Accuracy and Qualities of Real and Suggested Memories: Nonspecific Age Differences

Mara E. Karpel,1 William J. Hoyer,1 and Michael P. Toglia2

1Department of Psychology, Syracuse University, New York.
2State University of New York at Cortland.

This study examined adult age differences in the accuracy, confidence ratings, and vividness ratings of veridical and suggested memories. After seeing either one or two exposures of a vignette depicting a theft, young adults (M = 19 years) and older adults (M = 73 years) were given misleading information that suggested the presence of particular objects in the episode. Memory accuracy was higher for younger adults than for older adults, and the frequency of falsely reporting the presence of suggested objects was greater for older adults than for young adults. Further, levels of confidence and vividness ratings of the perceptual attributes (colors, locations) of falsely recognized items were higher for older adults than for young adults. Both young adults and older adults used more perceptual references when describing veridical memories than when describing suggested memories. Age differences in the suggestibility of memory were attributed to nonspecific or nondissociated memory aging effects.

The memory is sometimes so retentive, so serviceable, so obedient—at others, so bewildered and so weak—and at others again, so tyrannic, so beyond control!

—Jane Austin, Mansfield Park

In contrast to classic anecdotes and our intuitions about the dynamics of memory, it has been suggested that it may not be necessary to advance specific mechanisms to account for what appear to be different memory functions. In the study of self as a memory system, for example, Greenwald and Banaji (1989) demonstrated that the powerful mnemonic effects associated with self reference are best understood in terms of ordinary memory processes (for a meta-analysis, see Symons & Johnson, 1997). It has also been argued that it is unnecessary and inaccurate to explain memory aging phenomena in terms of specific processes because most of the age variance in memory performance can be attributed to a general aging factor (Salthouse, 1996). The evidence suggesting that memory aging is largely a nonspecific or process—general phenomenon comes mainly from recent meta-analytic aggregations of effect sizes from cross-sectional studies (La Voie & Light, 1994; Verhaeghen & Marcoen, 1993; Verhaeghen, Marcoen, & Goossens, 1993). In the Verhaeghen, Marcoen, and Goossens (1993) meta-analysis, the mean weighted effect size for the age difference between younger and older adults in recall from episodic memory was 0.99. In other words, there was an average age difference across studies using different measures of episodic memory and different subject samples of about one standard deviation (see Cohen, 1988). Similarly, La Voie and Light (1994) reported a mean weighted effect size of 0.97 for episodic recall and a somewhat smaller effect size for episodic recognition.

Descriptions of the effects of aging on episodic memory are drawn mainly from studies comparing the performance of younger and older adults in terms of the ability to remember previous occurrences of events or ideas (for reviews, see Craik & Jennings, 1992; Hultsch & Dixon, 1990; Spencer & Raz, 1995; Verhaeghen & Salthouse, 1997; Zacks, Hasher, & Li, 2000). However, little is known about the extent to which episodic memory is malleable and susceptible to distortion with age. Using a variety of experimental paradigms, it has been demonstrated that individuals of all ages report remembering events or items that were never encountered but that were implanted by suggestion (Pezdek & Banks, 1996). Further, people often report a high degree of confidence in implanted and imagined memories (Bothwell, Defenbacher, & Brigham, 1987; Tversky & Tuchin, 1989; Wells & Murray, 1984).

Much of the research on the suggestibility of memory has used the “misleading post-event information paradigm” originally developed by Elizabeth Loftus and her colleagues (Loftus, 1979; Loftus & Hoffman, 1989; Loftus, Miller, & Burns, 1978), and revised by others (Belli, Windschitl, McCarthy, & Winfrey, 1992; McCloskey & Zaragoza, 1985). In this procedure, individuals are subtly given misleading information about some of the details of an event such as a theft or an automobile accident that was previously shown in a video or slide sequence. The misleading post-event information impairs the memory report of the event, and there is some evidence to suggest developmental differences in the susceptibility to misleading information (Adams-Price, 1992; Brainerd & Reyna, 1988; Ceci, Ross, & Toglia, 1987; Ceci, Toglia, & Ross, 1988; Cohen & Faulkner, 1989; Coxon & Valentine, 1997; Foley & Johnson, 1985; List, 1986; Multhaup, De Leonardis, & Johnson, 1999; Searcy, Bartlett, & Memon, 1999). An age-related increase in memory suggestibility has been found in some of the developmental studies that compare younger adults and older adults. However, the age effects are typically small, and it is not clear if older adults are disproportionately affected by
suggestion when age-related differences in veridical memory are taken into account. Similarly, modest effects have been obtained in recent studies of aging and false memory (Balota et al., 1999; Koutstaal, Schacter, Galluccio, & Stofer, 1999; Norman & Schacter, 1997; Schacter, Koutstaal, & Norman, 1997; Tun, Wingfield, Rosen, & Blanchard, 1998), using a paradigm introduced by Deese (1959) and revived by Roediger and McDermott (1995).

A related question that has recently received some empirical attention is whether there are age-related differences in the qualitative characteristics of real and suggested memories (Hashtroudi, Johnson, & Chrosniak, 1989; Koutstaal, Schacter, Johnson, Angell, & Gross, 1998; Norman & Schacter, 1997). In studies with college students, it has been reported that veridical memories contain more details, more perceptual information, and more spatial and temporal information, and that suggested or imagined memories contain more information about the cognitive operations involved in creating the memory (Johnson, Foley, Suengas, & Raye, 1988; Mather, Henkel, & Johnson, 1997; Schoolder, Gerhard, & Loftus, 1986). For example, in the study by Schoolder and colleagues (1986), young-adult participants first viewed a slide sequence showing an automobile accident. For half the participants, the sequence contained a slide (the Critical slide) that included a Yield sign. For the other half of the participants, the slide sequence was identical except that the Critical slide did not contain a Yield sign. Following the presentation of the slide sequence, participants were given a questionnaire. The questionnaire for the participants who had not seen the Yield sign contained a question that suggested that there was a Yield sign. Participants were then asked whether or not they had seen a Yield sign. If participants answered affirmatively to this question, they were asked to describe the sign. Descriptions of real memories contained more perceptual details, whereas descriptions of suggested memories contained more verbal hedges (e.g., “I think”), more references to cognitive operations presumably used in remembering the item, and more references to the functions of items. Similar findings have been obtained in studies of aging and the qualitative characteristics of real and implanted or imagined memories. Older adults report fewer perceptual details compared with younger adults, and differences in the phenomenal characteristics of veridical and suggested memories are smaller for older adults than for younger adults. Such findings have been interpreted as support for an age-associated source monitoring deficit (Chalfonte & Johnson, 1996; Hashtroudi, Johnson, & Chrosniak, 1989, 1990; Henkel, Johnson, & De Leonards, 1998; Norman & Schacter, 1997).

One aim of this study was to describe possible adult age differences in the extent to which memories are susceptible to suggestive influences, and to compare the magnitude of age-related differences in the suggestibility of memory with the magnitude of age-related differences in memory for observed information. We gave young and older adults either one or two exposures of a slide sequence that depicted a theft. Participants then received misleading information about the presence of particular items in the sequence. Although it was expected that older adults would show a higher percentage of false recognition of suggested items than young adults, it was also expected that magnitude of age differences in false recognition would be about the same as the magnitude of age differences in accuracy of remembering presented items.

A second aim was to describe possible age differences in the qualitative characteristics of veridical and suggested memories. Based on previous findings, it was expected that there would be differences in the descriptions of real and suggested items. Consistent with an interpretation that emphasizes age-related inefficiency in source monitoring, it was expected that there would be fewer qualitative differences in the descriptions of real and suggested items for older adults than for younger adults. That is, compared with the descriptions of younger adults, it was expected that older adults’ descriptions of suggested items would contain more perceptual and location references, and older adults’ descriptions of real items would contain more functional references. Such findings would also support an interpretation that emphasizes an age-related deficit in the strength of the encoded trace.

A third aim was to examine the effects of number of exposures on memory suggestibility and on the qualitative aspects of real and suggested memories. Although the effects of exposure on adult age differences in real and suggested memories have not been systematically examined in previous research, both young and older adults were expected to show greater memory accuracy and less susceptibility to suggestion if they received two exposures rather than one exposure to the slide sequence.

**METHODS**

**Participants**

A total of 60 young women and 62 older women participated. The young adults were undergraduates recruited from the Syracuse University human subjects pool. Mean age of the young adults was 19.0 years (range = 17–25 years). The education level of the young adults ranged from freshman year in college (13 years) to senior year in college (16 years). Older adults were recruited from several senior organizations. Mean age of the older participants was 73.0 years (range = 65–85 years). Mean education level of the older participants was 14.0 years ($SD = 3$).

Participants rated their health and vision on a scale of 1 to 4 (1 = Excellent; 2 = Good; 3 = Fair; 4 = Poor). For young adults, the mean health rating was 1.65 ($SD = .52$), and the mean vision rating was 1.95 ($SD = .72$). For older adults, the mean health rating was 1.97 ($SD = .63$), and the mean vision rating was 2.21 ($SD = .55$). Participants were also administered the Digit Span and Vocabulary subtests from the WAIS-R (Wechsler, 1981). Mean digit span scores were 16.4 and 16.3 for the young adults and older adults, respectively. Mean vocabulary scores were 51.8 and 53.3 for the younger and older adults, respectively. Statistical analyses of educational level, self-ratings of health and vision, and Digit Span and Vocabulary scores revealed no age differences. At the beginning of the experimental session, participants were also screened for visual acuity. Three older participants who could not copy the 20/30 line of screen...
projected Snellen eye chart were excused from further testing and were replaced.

Materials

Participants were shown a sequence of 62 slides depicting a theft. To achieve counterbalancing, two sequences of the theft episode were used. The sequences were nearly identical to those previously used by McCloskey and Zaragoza (1985). Each of the sequences contained two of four Critical items (a can of Coke, a Vogue magazine, a jar of coffee, or a wrench). Half of the participants saw a theft sequence that had a slide that included a can of Coke and a slide that included a Vogue magazine, and half saw a sequence that had a slide that included a wrench and a slide that included a jar of coffee.

Design and Procedure

Half of the participants in each age group were given one exposure to the slide sequence, and half were given two exposures. In the 2-Exposure condition, there was a delay of 5 min between exposures. All participants were instructed to try to remember the objects contained in the slides and the events depicted in the sequence. Participants were directed to pay close attention to the colors and locations of objects. Two warm-up slides were presented, each for 8 s, to give participants practice in describing and remembering objects shown in the slides.

Following presentation of either one or two exposures of the slides, participants received 15 min of filler activity. During this interval, participants completed the subject information questionnaire and Digit Span and Vocabulary subtests.

Next, participants were asked about their memory for objects and events in the sequence via Questionnaire 1, consisting of 24 questions that were similar in format to those used by Schooler and associates (1986). Questions were read aloud, and participants made their responses on a data sheet by circling yes or no, and by circling 1 to indicate certainty, 2 to indicate a moderate level of confidence, or 3 to indicate no confidence in response accuracy. Two of the questions contained misleading information, suggesting the presence of a Critical item. The same questions were asked—except that no mention was made of the Critical items—to participants who saw the sequence containing those Critical items. For example, participants who did not see the can of Coke were asked, “Did the thief pick up the bottle of Elmer’s glue that was on the second desk in front of the can of Coke?” Participants who had seen the can of Coke were asked, “Did the thief pick up the bottle of Elmer’s glue?”

Following 5 min of filler activity, participants were given a second questionnaire consisting of 15 items. Two versions of Questionnaire 2 were constructed so as to match the two versions of the slide sequence. Items were read aloud, and participants were asked if they had seen these items in the slide sequence. For each object, participants made yes or no responses to questions of the form, “Did you see . . . ?” Participants also gave confidence ratings for each response, using a 3-point scale. Two of the questions pertained to the Critical items.

Next, following 5 min of filler activity, a third set of questions was administered. Three questions were asked for each of the items that participants affirmed on Questionnaire 2, and each of the Critical items. These questions, derived from the Memory Characteristics Questionnaire (Johnson et al., 1988), assessed general vividness of the item, vividness for the color of the item, and vividness for the location of the item. Participants responded by circling a number on a 7-point scale (vague 1 2 3 4 5 6 7 clear/distinct).

Last, participants were asked to provide written detailed descriptions of the two Critical items and four other items selected arbitrarily from Questionnaire 2. Participants were instructed to take as much time as they needed to describe the objects.

Results

Responses to Questionnaire 1 provided measures of the accuracy of memory for items in the sequence and subjects’ confidence in their responses. Responses to Questionnaire 2 provided a measure of susceptibility to suggestion in terms of the accuracy of answers to questions about the suggested or Critical items. Accuracy of responses to questions about real items, and confidence ratings in the subjects’ responses to suggested and real items were also assessed on Questionnaire 2. Responses to Questionnaire 3 provided measures of the general vividness of memory for observed and suggested objects and of the vividness of the colors and locations of these objects. A summary of the significant correlations of age with measures from the three questionnaires is provided in Table 1.

Separate analyses of variance (ANOVAs) were conducted to examine possible differences between the orders in which questions were asked and between the two versions of the slide sequences for memory accuracy and confidence ratings on Questionnaire 2. Neither the order of the questions nor the version of the slide sequence produced any effects, and we collapsed over order and version in subsequent analyses. A preliminary analysis of the differences in accuracy between real items and Noncritical items (i.e., the two of the four Critical items that were not suggested but were shown to half the subjects depending on counter-

Table 1. Significant Correlations of Dependent Variables With Age

<table>
<thead>
<tr>
<th>Variable</th>
<th>Pearson Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Questionnaire 1:</td>
<td></td>
</tr>
<tr>
<td>Accuracy of memory</td>
<td></td>
</tr>
<tr>
<td>Questionnaire 2:</td>
<td></td>
</tr>
<tr>
<td>Confidence in false recognition of suggested items</td>
<td></td>
</tr>
<tr>
<td>Questionnaire 3:</td>
<td></td>
</tr>
<tr>
<td>Vividness Ratings</td>
<td></td>
</tr>
<tr>
<td>Colors of Noncritical items incorrectly denied</td>
<td></td>
</tr>
<tr>
<td>Locations of Noncritical items incorrectly denied</td>
<td></td>
</tr>
<tr>
<td>Memory of suggested items incorrectly affirmed</td>
<td></td>
</tr>
<tr>
<td>Colors of suggested items incorrectly affirmed</td>
<td></td>
</tr>
<tr>
<td>Item Description</td>
<td></td>
</tr>
<tr>
<td>Location references</td>
<td></td>
</tr>
</tbody>
</table>

*p < .01; ***p < .0001.
balancing) revealed no effects, so we collapsed over responses to real and Noncritical items so as to simplify the presentation of the results. All analyses reported below were two-tailed, and a Bonferroni correction for multiple tests was used.

**Memory for Real Items**

The effects of age and number of exposures on general memory performance were examined by looking at the accuracy of responses to real items on Questionnaires 1 and 2. On Questionnaire 1, the level of accuracy of memory for real items was higher for young adults \((M = .90, SD = .06)\) than for older adults \((M = .81, SD = .12)\), \(F(1,117) = 29.72, p < .0001, MSe = .09\). A Pearson product moment correlation revealed that, as age increased, the proportion of accurate responses to real items decreased \((r = -.46, p < .0001)\). For both age groups, the benefit from two exposures to the sequence was not significant; the mean for the 2-Exposure condition was \(.87, (SD = .10)\) and the mean for the 1-Exposure condition was \(.84 (SD = .10)\).

An analysis of effects of Age and Exposure on accuracy for real items on Questionnaire 2 revealed main effects for Age, \(F(1,118) = 13.63, p < .0003, MSe = .01\), and Exposure, \(F(1,118) = 8.79, p < .004\), and an Age \(\times\) Exposure interaction, \(F(1,118) = 9.14, p < .003\). When given only one exposure to the slide sequence, there was no difference between age groups in accuracy for real items. However, young adults performed more accurately in the 2-Exposure condition \((M = .91, SD = .09)\) than they did in the 1-Exposure condition \((M = .78, SD = .13)\). In addition, young adults in the 2-Exposure condition performed more accurately than older adults in the 2-Exposure condition \((M = .77, SD = .13)\) and more accurately than older adults in the 1-Exposure condition \((M = .77, SD = .10)\). The accuracy of older adults was not different between Exposure conditions. No effects of Age or Exposure were obtained for the confidence ratings for real items.

**Memory for Critical Items**

Two of the questions on Questionnaire 2 were about items that were not shown, but their presence in the sequence was suggested in the wording of Questionnaire 1. These are the Critical items (either a can of Coke and a Vogue magazine or a wrench and a jar of coffee). Possible effects of age and number of exposures on the suggestibility of memory were examined by looking at the accuracy of responses to the Critical items.

In both exposure conditions, older adults were more susceptible to suggestion than young adults, as indicated by a main effect for Age, \(F(1,118) = 4.61, p < .03, MSe = .13\). Young adults responded more accurately to these questions, correctly denying the presence of the suggested items at a greater frequency \((M = .80, SD = .32)\) than older adults \((M = .66, SD = .38)\).

Furthermore, older adults reported a higher level of confidence in their affirmative responses to the Critical items than young adults, \(F(1,47) = 10.40, p < .002, MSe = .21\). Mean confidence ratings for affirmative answers to the suggested items were \(1.71 (SD = .45)\) and \(1.27 (SD = .49)\) for young adults and older adults, respectively. Note that a score of 1 indicated the highest level of confidence in the answer, and a score of 3 indicated that the answer was a guess.

The correlation between age and the confidence scores for affirmative answers to suggested items was reliable \((r = -.43, p < .002)\). Analysis of the correlations between accuracy and confidence ratings for correct rejections of Critical items \((r = .10, p < .33)\) and between accuracy and confidence ratings for affirmative responses to Critical items \((r = .01, p < .93)\) suggested that there was no relationship between confidence and accuracy of memory. Levels of confidence in affirmative responses to Critical items were higher for participants in the 1-Exposure condition \((M = 1.27, SD = .51)\) than for participants in the 2-Exposure condition \((M = 1.60, SD = .48)\), but this difference was not reliable at the .01 criterion, \(F(1,118) = 4.19, p < .04, MSe = .12\).

**Comparisons Among Memory Measures**

Note that the questions on Questionnaire 1 were longer than one clause and they included contextual information about the slide sequence. In contrast, the questions on Questionnaire 2 were in the form “Did you see...?” and did not provide a context for remembering objects and related details. The magnitude of age-related differences in memory revealed by these different measures was assessed by calculating the effect sizes for each of the three comparisons using Cohen’s \(d\) (Cohen, 1988). These values are reported in Table 2. The measure of memory for real items on Questionnaire 1 produced an age difference of almost one standard deviation \((d = .94, 95\%\text{ confidence interval} = [.87–1.00])\). In comparison, age differences were smaller for the measures of real and suggested memory taken from Questionnaire 2. Specifically, \(d = .67 (95\%\text{ CI} = .60–.73)\) for real items, and \(d = .39 (95\%\text{ CI} = .33–.46)\) for suggested items. These differences in effect sizes were taken to be reliable and robust because the 95% CIs for these estimates did not overlap (despite possible differences in the reliabilities of the two questionnaires). Note that Questionnaire 1 contained 24 items and provided relatively rich contextual support for item retrieval, whereas Questionnaire 2 contained 15 items and provided no support for item retrieval. Considering these format differences, it appears that the availability of context to support memory performance did not serve to attenuate the magnitude of age differences in memory. This finding contrasts with an environmental support hypothesis of memory aging (Craik & Jennings, 1993).

<table>
<thead>
<tr>
<th>Memory Type</th>
<th>Age</th>
<th>Mean</th>
<th>SD</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Questionnaire 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Real items</td>
<td>Young</td>
<td>.90</td>
<td>.06</td>
<td>.94</td>
</tr>
<tr>
<td></td>
<td>Old</td>
<td>.81</td>
<td>.12</td>
<td></td>
</tr>
<tr>
<td>Questionnaire 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Real items</td>
<td>Young</td>
<td>.84</td>
<td>.11</td>
<td>.67</td>
</tr>
<tr>
<td></td>
<td>Old</td>
<td>.77</td>
<td>.12</td>
<td></td>
</tr>
<tr>
<td>Suggested items</td>
<td>Young</td>
<td>.80</td>
<td>.32</td>
<td>.39</td>
</tr>
<tr>
<td></td>
<td>Old</td>
<td>.66</td>
<td>.38</td>
<td></td>
</tr>
</tbody>
</table>
REAL AND SUGGESTED MEMORIES

Vividness Ratings

Vividness ratings for correct and incorrect responses to real and suggested items are summarized in Table 3. The ratings are presented for young and older adults collapsed over the two Exposure conditions. Relatively few effects involving the Exposure factor were obtained. A consistent pattern of interactions emerged between Age and the differences between items. Separate mixed factorial ANOVAs confirmed the interactive effects of Age with the differences between item types for general memory (p < .01), colors (p < .001), and locations (p < .01).

Analyses of the vividness ratings also revealed a main effect of Exposure and an Age by Exposure interaction for falsely recognized Critical items for each of the vividness dimensions (general memory, color, and location). In each case, the second exposure to the slide sequence served to attenuate the reported vividness of false recognitions for young adults but not for older adults (all ps < .01). For each dimension, the Exposure manipulation did not influence the ratings of older adults.

The pattern of the data shown in Table 3 suggests that older adults rated suggested and observed items more similarly than did younger adults, suggesting more source confusion for older adults than for young adults.

Table 3. Mean Vividness Ratings for Real Items and Suggested Items by Age

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Item Type</th>
<th>Age</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>General Memory*</td>
<td>Real correct*</td>
<td>Young</td>
<td>6.54</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Old</td>
<td>6.31</td>
<td>1.15</td>
</tr>
<tr>
<td></td>
<td>Real incorrect***</td>
<td>Young</td>
<td>1.66</td>
<td>1.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Old</td>
<td>3.28</td>
<td>2.22</td>
</tr>
<tr>
<td></td>
<td>Suggested correct***</td>
<td>Young</td>
<td>1.92</td>
<td>1.08</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Old</td>
<td>2.04</td>
<td>1.30</td>
</tr>
<tr>
<td></td>
<td>Suggested incorrect***</td>
<td>Young</td>
<td>4.18</td>
<td>2.04</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Old</td>
<td>5.65</td>
<td>1.47</td>
</tr>
<tr>
<td>Color***</td>
<td>Real correct*</td>
<td>Young</td>
<td>6.01</td>
<td>1.12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Old</td>
<td>6.11</td>
<td>1.22</td>
</tr>
<tr>
<td></td>
<td>Real incorrect***</td>
<td>Young</td>
<td>1.73</td>
<td>1.29</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Old</td>
<td>3.07</td>
<td>2.25</td>
</tr>
<tr>
<td></td>
<td>Suggested correct***</td>
<td>Young</td>
<td>2.00</td>
<td>1.37</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Old</td>
<td>2.19</td>
<td>1.33</td>
</tr>
<tr>
<td></td>
<td>Suggested incorrect***</td>
<td>Young</td>
<td>3.42</td>
<td>1.95</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Old</td>
<td>5.18</td>
<td>1.89</td>
</tr>
<tr>
<td>Location*</td>
<td>Real correct</td>
<td>Young</td>
<td>6.24</td>
<td>1.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Old</td>
<td>6.09</td>
<td>1.17</td>
</tr>
<tr>
<td></td>
<td>Real incorrect***</td>
<td>Young</td>
<td>1.58</td>
<td>1.08</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Old</td>
<td>3.29</td>
<td>2.28</td>
</tr>
<tr>
<td></td>
<td>Suggested correct***</td>
<td>Young</td>
<td>0.95</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Old</td>
<td>1.33</td>
<td>0.90</td>
</tr>
<tr>
<td></td>
<td>Suggested incorrect***</td>
<td>Young</td>
<td>4.11</td>
<td>1.96</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Old</td>
<td>5.47</td>
<td>1.55</td>
</tr>
</tbody>
</table>

*p < .01; ***p < .0001.

Item Descriptions

Participants’ descriptions of real and suggested items were scored in terms of the following qualities: (a) perceptual attributes (color, size, shape); (b) location references, including their spatial relationships to other objects or any scenery in the slides; (c) function or purpose of the items; and (d) use of verbal hedges, such as “I think” and “I’m not sure but.” Inspection of the verbal hedges data suggested that many of the individual scores were idiosyncratic or meaningless, and these were not analyzed further. A series of 2 x 2 ANOVAs were conducted to examine the effects of Age and Exposure on the numbers of perceptual details, location references, and functional references used by participants to describe real and suggested items. The data are presented in Table 4.

Separate analyses were performed on suggested and real items that were correctly rejected and falsely recognized on Questionnaire 2. A main effect of Age was found for the number of location references used in the descriptions of real items, F(1,118) = 25.61, p < .0001, MSe = 1.43. Young adults used more location references (M = 2.30, SD = 1.03) than older adults (M = 1.16, SD = 1.03). The correlation between age and the number of location references used in describing real items was significant, r = -.49, p < .0001.

Also, as shown in Table 4, more perceptual and location details were used to describe objects that were actually presented than were used to describe suggested objects. Pairwise comparisons for correct and incorrect responses indicated that more perceptual and location details were given for correct than for incorrect responses. Contrary to expectation, no age differences in the numbers of functional references used to describe suggested and real objects were found.

DISCUSSION

The present study examined the effects of age and exposure on (a) memory for observed and suggested information, (b) confidence ratings in responses to real and suggested information, and (c) possible differences in the phenomenal characteristics of real and suggested memories.

Overall, the results were consistent with previous findings, suggesting greater susceptibility to suggestion for

Table 4. Mean Numbers of Attributes in the Descriptions of Real and Suggested Objects for Correct and Incorrect Responses: Results of Pair-Wise Comparisons

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Real Correct</th>
<th>Real Incorrect</th>
<th>Suggested Correct</th>
<th>Suggested Incorrect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceptual details***</td>
<td>1.67 (1.4)</td>
<td>0.23 (.6)</td>
<td>0.66 (.8)</td>
<td>0.14 (.4)</td>
</tr>
<tr>
<td>Location references***</td>
<td>1.87 (1.3)</td>
<td>0.07 (.3)</td>
<td>0.28 (.6)</td>
<td>1.11 (.9)</td>
</tr>
<tr>
<td>Functional references</td>
<td>0.05 (.2)</td>
<td>0.04 (.1)</td>
<td>0.03 (.1)</td>
<td>0.20 (.4)</td>
</tr>
</tbody>
</table>

Note: Standard deviations in parentheses.

***p < .0001.
older adults than for young adults (Cohen & Faulkner, 1989; List, 1986; Loftus, Leavdow, & Dunsing, 1992; Muchtaup et al., 1999; Schacter et al., 1997). Older adults affirmed having seen the suggested items with greater frequency as well as endorsing higher levels of confidence in their affirmative responses to suggested items than young adults. Interestingly, the magnitudes of age-related differences in the suggestibility of memory were actually smaller than size of the effect of age on veridical memory. Another important finding was that the vividness ratings for the suggested items were greater for older adults than for young adults.

It is useful to look at the results of this study in terms of their implications for specifying the processes and mechanisms that underlie the aging of memory and memory suggestibility. Candidate explanations for the observed age effects include source monitoring, encoding strength, feature binding, and discrepancy detection. Although the design of the present study does not permit us to precisely distinguish among the various alternative accounts, explanations that emphasize relatively weak encoded traces and ineffective source monitoring seem most plausible in our view. Compared to younger adults, the memories of older adults may be more vulnerable to interference from competing subsequent events and destructive updating because of a relatively weaker initial trace (Ceci et al., 1988; Coxon & Valentine, 1997; Loftus et al., 1978; Tun et al., 1998). The notion of a weaker trace is supported by the pattern of age differences in the accuracy of memory for real objects and age differences in false recognition of suggested objects.

The idea that age-related differences in the strength of the original trace can account for suggestibility effects is consistent with the data showing age differences in confidence ratings and in the phenomenal characteristics of veridical and suggested memories. Consistent with previous findings (Johnson et al., 1988; Norman & Schacter, 1997; Schooler et al., 1986; Suengas & Johnson, 1988), we found qualitative differences in the vividness ratings of real and suggested memories. Descriptions of real memories contained more perceptual details and more location information for both age groups, and the differences in the vividness ratings of real and suggested memories were smaller for older adults than for younger adults. Age differences in the vividness of veridical and suggested memories suggest corresponding differences in the strength or specificity of encoding of the initial trace. Of course, the findings reported here are also consistent with an account that emphasizes inefficient source monitoring on the part of older adults. It is not possible for us to distinguish among these accounts, either in terms of the data from our study or the data from other studies (e.g., Hashtroudi et al., 1990; Norman & Schacter, 1997).

A related explanation, called discrepancy detection, suggests that individuals with poor memories to begin with will be more susceptible to suggestion than individuals with good memories (Loftus et al., 1992). That is, people with good memories are good learners and ordinarily pay attention to details when learning. Similarly, Brainerd and Reyna (1988) pointed out that age-related differences in suggestibility can be viewed as a manifestation of process—general chronological age changes in memory and forgetting. Differences among individuals with good and poor memories can be related to strategy differences that make for stronger traces, or they can be linked to underlying brain physiology such as prefrontal function (Norman & Schacter, 1997). Another account that could be applied to the present findings points to age differences in feature memory and binding (Chalfonte & Johnson, 1996). Because most or all of the various candidate accounts are indistinguishable at an empirical level, an explanation emphasizing nonspecific effects of memory aging on memory suggestibility seems most parsimonious.

The finding that young adults benefited more from a second exposure than older adults supports a trace strength interpretation of the data if it is assumed that two exposures were insufficient to build a strong trace for older adults. These findings are in line with the results from a recent study by Koutstaal and colleagues (1999), which found that encoding and retrieval manipulations designed to reduce gist-based false recognition in older adults were insufficient for the purposes of eliminating adult age differences in false recognition. Dunlosky and Salthouse (1996) also reported that multitrial free recall was less efficient for older adults than for younger adults, and interpreted their findings in terms weaker representations of the to-be-remembered items. Future studies ought to test the trace strength account of aging and memory suggestibility by bringing young adults and older adults to the same criterion of episodic learning prior to the introduction of misleading information. If distortion is associated with incomplete traces, older adults should perform as accurately as younger adults in memory situations when trace or source of event is completely encoded. When events are incompletely encoded or when it is easier to rely on a schematic rather than true representation of specific episodes, older adults will show distortion commensurate with general memory aging effects. In a study examining age differences in drawing inferences, for example, Reder, Wible, and Martin (1986) found that exact retrieval of textual information was more difficult for older adults than for younger adults, but that the use of inferences and reasoning based on prior knowledge was easy for both groups.

Some aspects of our results have practical implications. Specifically, subjects' confidence in their memories was not related to memory accuracy. This finding is consistent with the conclusion reached by Adams-Price and Perlmuter (1992), in their review of the literature, that confidence appears to be related to factors other than accuracy, such as repetition. The implication is that the accuracy of an eyewitness testimony cannot be accurately assessed by the amount of confidence the witness appears to have in the testimony. Further, the accuracy of memories of older adults cannot be evaluated through the assessment of the vividness of their memories because the vividness of real and suggested memories is not very different.

Possible age-related differences in sensitivity to the interpersonal demands of the testing situation might have some bearing on the obtained findings. It may be that older adults tend to be more compliant with demand characteristics of this type of investigation (McCloskey & Zaragoza, 1985;
REAL AND SUGGESTED MEMORIES

P109

Smith & Ellsworth, 1987; Toglia, Ross, Ceci, & Hembrooke, 1992). That is, older adults may be more likely than young adults to behave in a way that they believe is expected of them, such as agreeing with the experimenter about the existence of an item even if they do not remember it. Or, older adults might endorse a higher rating of vividness for items they do not remember because they are self-conscious about memory loss and do not want to appear as not vividly remembering the objects asked about. However, our findings of age differences in memory accuracy and in item descriptions suggest that noncognitive factors probably did not play as much of a role in the suggestibility effects obtained.

Studies of age-related differences in memory and other aspects of cognitive function sometimes yield findings that clarify as well as extend the understanding of general cognitive phenomena. The results of the present study have implications for developing precise and more parsimonious descriptions of the effects of suggestion on memory distortion. Many explanations have been suggested to account for the susceptibility of memory to misleading information (e.g., Belli, 1989; Brainerd & Reyna, 1988; Ceci et al., 1987; Loftus, 1995; McCloskey & Zaragoza, 1985; Pezdek & Banks, 1996; Toglia et al., 1992; Tversky & Tuchin, 1989; Zaragoza & McCloskey, 1989). However, the data reported here imply that memory suggestibility is nonspecific and largely a function of the strength of encoded trace and the cognitive operations that ordinarily underlie the formation and retrieval of episodic memories.

Acknowledgments

This research was supported by Grant AG11451 from the National Institute on Aging. We gratefully acknowledge Maria Zaragoza and Elizabeth Loftus for sharing their stimulus materials. We thank Michael Levy for assistance with data collection and Blair Johnson and Paul Verhaeghen for valuable comments on earlier drafts of the manuscript.

Address correspondence to William J. Hoyer, Department of Psychology, Syracuse University, Syracuse, NY 13244-2344. E-mail: wjhoyer@syr.edu

References


