at the longer lags, such as Lag 24 or Lag 48, are falsely recognized significantly more often than the baseline rate because individuals do not recollect seeing the word earlier on the test and misattribute the item’s familiarity for having studied it previously. Finally, the distinctiveness heuristic is invoked when participants encounter a familiar test word, such as a repeated new word, and they do not recollect source information about either seeing the item earlier on the test or studying the item. In this situation, an item is presumed to be novel when it does not elicit expected memory information.

In terms of the foregoing three factors, familiarity is likely intact in elderly adults and, when unchecked by opposing influences, can lead to increased false recognition (e.g., Jennings & Jacoby, 1997; Schacter, Koutstaal, & Norman, 1997). Elderly adults are clearly impaired in their ability to use a recall-to-reject process, that is, correctly rejecting repeated new words on the basis of recollecting the word earlier on the test. Indeed, Jennings and Jacoby (1997) found that after a word-encoding condition, older and younger adults showed equivalent false recognition rates to new words only when they repeated at Lag 0 (i.e., a consecutive repetition of the new word). When new words repeat after a lag even as short as 4 items, older but not younger adults are prone to falsely recognizing the repeated new words significantly above the baseline rate, thereby indicating a failure of the recall-to-reject process in the elderly individuals. In addition, the elderly’s impaired ability to recall and reject repeated new words contributed to their relatively high false recognition rate to the repeated new words, relative to the young, in the word encoding condition, which contributed to the aforementioned trends toward an Age × Study Condition interaction shown in Figures 1 and 2. A similar failure of a recall-to-reject process may also account for the findings discussed earlier that older adults do not suppress false recognition in the DRM paradigm when lists are studied and tested repeatedly (Kensinger & Schacter, 1999). Older adults’ diminished capacities to use the recall-to-reject process are consistent with the many studies showing impaired source memory in the old, relative to the young (e.g., Johnson et al., 1993; see Spencer & Raz, 1995, for a review).

Our study, together with the findings of Schacter et al. (1999), shows that despite their apparent inability to use recall-to-reject as a mechanism for reducing false recognition, older adults can suppress the occurrence of false recall and recognition. Hicks and Marsh (1999) showed that the lack of expected memory or knowledge about a particular item to reject new items, such as remembering that Houston was studied in order to reject the related lure, Phoenix.
nonetheless use the distinctiveness heuristic to accomplish this task. The distinctiveness heuristic consists of three components: (a) monitoring or noticing what has been remembered about a test item, (b) expectations about what should be remembered if the test item had been studied in a distinctive manner (e.g., as a picture), and (c) responding on the basis of contrasting the first two components.

Because the distinctiveness heuristic depends on metacognitive processes such as monitoring, it is important to relate our findings to other observations concerning aging and metacognition. Nelson and colleagues have argued that a key component of metacognition involves monitoring one’s current state in order to guide behavior, such as predicting subsequent memory performance on the basis of the present phenomenal state (Barnes, Nelson, Dunlosky, Mazzoni, & Narens, 1999; Nelson, 1996; Nelson & Narens, 1990). Older and younger adults generally exhibit comparable monitoring skills (e.g., Butterfield, Nelson, & Peck, 1988; Connor, Dunlosky, & Hertzog, 1997; see Hertzog & Hultsch, 2000, for review). For instance, the accuracy of judgments of learning—predictions during encoding about the probability of subsequently remembering an item—typically does not differ between younger and older adults (Connor et al., 1997). Similarly, the accuracy of feeling-of-knowing judgments—the predicted likelihood of recognizing a currently unrecognizable item—is comparable in the young and old (Butterfield et al., 1988; although see Souchay, Isingrini, & Espagnol, 2000, for age differences in feeling-of-knowing judgments about episodic memory). Hence, our results showing preserved use of the distinctiveness heuristic in the elderly are consistent with studies showing preserved monitoring skills on the part of the elderly.

In addition to requiring preserved monitoring abilities, effective use of the distinctiveness heuristic also requires that individuals expect to remember distinctive information following encoding conditions, such as the picture condition used here, that promote encoding of distinctive information. We are not aware of any research that has specifically examined what older adults expect to remember following distinctive (or nondistinctive) encoding conditions. However, one reasonable inference from our findings of reduced false recognition is that older adults expect to remember the same kinds of vivid, distinctive information following encoding of pictures that younger adults do. That older adults showed just as much reduction of false recognition as younger adults in the 50% picture condition indicates that distinctive information need not be perfectly diagnostic for older adults to invoke the distinctiveness heuristic. Further research is needed to clarify the relation between expectations about what should be remembered and the diagnostic value of distinctive information in both older and younger adults.

In summary, much research has documented the costs to memory from aging (e.g., Johnson et al., 1993; Ka'sker, 1994; Light, 1991; Salthouse, 1991; Schacter, Koutstaal, & Norman, 1997). By contrast, we demonstrate that the strategic retrieval process known as the distinctiveness heuristic appears preserved in elderly adults. Our data show clearly that elderly adults can invoke the distinctiveness heuristic even when study and test items are unrelated and even when distinctive information is not perfectly diagnostic of prior study. Considered together with previous research concerning aging and the distinctiveness heuristic (Schacter et al., 1999), our data indicate that this heuristic is a powerful and general mechanism for reducing the occurrence of false memories in older and younger adults.

References


Received March 13, 2001
Revision received October 23, 2001
Accepted October 30, 2001

---

**Call for Nominations**

The Publications and Communications (P&C) Board has opened nominations for the editorships of *Contemporary Psychology: APA Review of Books, Developmental Psychology*, and *Psychological Review* for the years 2005–2010. Robert J. Sternberg, PhD, James L. Dannemiller, PhD, and Walter Mischel, PhD, respectively, are the incumbent editors.

Candidates should be members of APA and should be available to start receiving manuscripts in early 2004 to prepare for issues published in 2005. Please note that the P&C Board encourages participation by members of underrepresented groups in the publication process and would particularly welcome such nominees. Self-nominations are also encouraged.

Search chairs have been appointed as follows:

- **Contemporary Psychology: APA Review of Books**: Susan H. McDaniel, PhD, and Mike Pressley, PhD
- **Developmental Psychology**: Joseph J. Campos, PhD
- **Psychological Review**: Mark I. Appelbaum, PhD

To nominate candidates, prepare a statement of one page or less in support of each candidate. Address all nominations to the appropriate search committee at the following address:

Karen Sellman, P&C Board Search Liaison
Room 2004
American Psychological Association
750 First Street, NE
Washington, DC 20002-4242

The first review of nominations will begin November 15, 2002. The deadline for accepting nominations is November 25, 2002.