

# Misattribution Errors in Alzheimer's Disease: The Illusory Truth Effect

Jason P. Mitchell  
Harvard University

Daniel L. Schacter  
Harvard University

Alison L. Sullivan  
Brigham and Women's Hospital

Andrew E. Budson  
Brigham and Women's Hospital and Edith Nourse Rogers  
Memorial Veterans Hospital

Patients with mild Alzheimer's disease (AD) and age-matched controls were compared on a series of tasks designed to measure errors of misattribution, the act of attributing a memory or idea to an incorrect source. Misattribution was indexed through the illusory truth effect, the tendency for participants to judge previously encountered information to be true. Cognitive theories have suggested that the illusory truth effect reflects the misattribution of experimentally produced familiarity (a nonspecific sense that an item has been previously encountered) to the veracity of previously encountered information. Consistent with earlier suggestions that AD impairs both familiarity and recollection (specific memory for contextual details of the study episode), AD patients demonstrated significantly fewer misattribution errors under conditions in which the illusory truth effect is thought to rely on relative familiarity (uncued condition), but more misattribution errors under conditions thought to rely on relative amounts of contextual recollection (cued condition). These results help further specify the precise nature of memory impairments in AD.

*Keywords:* memory errors, Alzheimer's disease, misattribution, recollection, familiarity

Although impairments in episodic memory in Alzheimer's disease (AD) have been well-characterized cognitively (see Hodges, 2000, for an overview), less is known regarding the susceptibility of AD patients to various forms of memory distortions. One type of memory distortion that has begun to be studied in AD is that of *misattribution*, the act of attributing a recollection or idea to an incorrect source (Budson, Daffner, Desikan, & Schacter, 2000, 2001; Budson et al., 2002; Schacter, 1999). For example, to examine misattribution errors that result when one falsely remembers an event that never occurred, Budson et al. (2000) tested AD patients using the Deese–Roediger–McDermott (DRM) paradigm (Deese, 1959; Roediger & McDermott, 1995). In the DRM paradigm, participants study lists of semantic associates that all converge on an unrepresented “theme” word, and, at remarkably high

rates, normal participants subsequently claim that this theme word had been presented earlier. Intriguingly, Budson et al. (2000) observed that, whereas AD patients were significantly less likely than controls to claim erroneously that the theme word had been presented after a single presentation of the study lists, this pattern reversed when study lists were repeated multiple times—after seeing the lists five times, AD patients demonstrated greater rates of false memory than controls.

These results have been interpreted as reflecting parallel impairments in AD of two components of recognition memory: *familiarity* (a nonspecific sense that an item has been encountered earlier, devoid of additional contextual details about the initial encounter) and *recollection* (memory for the specific aspects of a previous encounter with the item). By converging strongly on a single theme word, the DRM lists are thought to spuriously augment the familiarity of this unrepresented word. Budson et al. (2000) suggested that AD patients may accumulate familiarity relatively slowly, thereby initially demonstrating relatively low levels of false recognition compared with controls (see also Schacter, Verfaellie, & Anes, 1997; Schacter, Verfaellie, Anes, & Racine, 1998; Schacter, Verfaellie, & Pradere, 1996; Verfaellie, Schacter, & Cook, 2002). However, over several presentations of the lists, AD patients—like normal controls—may begin to accumulate familiarity for the unrepresented items. Left unchecked, this increased familiarity will tend to produce greater false memory for unrepresented items. Over several presentations of the lists, however, healthy controls will tend to develop increasing contextual recollection for the stimulus items and may be able to use such recollection to suppress false memory for the unrepresented item. In contrast, AD patients may fail to accumulate contextual recollection over repeated presentation of the lists and, accordingly, be unable to counter the increasing effect of familiarity on false memory. In

---

Jason P. Mitchell and Daniel L. Schacter, Department of Psychology, Harvard University; Alison L. Sullivan, Division of Cognitive and Behavioral Neurology, Department of Neurology, Brigham and Women's Hospital, Boston, MA; and Andrew E. Budson, Division of Cognitive and Behavioral Neurology, Department of Neurology, Brigham and Women's Hospital, and Geriatric Research Education Clinical Center, Edith Nourse Rogers Memorial Veterans Hospital, Bedford, MA.

This research was supported by a National Research Service Award predoctoral fellowship to Jason P. Mitchell, by National Institute of Aging Grant 08441 to Daniel L. Schacter, and by National Institute of Mental Health Grant K23 MH01870 and a Brigham and Women's Hospital Faculty Award in Translational Neurosciences to Andrew E. Budson.

Correspondence concerning this article should be addressed to Jason P. Mitchell, Department of Psychology, Harvard University, William James Hall, 33 Kirkland Street, Cambridge, MA 02138, or Andrew E. Budson, Geriatric Research Education Clinical Center, Building 62, Room B30, Edith Nourse Rogers Memorial Veterans Hospital, 200 Springs Road, Bedford, MA 01730. Electronic mail may be sent to [jmitchel@wjh.harvard.edu](mailto:jmitchel@wjh.harvard.edu) or [abudson@partners.org](mailto:abudson@partners.org)

other words, although AD patients may initially demonstrate less false recognition because they acquire familiarity relatively slowly, they may nevertheless eventually come to demonstrate greater false recognition as the accumulation of familiarity outstrips the recollection needed to suppress increased familiarity-based false recognition over repeated study list iterations.

In addition to studies that use the DRM paradigm, misattribution has been investigated most extensively in normal populations by examining perceivers' tendency to alter their subjective judgments of stimuli on the basis of incomplete memory for one's initial encounter with them. Perhaps the most robust example of such judgment-based misattribution is indexed by the *illusory truth effect*, the tendency for perceivers to increase their judgments of the truth of previously encountered information (Bacon, 1979; Begg, Anas, & Farinacci, 1992; Gilbert, Krull, & Malone, 1990). For example, after studying a series of trivia statements (e.g., "It takes 4 hours to hard boil an ostrich egg"), perceivers more frequently judge items to be true (i.e., an accurate reflection of fact) relative to items that are novel to the experiment (Bacon, 1979; Gilbert, 1991, 1993; Gilbert et al., 1990). Moreover, perceivers continue to demonstrate the illusory truth effect even after being informed at study about which statements are true and which are false (Begg et al., 1992; Gilbert et al., 1990); that is, statements that are initially cued as false are nevertheless judged to be true more frequently than novel statements.

Similar to accounts of misattribution on the DRM paradigm, cognitive theories of judgment-based misattribution errors have generally focused on the competing influences of familiarity and recollection (Begg et al., 1992; Jacoby, 1991; Jacoby, Kelley, Brown, & Jasechko, 1989; Johnson, Hashtroudi, & Lindsay, 1993; Schacter, 1999). According to such accounts, prior exposure to an item will increase an item's familiarity, and perceivers may systematically misattribute this relatively high familiarity to a variety of extraneous sources, such as a statement's truth, instead of an earlier experimental encounter (the actual source of increased familiarity). In such cases, recollection of the contextual details surrounding this earlier encounter may be needed to avoid misattribution errors, as perceivers need not misattribute familiarity to an item's truth if they can recollect the item's earlier presentation. The notion that recollection underlies the ability to avoid judgment-based misattribution is supported by a variety of empirical evidence, including manipulations that impair recollection, such as introducing a secondary task during encoding or confusing the distinction between sources (Johnson et al., 1993), analytic techniques designed to separate recollection and familiarity (Begg et al., 1992; Jacoby, 1991; Jacoby et al., 1989), and recent neuroimaging evidence (Mitchell, Dodson, & Schacter, 2005).

In the present experiment, we examined the illusory truth effect in AD patients and healthy elderly controls. As described above, the illusory truth effect has two different behavioral manifestations, each of which may be differently affected by AD. First, in the absence of any information at study about the actual truth or falseness of information, perceivers later tend to judge previously encountered information to be true relative to novel information. It is interesting that this effect occurs even under conditions in which recognition memory for the information is near ceiling (Begg et al., 1992), suggesting that, when information is uncued as true or false, recollection may not be sufficient to suppress familiarity-

based misattribution. Accordingly, to the extent that AD is marked by deficits in familiarity (Budson et al., 2000, 2001), we expect that AD patients should demonstrate fewer misattribution errors than healthy controls when information is not cued during the study phase. To test this prediction, we asked participants to study a set of uncued trivia statements. Subsequently, we presented these items among novel statements and asked participants to judge each as either true or false (uncued truth task). Relative to control participants, AD patients were expected to judge fewer previously encountered items to be true compared with novel items on the uncued truth task.

In contrast, a second manifestation of the illusory truth effect occurs when participants are told during study about which information is to be believed and nevertheless tend to misjudge cued-false information to be true (often more frequently than novel information; Begg et al., 1992). Statistical techniques (Begg et al., 1992) and recent neuroimaging research (Mitchell et al., 2005) have demonstrated that this effect relies critically on the absence of recollection for the study episode; participants tend not to misjudge cued-false information to be true when they have access to contextual memory for the cue with which information was initially presented. In contrast to the first prediction, AD patients should produce more misattribution errors than healthy controls when information is cued as being true or false during the study phase, as the absence of recollection augments the number of illusory truth errors under these conditions. To test this second prediction, we asked participants to study an additional set of trivia statements that were each cued as either true or false. Subsequently, we presented these items a second time and asked participants to judge each as either true or false (cued truth task). Relative to control participants, AD patients were expected to misjudge more cued-false items as true but to correctly judge an equivalent number of cued-true items.

Finally, we expected that performance on the cued truth task would be related to overall recollection deficits among AD patients. Specifically, the correct judgment of cued-false items is thought to require access to recollection; accordingly, the preserved ability to recollect contextual details of the study episode should enable an individual to avoid incorrectly judging cued-false items as true. To test this prediction, we also measured participants' performance on a standard source memory task (Johnson et al., 1993). Participants studied a set of statements that were each read in either a female or male voice. Subsequently, we presented these items a second time and asked participants to indicate the voice in which each was originally presented (gender source task). Consistent with previous demonstrations of source memory deficits in AD (e.g., Dalla Barba, Nedjam, & Dubois, 1999), AD patients were expected to demonstrate significantly impaired performance on the gender source task relative to control participants. Critically, however, we predicted that overall performance on the gender source task would correlate with the ability to correctly judge cued-false items to be false, but not with the ability to judge cued-true items to be true (as the former, but not the latter, items are thought to require access to recollection). Such a finding would provide confirmatory evidence that recollection plays a critical role in suppressing misattribution errors, such as those captured by the illusory truth effect.

## Method

### Participants

Participants were 16 patients with a clinical diagnosis of probable AD (9 women, 7 men), as indexed by the criteria of the National Institute of Neurological and Communicative Disorders and Stroke–Alzheimer's Disease and Related Disorders Association (McKhann et al., 1984) and 16 healthy older adults (8 women, 8 men). AD patients were matched to control participants on the basis of age ( $M = 77.4$  and  $77.6$  years, respectively) and education ( $M = 16.4$  and  $15.8$  years, respectively). Patients were recruited from the clinical population at the Memory Disorders Unit, Brigham and Women's Hospital in Boston, MA. Older adults were recruited from participants in a longitudinal study of normal aging at Brigham and Women's Hospital, from spouses and friends (but not blood relatives) of the AD patients, and from flyers and posters placed in senior centers in and around Boston. Written informed consent was obtained from all participants and their caregivers (when appropriate). The study was approved by the Human Subjects Committee of Brigham and Women's Hospital. As a group, AD patients showed mild impairment on the Mini-Mental Status Examination (Folstein, Folstein, & McHugh, 1975; AD,  $M = 25.4$ ; healthy controls,  $M = 29.3$ ). All participants had normal or corrected-to-normal vision and hearing and were excluded if they were characterized by clinically significant depression, alcohol or drug use, brain damage due to stroke or tumor, traumatic head injury, or if English was not their primary language.

### Materials

Stimuli consisted of 88 trivia statements taken from previous work on the illusory truth effect (Bacon, 1979). Each statement conveyed information with which participants were unlikely to be familiar, for example, "It takes 4 hours to hard boil an ostrich egg." Half of the statements were factually true, whereas the remaining half were factually false. For each statement, we created a digital audio recording of an actor reading the statement aloud. Each statement was recorded three times, once in a female voice and twice in different male voices; the length of each statement's recording was roughly equal across the different actors. The stimulus set was randomly divided into four equal lists that were rotated through conditions across participants.

### Procedure

Participants completed three separate tasks: uncued truth, cued truth, and gender source. Different lists of stimuli were used across the three tasks (two lists were used in uncued truth, and one each in cued truth and gender source). All participants first performed the uncued truth task and then performed the cued truth and gender source tasks (the order of cued truth and gender source was counterbalanced across participants).

The uncued truth task indexed each participant's tendency to commit misattribution errors (i.e., judge statements to be true) in the absence of information about the veracity of the statements. During the study phase of the uncued truth task, 22 statements were presented sequentially. Stimuli were simultaneously presented visually on a computer screen and read aloud through stereo headphones by one of the male voices. Each statement appeared for 7 s, and participants were instructed simply to attend to the stimuli for a later memory test. Following a 5-min delay, each of the 22 old statements was randomly presented among 22 novel statements. At test, stimuli were only presented visually. Participants were asked to consider each statement and indicate whether they believed it to be true or false. Participants indicated their judgments verbally, and the experimenter pressed one of two computer keys to record their response. Statements remained onscreen until a response was recorded.

The cued truth task indexed participants' tendency to commit misattribution errors for information that was explicitly presented as false.

The study phase of the cued truth task was identical to that of the uncued truth task with the exception that 44 statements were presented and each statement was preceded by one of two cues (*true* or *false*). Cues were both read aloud and presented visually above the statement (where they remained onscreen throughout the presentation of the statement). Cue onset preceded the statement by 1 s. During this phase, participants were instructed to memorize each of the 44 statements along with its associated cue. During the test phase of the cued truth task, each of these statements was presented visually (not accompanied by the cue word), and participants verbally indicated whether they believed each to be true or false. No novel statements were presented during the test phase of the cued truth task.

The gender source task indexed participants' source memory by assessing participants' ability to recollect whether statements had originally been presented by a female or male voice. As for the cued truth task, the study phase of the gender source task consisted of 44 statements presented audibly and visually. However, in an attempt to equate overall performance on the gender source and cued truth tasks, statements were each presented twice; specifically, after each statement was presented once, the entire list was repeated in a new random order. Half of the statements were read in a female voice, whereas half were read in a male voice. To increase the salience of the gender manipulation and to parallel the cued truth task, the word *female* or *male* (consistent with the gender of the voice) was read immediately prior to the statement and remained onscreen throughout its presentation. Participants were instructed to memorize each statement along with the gender of the voice in which it was read. During the test phase of this task, each of the 44 statements was presented visually (not accompanied by the cue word), and participants verbally indicated whether the statement had been read in a female or male voice. No novel statements were presented during the test phase of the gender source task.

## Results

In the uncued truth task, participants studied a series of trivia statements without reference to their truth value and later judged each of these old statements and an equal number of novel statements as either true or false. To index the tendency to judge previously encountered statements to be true, we calculated the proportion of old and novel items that each participant judged to be true (see Table 1). A 2 (item type: old, novel)  $\times$  2 (group: AD, controls) mixed-model analysis of variance (ANOVA) revealed a significant main effect of item type, such that old items were significantly more likely to be judged true ( $M = .74$ ) than novel items ( $M = .48$ ),  $F(1, 30) = 29.58$ ,  $p = 10^{-5}$ , Cohen's  $d = 0.99$ . This effect was qualified by a marginally significant two-way interaction,  $F(1, 30) = 3.95$ ,  $p < .06$ ,  $d = 0.73$ , suggesting that when information was uncued, AD patients were somewhat less likely to increase truth judgments of old relative to novel statements

Table 1  
Proportion of Items Judged "True" in the Uncued Truth Task

Item type	Participant group		Difference
	AD patients	Controls	
Old	.70	.77	.07
New	.54	.43	-.11

Note. AD = Alzheimer's disease.

( $M = .16$ ) than control participants ( $M = .34$ ).<sup>1</sup> No main effect of group was observed,  $F(1, 30) = 0.11$ ,  $ns$ ,  $d = 0.12$ .

In the cued truth phase of the experiment, statements were cued as either *true* or *false*, and participants subsequently judged the truth value of each statement. To index performance, we calculated the proportion of cued-false and cued-true statements for which each participant correctly judged the truth value (see Table 2). A 2 (item type: cued-false, cued-true)  $\times$  2 (group: AD, controls) mixed-model ANOVA revealed significant main effects of both item type and group. As a whole, participants were significantly more likely to correctly judge a cued-true item ( $M = .73$ ) than a cued-false item ( $M = .51$ ),  $F(1, 30) = 20.80$ ,  $p = 10^{-5}$ ,  $d = 0.83$ ; moreover, AD patients were significantly worse at correctly judging the truth value of items ( $M = .55$ ) than control participants ( $M = .69$ ),  $F(1, 30) = 6.70$ ,  $p < .02$ ,  $d = 0.95$ . Although no significant two-way interaction was observed,  $F(1, 30) = 1.63$ ,  $p > .20$ ,  $d = 0.47$ , planned comparisons revealed the expected performance asymmetry between cued-true and cued-false items. Specifically, whereas both groups of participants demonstrated equivalent performance on cued-true items,  $t(30) = 1.22$ ,  $p > .20$ ,  $d = 0.45$ , AD patients were significantly less likely than control participants to correctly judge the truth value of cued-false items,  $t(30) = 2.83$ ,  $p < .01$ ,  $d = 1.03$ .

In the gender source phase of the experiment, statements were presented in either a female or a male voice, and participants subsequently indicated which of these two sources had originally presented each statement. To index source memory, we calculated the proportion of cued-female and cued-male statements that each participant correctly associated with its original source (see Table 3). A 2 (item type: cued-female, cued-male)  $\times$  2 (group: AD, controls) mixed-model ANOVA revealed a significant main effect of item type, such that participants were significantly more likely to indicate that a cued-male item was originally spoken by a male voice ( $M = 0.71$ ) than to indicate that a cued-female was originally spoken by a female voice ( $M = 0.61$ ),  $F(1, 30) = 15.03$ ,  $p < .001$ ,  $d = 0.71$ . This effect indicated that participants had a general (and unexpected) bias toward judging the source to be male, perhaps because throughout the experiment, two lists were read by male voices but only one by a female voice. The distributions of overall performance on the gender source were almost nonoverlapping between AD patients (range = 0.36–0.64) and control participants (range = 0.61–0.98), resulting in a highly reliable main effect of group,  $F(1, 30) = 10^{13}$ ,  $p < 10^{-11}$ . No significant two-way interaction was observed,  $F(1, 30) = 0.66$ ,  $ns$ ,  $d = 0.30$ , indicating that AD and control participants demonstrated an equivalent bias toward judging the source to be male.

Finally, we examined the extent to which each participant's source memory performance (as measured by the gender source

Table 2  
*Proportion of Correct Judgments in Cued Truth Task*

Item type	Participant group		Difference
	AD patients	Controls	
Cued-true	.69	.77	.08
Cued-false	.41	.61	.20

Note. AD = Alzheimer's disease.

Table 3  
*Proportion of Correct Source Judgments in Gender Source Task*

Item type	Participant group		Difference
	AD patients	Controls	
Cued-female	.47	.75	.28
Cued-male	.60	.83	.20

Note. AD = Alzheimer's disease.

task) correlated with his or her ability to suppress misattribution in the cued truth task, in which recollection is necessary to correctly judge cued-false items to be false. That is, treating recollection as an individual difference variable, we expected that the ability to correctly remember the voice in which statements were originally presented would correlate with the ability to correctly judge the truth value of cued-false statements. Because recollection is not needed to correctly judge cued-true items (which may be so judged on the basis of misattribution), no such correlation should be observed for cued-true items. Consistent with these predictions, we observed a significant correlation between participants' gender source performance and the proportion of cued-false items correctly judged to be false on the cued truth task,  $r(31) = .50$ ,  $p < .005$ . As displayed in the left panel of Figure 1, this correlation resulted primarily from the group differences between AD patients and controls. However, no such correlation was observed between gender source performance and the proportion of cued-true items correctly judged to be true,  $r(31) = -.01$ ,  $ns$  (Figure 1, right panel). Statistical comparison of these two correlations following Fisher's  $r$ -to- $Z$  transformation revealed that the two correlations differed significantly,  $Z = 2.05$ ,  $p < .05$ . As predicted, no significant correlation was observed between the gender source task and performance on the uncued truth task, in which recollection is not believed to influence misattribution errors.

## Discussion

The present study investigated the susceptibility of AD patients to make errors of misattribution, whereby one attributes traces of past experiences to an inappropriate source. Misattribution was indexed by two forms of the illusory truth effect, the tendency for perceivers to judge previously encountered information to be true (even when initially cued as false). Taken together, our results suggest that AD patients' tendency to commit such errors of misattribution depends critically on whether performance draws primarily on familiarity-based or recollection-based memory. When study items were uncued—a condition in which misattribution is thought to arise primarily from experimentally induced familiarity—AD patients proved relatively resilient to misattribution; that is, relative to healthy controls, AD patients were less likely to judge previously presented statements to be true. In contrast, when study items were cued as true or false—a condition in which misattribution can be prevented by recollection for details of the study episode—AD patients were marginally more likely

<sup>1</sup> Although marginally significant ( $p < .06$ ), this effect exceeded the cutoff recommended by Cohen (1988) for a medium-sized effect ( $d = 0.5$ ) and neared that cutoff for a large effect ( $d = 0.8$ ).

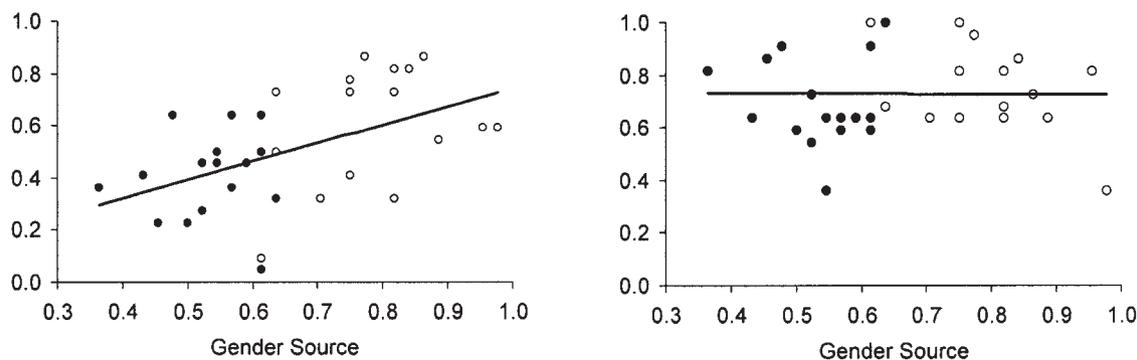


Figure 1. Relation between gender source and the proportion cued-false (left panel) and cued-true (right panel) items that were correctly judged in the cued truth task. Open circles represent Alzheimer's patients and the closed circles represent controls (one circle per participant).

than controls to succumb to misattribution. Critically, although both groups were as likely to correctly judge a cued-true item to be true, AD patients were significantly less likely to correctly judge a cued-false item to be true. In other words, AD patients were significantly more likely than controls to erroneously misattribute truth to an item that had initially been presented as false.

This pattern of results is highly consistent with both the deficits in recollection and familiarity previously observed in AD as well as the contributions made by these components of recognition memory to misattribution errors. However, despite this overall pattern of consistency with our a priori predictions, we note that these data should be interpreted somewhat cautiously given the marginal statistical reliability associated with the interaction for the uncued truth task and the nonsignificant interaction observed for the cued truth task. Nevertheless, in both cases, failure to achieve statistical significance seems most likely the result of the relatively small sample of participants. Specifically, the effect size associated with the two-way interaction in the uncued truth task well exceeded the cutoff for a medium-sized effect ( $d = 0.5$ ) and approached the cutoff for a large effect ( $d = 0.8$ ). Likewise, planned comparisons for cued truth data demonstrated the expected difference between AD patients and controls for cued-false items and the expected correspondence between groups for cued-true items. Future replication of these results, however, will be needed to firmly establish the reliability of these effects.

It is worthwhile to point out that familiarity and recollection are thought to contribute to misattribution errors in opposing ways, and that the performance of AD patients in this experiment directly reflects the counterforce of these two components of memory. Cognitive accounts of misattribution have generally posited that the tendency for misattribution increases as stimulus familiarity increases, but at least under some circumstances, misattribution decreases as recollection increases. Capitalizing on two manifestations of the illusory truth effect that rely differentially on familiarity and recollection, the present results are consistent with earlier suggestions that both these memory components are impaired in AD (Budson et al., 2000). Provocatively, these results suggest that AD patients are simultaneously both less and more susceptible to misattribution errors than healthy elderly controls. When misattribution depends solely on familiarity (i.e., when the presence or absence of recollection does not account for variability

in performance, as in the uncued condition), AD patients can be expected to demonstrate relatively few misattribution errors; however, when misattribution depends critically on the presence or absence of recollection (as in the cued condition), AD patients may be especially likely to demonstrate misattribution.

This latter point is underscored by the correlation between participants' ability to recollect the voice in which statements were presented (gender source, a frequently used task of source memory; Johnson et al., 1993) and their ability to prevent misattribution errors in the cued condition (cued truth). Specifically, performance on the gender source task correlated directly with the proportion of cued-false items that were later judged correctly as false, but not with the proportion of cued-true items that were later judged correctly. These results are consistent with the notion that recollection contributes critically to the ability to avoid misattribution errors in the cued condition; those participants who could consistently recollect the voice in which statements were presented were also those participants who could successfully avoid the tendency to misjudge cued-false statements as true. It is important to note that this correlation was obtained only for items for which performance depends on the presence of recollection (i.e., cued-false statements) but not for items for which performance could result from familiarity alone (cued-true statements) or for performance in the uncued truth task.

The deficits in recollection observed in the patients with AD in the present study may be attributable to pathology in their medial temporal lobes (Price & Morris, 1999). Amnesic patients with isolated lesions in the medial temporal lobes show deficits of recollection in different experimental paradigms (e.g., Cermak, Verfaellie, Sweeney, & Jacoby, 1992; Giovanello & Verfaellie, 2001; Giovanello, Verfaellie, & Keane, 2003). Although controversy still exists over whether the contributions of the hippocampus to recognition memory are specifically recollective (Manns & Squire, 1999; Reed, Hamann, Stefanacci, & Squire, 1997; Reed & Squire, 1997), patients with damage limited to the hippocampus show particularly pronounced impairment when memory performance requires recollection. Whether lesions restricted to the hippocampus are sufficient to impair recollection is less clear (Baddeley, Vargha-Khadem, & Mishkin, 2001; Holdstock et al., 2002; Mayes, Holdstock, Isaac, Hunkin, & Roberts, 2002; Yonelinas et al., 2002). These findings are complemented by recent

neuroimaging work (Henke, Weber, Kneifel, Wieser, & Buck, 1999; Jackson & Schacter, 2004; Sperling et al., 2001, 2003) that has demonstrated a role for the hippocampus in binding together contextual details into an associative recollection. Given the importance of the hippocampus to recollection, the relative inability of AD patients to suppress misattribution errors on the cued truth task may result from the known pathology in the medial temporal lobe in AD.

However, another possible explanation of these recollection-based deficits is that even mild AD patients have dysfunction of frontal networks, as suggested by earlier neuropsychological and neuroimaging research on AD (Baddeley, Bressi, Della Sala, Logie, & Spinnler, 1991; Dalla Barba et al., 1999; Haxby et al., 1988; Mountjoy, Roth, Evans, & Evans, 1983). Neuropsychological studies of patients with frontal lobe damage have strongly implicated the frontal cortex in recollection-based memory. For example, although frontal patients have shown relative sparing of item recognition, they typically demonstrate impaired use of recollective information (Janowsky, Shimamura, & Squire, 1989; Shimamura, Janowsky, & Squire, 1990). Patients with frontal lobe lesions are also impaired in their ability to use recollection to counter misattributions due to gist-based familiarity, whereas their recognition of studied items is unimpaired (Budson, et al., 2002). In addition, functional MRI research on normal populations has demonstrated greater prefrontal activation during the encoding of items later accompanied by contextual recollection, as indexed both by source memory (Davachi, Mitchell, & Wagner, 2003) or by "remember" responses in the remember/know paradigm (Henson, Rugg, Shallice, Josephs, & Dolan, 1999). At the same time, researchers have also reported greater activation in an extensive prefrontal region when comparing recognition tasks that require recollection to recognition tasks that do not require such reinstatement of the study context (Dobbins, Foley, Schacter, & Wagner, 2002; Dobbins, Rice, Wagner, & Schacter, 2003).

One idea for how frontal networks dysfunction may lead to misattribution errors in AD patients is related to the fact that intact frontal networks are needed for an efficient attentional-activation monitoring system. To distinguish whether test items on the cued truth task are true or false, discrimination between sources of activation at retrieval is necessary. Specifically, participants need to determine whether activation reoccurring at test is present because it was acquired prior to the experiment and is therefore true, acquired during the experiment as a true item, or acquired during the experiment as a false item. Because AD patients show impairments in their attentional-activation monitoring system (Balota et al., 1999; Spieler, Balota, & Faust, 1996; Sullivan, Faust, & Balota, 1995), they may have experienced difficulty in discriminating between these multiple sources of activation at test, leading them to overly attribute false items as true. Furthermore, if the AD patients' encoding during the uncued truth task was also impaired by a faulty attentional-activation monitoring system, then impairment of this system could also provide another explanation for the finding that AD patients were less likely than healthy controls to judge previously presented statements as true.

One recent functional MRI study (Mitchell et al., 2005) has directly linked hippocampal and prefrontal activation to the successful avoidance of misattribution errors. In this study, participants were scanned while encoding a series of trivia statements like the ones used in the current experiment. Some statements were

cued as true or false, whereas others were uncued. Subsequently, participants saw each statement a second time (along with a number of novel distractors) and were asked to indicate whether they believed each to be true, to be false, or to provide no basis for judging. Encoding activation in three regions—bilateral ventrolateral prefrontal regions and the left hippocampus—correlated with whether cued-false items were successfully judged to be false (but not whether cued-true items were judged to be true). These results provide additional evidence that successful avoidance of misattribution errors on the illusory truth task may rely on the function of hippocampal and prefrontal regions and are consistent with the notion that AD patients' particular susceptibility to recollection-based misattribution errors may be attributable to their frontal and medial temporal impairments.

Finally, the present results have implications for some of the memory distortions that AD patients experience in their everyday lives. The results of the cued truth condition suggested that, although AD patients are as likely as healthy older adults to correctly remember that a true item was true, they are impaired in their ability to remember that a false item was false. Indeed, AD patients correctly judged false items only 41% of the time; more than half of the time, they incorrectly believed an item that had originally been presented as false to be true. This observation may have profound implications for how clinicians and caregivers might optimally communicate information to AD patients. For example, the statement, "The 59 bus will not take you to your sister's house, take the 67 bus instead" can be thought of as two assertions, one cued false (the 59 bus *will not* take you to your sister's house) and one cued true (the 67 bus *will* take you to your sister's house). If this statement were communicated to a mild AD patient such as those who participated in the present study, our data suggest that she would be more likely to believe the false statement to be true than to remember that the false statement was false (and would be therefore quite likely to take the wrong bus). Thus, because AD patients appear particularly likely to think that any statement that they remember is true, clinicians and caregivers should be careful to provide patients with true statements only. Despite the goal of using false statements to provide additional helpful information, they are more likely to confuse, rather than aid, these patients.

## References

- Bacon, F. T. (1979). Credibility of repeated statements: Memory for trivia. *Journal of Experimental Psychology: Learning and Memory*, 5, 241–252.
- Baddeley, A. D., Bressi, S., Della Sala, S., Logie, R., & Spinnler, H. (1991). The decline of working memory in Alzheimer's disease. A longitudinal study. *Brain*, 114, 2521–2542.
- Baddeley, A. D., Vargha-Khadem, F., & Mishkin, M. (2001). Preserved recognition in a case of developmental amnesia: Implications for the acquisition of semantic memory? *Journal of Cognitive Neuroscience*, 13, 357–369.
- Balota, D. A., Cortese, M. J., Duchek, J. M., Adams, D., Roediger, H. L., McDermott, K. B., et al. (1999). Veridical and false memories in healthy older adults and in dementia of the Alzheimer's type. *Cognitive Neuropsychology*, 16, 361–384.
- Begg, I. M., Anas, A., & Farinacci, S. (1992). Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *Journal of Experimental Psychology: General*, 121, 446–458.

- Budson, A. E., Daffner, K. R., Desikan, R., & Schacter, D. L. (2000). When false recognition is unopposed by true recognition: Gist-based memory distortion in Alzheimer's disease. *Neuropsychology, 14*, 277–287.
- Budson, A. E., Desikan, R., Daffner, K. R., & Schacter, D. L. (2001). Perceptual false recognition in Alzheimer's disease. *Neuropsychology, 15*, 230–243.
- Budson, A. E., Sullivan, A. L., Mayer, E., Daffner, K. R., Black, P. M., & Schacter, D. L. (2002). Suppression of false recognition in Alzheimer's disease and in patients with frontal lobe lesions. *Brain, 125*, 2750–2765.
- Cermak, L. S., Verfaellie, M., Sweeney, M., & Jacoby, L. L. (1992). Fluency versus conscious recollection in the word completion performance of amnesic patients. *Brain and Cognition, 20*, 367–377.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum Associates.
- Dalla Barba, G., Nedjam, Z., & Dubois, B. (1999). Confabulation, executive functions, and source memory in Alzheimer's disease. *Cognitive Neuropsychology, 27*, 1–16.
- Davachi, L., Mitchell, J. P., & Wagner, A. D. (2003). Multiple routes to memory: Distinct medial temporal lobe processes build item and source memories. *Proceedings of the National Academy of Sciences, USA, 100*, 2157–2162.
- Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58*, 17–22.
- Dobbins, I. G., Foley, H., Schacter, D. L., & Wagner, A. D. (2002). Executive control during episodic retrieval: Multiple prefrontal processes subservise source memory. *Neuron, 35*, 989–996.
- Dobbins, I. G., Rice, H. J., Wagner, A. D., & Schacter, D. L. (2003). Memory orientation and success: Separable neurocognitive components underlying episodic recognition. *Neuropsychologia, 41*, 318–333.
- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*, 189–198.
- Gilbert, D. T. (1991). How mental systems believe. *American Psychologist, 46*, 107–119.
- Gilbert, D. T. (1993). The assent of man: Mental representation and the control of belief. In D. M. Wegner & J. W. Pennebaker (Eds.), *Handbook of Mental Control* (pp. 57–87). Englewood Cliffs, NJ: Prentice Hall.
- Gilbert, D. T., Krull, D. S., & Malone, P. S. (1990). Unbelieving the unbelievable: Some problems in the rejection of false information. *Journal of Personality and Social Psychology, 59*, 601–613.
- Giovanello, K. S., & Verfaellie, M. (2001). The relationship between recall and recognition in amnesia: Effects of matching recognition between patients with amnesia and controls. *Neuropsychology, 15*, 444–451.
- Giovanello, K. S., Verfaellie, M., & Keane, M. M. (2003). Disproportionate deficit in associative recognition relative to item recognition in global amnesia. *Cognitive Affective and Behavioral Neuroscience, 3*, 186–194.
- Haxby, J. V., Grady, C. L., Koss, E., Horwitz, B., Schapiro, M., Friedland, R. P., et al. (1988). Heterogeneous anterior-posterior metabolic patterns in dementia of the Alzheimer type. *Neurology, 38*, 1853–1863.
- Henke, K., Weber, B., Kneifel, S., Wieser, H. G., & Buck, A. (1999). Human hippocampus associates information in memory. *Proceedings of the National Academy of Sciences, USA, 96*, 5884–5889.
- Henson, R. N. A., Rugg, M. D., Shallice, T., Josephs, O., & Dolan, R. J. (1999). Recollection and familiarity in recognition memory: An event-related functional magnetic resonance imaging study. *Journal of Neuroscience, 19*, 3962–3972.
- Hodges, J. R. (2000). Memory in the dementias. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford Handbook of Memory* (pp. 441–459). Oxford, United Kingdom: Oxford University Press.
- Holdstock, J. S., Mayes, A. R., Roberts, N., Cezayirli, E., Isaac, C. L., O'Reilly, R. C., et al. (2002). Under what conditions is recognition spared relative to recall after selective hippocampal damage in humans? *Hippocampus, 12*, 341–351.
- Jackson, O., & Schacter, D. L. (2004). Encoding activity in anterior medial temporal lobe supports subsequent associative recognition. *NeuroImage, 21*, 456–462.
- Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30*, 513–541.
- Jacoby, L. L., Kelley, C., Brown, J., & Jasechko, J. (1989). Becoming famous overnight: Limits on the ability to avoid unconscious influences of the past. *Journal of Personality and Social Psychology, 56*, 326–338.
- Janowsky, J. S., Shimamura, A. P., & Squire, L. R. (1989). Source memory impairment in patients with frontal lobe lesions. *Neuropsychologia, 27*, 1043–1056.
- Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114*, 3–28.
- Manns, J. R., & Squire, L. R. (1999). Impaired recognition memory on the Doors and People Test after damage limited to the hippocampal region. *Hippocampus, 9*, 495–499.
- Mayes, A. R., Holdstock, J. S., Isaac, C. L., Hunkin, N. M., & Roberts, N. (2002). Relative sparing of item recognition memory in a patient with adult-onset damage limited to the hippocampus. *Hippocampus, 12*, 325–340.
- McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology, 34*, 939–944.
- Mitchell, J. P., Dodson, C. S., & Schacter, D. L. (2005). fMRI evidence for the role of recollection in suppressing misattribution errors: The illusory truth effect. *Journal of Cognitive Neuroscience, 17*, 800–810.
- Mountjoy, C. Q., Roth, M., Evans, N. J., & Evans, H. M. (1983). Cortical neuronal counts in normal elderly controls and demented patients. *Neurobiology of Aging, 4*, 1–11.
- Price, J. L., & Morris, J. C. (1999). Tangles and plaques in nondemented aging and “preclinical” Alzheimer's disease. *Annals of Neurology, 45*, 358–368.
- Reed, J. M., Hamann, S. B., Stefanacci, L., & Squire, L. R. (1997). When amnesic patients perform well on recognition memory tests. *Behavioral Neuroscience, 111*, 1163–1170.
- Reed, J. M., & Squire, L. R. (1997). Impaired recognition memory in patients with lesions limited to the hippocampal formation. *Behavioral Neuroscience, 111*, 667–675.
- Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 803–824.
- Schacter, D. L. (1999). The seven sins of memory: Insights from psychology and cognitive neuroscience. *American Psychologist, 54*, 182–203.
- Schacter, D. L., Verfaellie, M., & Anes, M. D. (1997). Illusory memories in amnesic patients: Conceptual and perceptual false recognition. *Neuropsychology, 11*, 331–342.
- Schacter, D. L., Verfaellie, M., Anes, M. D., & Racine, C. (1998). When true recognition suppresses false recognition: Effects of amnesia. *Journal of Cognitive Neuroscience, 10*, 668–679.
- Schacter, D. L., Verfaellie, M., & Pradere, D. (1996). The neuropsychology of memory illusions: False recall and recognition in amnesic patients. *Journal of Memory and Language, 35*, 319–334.
- Shimamura, A. P., Janowsky, J. S., & Squire, L. R. (1990). Memory for the temporal order of events in patients with frontal lobe lesions and amnesic patients. *Neuropsychologia, 28*, 803–813.

- Sperling, R. A., Bates, J. F., Cocchiarella, A. J., Schacter, D. L., Rosen, B. R., & Albert, M. S. (2001). Encoding novel face-name associations: A functional MRI study. *Human Brain Mapping, 14*, 129–139.
- Sperling, R. A., Chua, E., Cocchiarella, A., Rand-Giovannetti, E., Poldrack, R., Schacter, D. L., et al. (2003). Putting names to faces: Successful encoding of associative memories activates the anterior hippocampal formation. *NeuroImage, 20*, 1400–1410.
- Spieler, D. H., Balota, D. A., & Faust, M. E. (1996). Stroop performance in healthy younger and older adults and in individuals with dementia of the Alzheimer's type. *Journal of Experimental Psychology: Human Perception and Performance, 22*, 461–479.
- Sullivan, M. P., Faust, M. E., & Balota, D. A. (1995). Identity negative priming in older adults and individuals with dementia of the Alzheimer type. *Neuropsychology, 9*, 537–555.
- Verfaellie, M., Schacter, D. L., & Cook, S. P. (2002). The effect of retrieval instructions on false recognition: Exploring the nature of the gist memory impairment in amnesia. *Neuropsychologia, 40*, 2360–2368.
- Yonelinas, A. P., Kroll, N. E., Quamme, J. R., Lazzara, M. M., Sauve, M. J., Widaman, K. F., et al. (2002). Effects of extensive temporal lobe damage or mild hypoxia on recollection and familiarity. *Nature Neuroscience, 5*, 1236–1241.

Received May 18, 2004

Revision received October 13, 2004

Accepted January 18, 2005 ■