Journal of Experimental Psychology:
Learning, Memory, and Cognition
2007, Vol. 33, No. 1, 169–181

0278-7393/07/$12.00 DOI: 10.1037/0278-7393.33.1.169

Aging, Source Memory, and Misrecollections

Chad S. Dodson and Sameer Bawa
University of Virginia

Scott D. Slotnick
Boston College

The authors propose an illusory recollection account of why cognitive aging is associated with episodic memory deficits. After listening to statements presented by either a female or a male speaker, older adults were prone to misrecollecting past events. The authors' illusory recollection account is instantiated in a new illusory recollection signal detection model that provides a better fit of older adults' data than does the standard signal detection model. They observed that age-related differences in source memory (as measured by source d' scores) virtually disappear after accounting for the occurrence of illusory recollections. These data suggest that age-related source memory impairments are not due to older adults' remembering less diagnostic source information and having to guess more. Instead, older adults appear to misremember past events more often than younger adults.

Keywords: recollection, signal detection theory, false memory, source memory, aging

Older adults have difficulty remembering source information—specific information about the circumstances under which an event was encountered (e.g., Johnson, Hashtroudi, & Lindsay, 1993). They are generally less accurate than younger adults at remembering which of two speakers presented an item (e.g., McIntyre & Craik, 1987; Schacter, Kaszniak, Kihlstrom, & Valdiserri, 1991; Siedlecki, Salthouse, & Berish, 2005; Simons, Dodson, Bell, & Schacter, 2004), whether an event was seen in a videotape or a photograph (Schacter, Koutstaal, Johnson, Gross, & Angell, 1997), and whether an event was thought about or spoken aloud (Hashtroudi, Johnson, & Chrosniak, 1989). Older adults are also less likely to correctly remember contextual features of events, such as their color or location (e.g., Chalfonte & Johnson, 1996; Kausler & Puckett, 1980; Park & Puglisi, 1985).

Although there are a variety of different accounts for this age-related source memory impairment, many of them share the view that older adults remember (or rely on) fewer source-identifying characteristics than the young. Consequently, older adults must more often guess or base a response on less specific information, such as familiarity (e.g., Anderson & Craik, 2000; Balota, Dolan, & Duchek, 2000; Castel & Craik, 2003; Dodson & Schacter, 2002; Glisky, Rubin, & Davidson, 2001; Jacoby, 1999; Johnson et al., 1993; Memon, Bartlett, Rose, & Gray, 2003; Schacter, Norman, & Koutstaal, 1998). Consistent with this account, several studies have shown that older adults are less likely than younger adults to "remember" specific details about past events (e.g., Gardiner & Richardson-Klavehn, 2000). Moreover, it is generally thought that the reduced ability to remember source information prevents older adults from counteracting false memories about events that never occurred (e.g., Dodson, Koutstaal, & Schacter, 2000; Schacter et al., 1998). We refer to this theory of age-related source memory deficits as the *reduced memory hypothesis*.

We propose an alternative account of older adults' source memory impairment: the *misrecollection hypothesis*. We suggest that older adults are prone to experience convincing false recollections. The critical element in this account is the term *convincing*. Thus, whereas the reduced memory hypothesis traces older adults' source memory deficits to diminished use of specific information, the misrecollection account argues that they have a propensity to miscombine features of different events that causes high-confidence source monitoring errors as well as other false memories (for related binding views, see Chalfonte & Johnson, 1996; Henkel, Johnson, & DeLeonardis, 1998; Koutstaal, Schacter, & Brenner, 2001; Kroll, Knight, Metcalfe, Wolf, & Tulving, 1996; Lyle & Johnson, 2006). The novel element of our account is that miscombined binding of features in older adults produces false memories that are as confidently held as true memories. Although we are contrasting our misrecollection account with the reduced memory account, these two hypotheses are not mutually exclusive. It is possible that both accounts can describe some kinds of age-related memory impairments.

The Misrecollection Signal Detection Model

Consider a typical source memory experiment in which either a male or a female voice presents words during an encoding phase. Source memory for these words can be represented using a standard signal detection model with two normal distributions in decision space, shown in Figure 1a, which represents the amount and/or quality of the information that individuals remember about each word (e.g., Banks, 2000; DeCarlo, 2003; Glanzer, Hilford, & Kim, 2004; Johnson et al., 1993; Qin, Raye, Johnson, & Mitchell, 2001; Slotnick & Dodson, 2005; Slotnick, Klein, Dodson, & Shimamura, 2000). This model is formally described by the pa-

---

Chad S. Dodson and Sameer Bawa, Department of Psychology, University of Virginia; Scott D. Slotnick, Department of Psychology, Boston College.

We thank David Riefer and Bill Rohwer for their helpful comments. We are also grateful for support from the Institute on Aging at the University of Virginia.

Correspondence concerning this article should be addressed to Chad S. Dodson, Department of Psychology, P.O. Box 400400, 102 Gilmer Hall, University of Virginia, Charlottesville, VA 22904-4400. E-mail: cdodson@virginia.edu

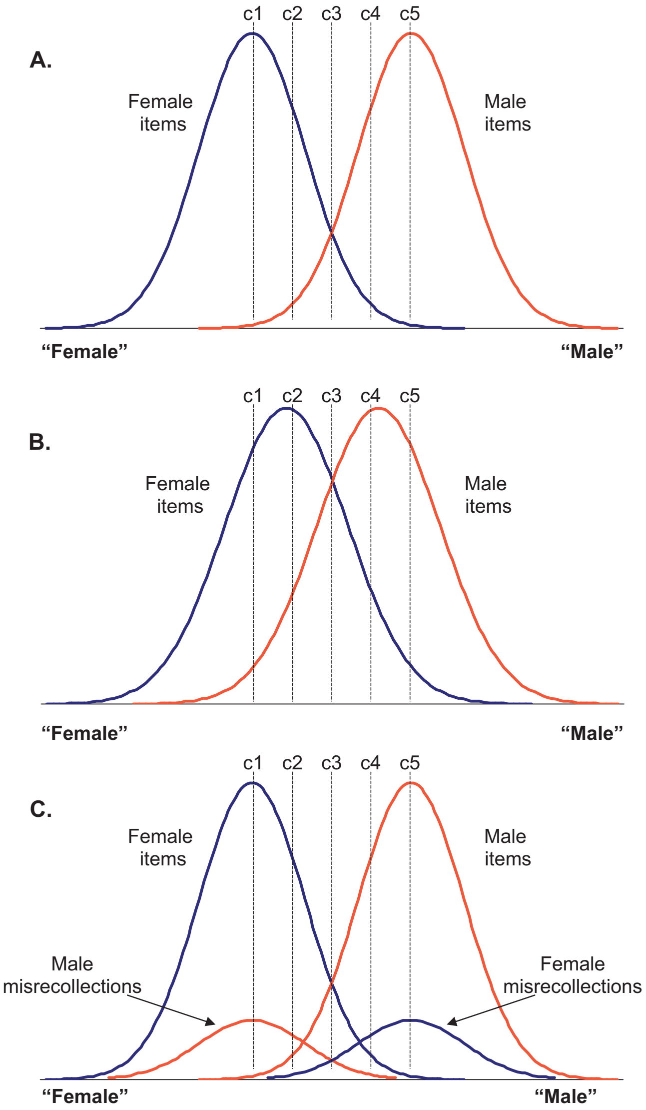

Figure 1. Standard signal detection conception of source memory (depicting memory strength of male and female items in decision space) in paradigmatic younger (A) and older adults (B) and the illusory recollection signal detection conception of source memory in older adults (C). Lines marked c1 through c5 reflect criteria for confidence ratings.

rameters d', the distance between the distribution means, and σ_{s1}/σ_{s2}, the ratio of the standard deviations for the two distributions:

$$y_{s1} = \frac{1\ \ \sigma_{s1}}{\sqrt{2\pi}\sigma_{s2}} e^{-1/2(\sigma_{s1}/\sigma_{s2}(x-d'/2))^2} \qquad (1)$$

$$y_{s2} = \frac{1}{\sqrt{2\pi}} e^{-1/2(x+d'/2)^2}, \qquad (2)$$

where x reflects source memory strength (the x-axis in Figure 1) and y reflects the probability of experiencing words from Source 1 (s1) or Source 2 (s2) with that memory strength (the y-axis in Figure 1). This model is referred to as the unequal variance (UEV) signal detection model because the standard deviations of the distributions can be unequal (i.e., the ratio of standard deviations is not restricted to be 1). The distribution endpoints refer to information that is highly diagnostic of the female voice and male voice, respectively. By contrast, the middle of the figure represents information that is less diagnostic about the speaker. The vertical lines refer to response criteria for making a source judgment. For instance, items of the two distributions that lie to the right of line c5 would receive a confidence rating of 6 (i.e., very confident that the male speaker presented the word) and items lying to the left of c1 would receive a confidence rating of 1 (i.e., very confident that the female speaker presented the word). This characterization of source memory as a signal detection process has been validated by a number of reliable findings, such as that receiver operating characteristics (ROCs) are curvilinear and *z*-transformed ROCs (*z*-ROCs) are linear or slightly curvilinear (the curvature can stem from collapsing over recognition confidence ratings; cf. Slotnick & Dodson, 2005), indicating that retrieval of source information is a continuous process (e.g., Banks, 2000; Glanzer et al., 2004; Hilford, Glanzer, Kim, & DeCarlo, 2002; Qin et al., 2001; Slotnick & Dodson, 2005; Slotnick et al., 2000). We use Figure 1a to illustrate source memory of a typical younger adult.

With respect to the finding that older adults generally show worse source memory than younger adults, the standard signal detection explanation corresponds to the foregoing reduced memory account. That is, according to this view a typical older adult is depicted in Figure 1b. An examination of Figures 1b and 1a shows that they differ in terms of the closeness and amount of overlap of the two distributions. In other words, an older adult experiences difficulty remembering who spoke a word because this individual is less likely than a typical younger adult to remember information that is highly diagnostic of one source or the other.

In contrast to the reduced memory account, Figure 1c depicts our proposed illusory recollection (IR) account of source memory. This account holds the signal detection view that retrieval of source information is a continuous process. However, the difference between the IR model and standard signal detection models is the occurrence of misrecollections. By *misrecollection* we refer to the subjective experience of misremembering the source of an event, such as incorrectly remembering that the male voice spoke a word that was actually spoken by the female voice. For example, in Figure 1c the red-colored distribution, which corresponds to memory for items spoken by the male voice, has one large peak located in the right area of the x-axis that corresponds to correct responses that the male speaker was the source. Critically, there is another smaller peak of misrecollections within the entire red distribution of items presented by the male speaker that is located in the left area of the x-axis and corresponds to incorrect responses that the female speaker was the source. The distribution of these male misrecollections is located at the same mean distance on the x-axis and has the same standard deviation as the distribution of items presented by the female speaker that are correctly attributed to the female speaker (i.e., the large blue peak located in the left area of the x-axis). Likewise, there is a small peak in the female distribution (i.e., small blue peak in the right area of the x-axis) reflecting female misrecollections.

Our misrecollection model is formally described by the parameters d', the distance between the distribution means; σ_{s1}/σ_{s2}, the ratio of distribution standard deviations; and ζ, the proportion of items that are misrecollected:

$$y_{s1} = (1-\zeta)\frac{1}{\sqrt{2\pi}}\frac{\sigma_{s1}}{\sigma_{s2}}e^{-1/2(\sigma_{s1}/\sigma_{s2}(x-d'/2))^2} + \zeta\frac{1}{\sqrt{2\pi}}e^{-1/2(x+d'/2)^2} \tag{3}$$

$$y_{s2} = (1-\zeta)\frac{1}{\sqrt{2\pi}}e^{-1/2(x+d'/2)^2} + \zeta\frac{1}{\sqrt{2\pi}}\frac{\sigma_{s1}}{\sigma_{s2}}e^{-1/2(\sigma_{s1}/\sigma_{s2}(x-d'/2))^2}. \tag{4}$$

The latter term in Equations 3 and 4 reflects the degree of misrecollections—that is, the proportion of items that are misrecollected as originating from the incorrect source. Our misrecollection model builds on the ideas of DeCarlo (2002), who has shown that signal detection theory can be extended by incorporating a mixture of underlying distributions. It is important to emphasize that the misrecollection model becomes the standard UEV model when ζ—the proportion of items that are misrecollected—is zero. Although it is just a modification of the standard UEV signal detection model, we refer to this as the IR model to highlight the additional misrecollection parameter. It is also important to note that because the UEV model and the IR model are nested (i.e., the IR model becomes the UEV model when ζ is equal to zero), these models can be directly compared with each other.

The misrecollection model and the standard signal detection model offer different accounts with regard to the effects of cognitive aging on source memory. Critically, these models make different and testable predictions as to the shape of the source memory ROC/*z*-ROC. Figure 2 (left panel) shows illustrative source ROCs that correspond to the UEV model (dashed line) and the IR model (solid line). Each source ROC was created by sweeping a criterion across decision space (e.g., Figures 1b and 1c) and plotting the source hit rate versus the source false alarm rate. It is notable that the IR model ROC, as compared with the UEV model ROC, has relatively higher source hit rates at lower-intermediate source false alarm rates (i.e., it protrudes toward the upper left corner) and has relatively lower source hit rates at lower and higher false alarm rates. These deviations follow directly from the underlying distributions; for example, the IR model protrusion is due to the relative increase in source hit rates associated with summation of the standard and misrecollection distributions at intermediate source false alarm rates (see Figure 1c). It follows then that a distinct prediction of the IR model is that the *z*-ROC will show negative curvature, as illustrated in Figure 2 (right panel). This negative curvature reflects the pattern that the IR model *z*-ROC will have relatively higher source hit rate values at lower-intermediate false alarm rate values and will otherwise have relatively lower source hit rate values. By contrast, the *z*-ROC corresponding to the UEV model is always linear, as illustrated by the dashed line in Figure 2 (right panel).

To test our misrecollection hypothesis, we fit the IR model and the standard UEV model to the source ROCs that have been derived from the data of older adults and younger adults. If our hypothesis is correct, the IR model should provide a significantly better fit of the older adults' data than the standard UEV model. In addition, the misrecollection parameter (ζ) associated with the best fit IR model should be significantly greater than zero for the older adults' data. To complement the model-fitting results, we also assess whether the *z*-ROC for older adults has curvature; if our hypothesis is correct, this curvature will be negative (Figure 2, right panel). It should be mentioned that the negative *z*-ROC curvature associated with the IR model is distinct from the linear *z*-ROC associated with the UEV model and in direct opposition to the positive *z*-ROC curvature associated with the two-high-threshold model (the latter model is evaluated to complement previous work; for a review, see Slotnick & Dodson, 2005).

We would also like to underscore the exciting possibility that is represented by our misrecollection account: Older adults' impaired source memory performance may disappear after controlling for the occurrence of misrecollections. In other words, under some circumstances, older and younger adults' source d' scores may be compa-

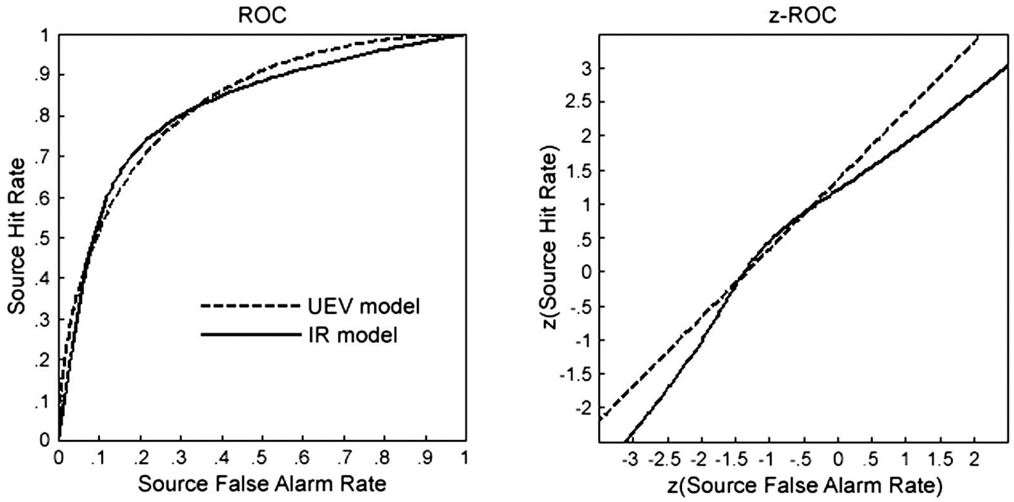

Figure 2. Receiver operating characteristics (ROCs) illustrating the unequal variance (UEV) model and the illusory recollection (IR) model (left panel) and corresponding *z*-transformed ROCs (*z*-ROCs) (right panel).

rable when misrecollections have been factored out. This possibility is visible by comparing the distance between the two large peaks in Figure 1a, which depicts a hypothetical younger adult, and Figure 1c, which depicts the misrecollection account of a hypothetical older adult. This outcome would suggest that misrecollections completely cause age-related source memory impairments.

Experiment 1

To test the foregoing predictions of the IR model, we used a typical source monitoring paradigm. Participants heard sentences presented by either a male or a female speaker, and on a subsequent source identification test they identified who had presented the sentence during the study phase. They also rated their confidence in their response. Finally, when investigating the IR model, it is critical to control for age-related differences in overall source accuracy. To match source accuracy between the older and younger adults, we introduced a 1-day delay between the study and test phases for the younger adults. Matching source accuracy means that whatever differences we observe in the occurrence of misrecollections on the part of the two age groups cannot be attributed to differences in source accuracy.

Method

Participants. Eighteen undergraduates (range of 18–23 years) participated in the young–delay condition, and 18 older adults (range of 60–80 years) participated in the older–1Rep condition. Table 1 contains demographic information for all participants. We assessed overall cognitive ability by examining performance on the logical memory and word-list recall tests of the Wechsler Memory Scale (Wechsler, 1997b) and the vocabulary test of the Wechsler Adult Intelligence Scale—Third edition (Wechsler, 1997a). As seen in Table 1, older adults remembered less than younger adults on the logical memory test, $t(34) = 2.48$, $p < .05$, and the word-list recall test, $t(34) = 1.62$, although this latter difference was not significant. On the vocabulary test, older adults performed slightly but nonsignificantly better than younger adults, $t(34) = 0.37$. Overall, these results are consistent with the general findings in the literature of age-related differences in measures of fluid intelligence (e.g., logical memory test) and lack of age-related differences (or even age-related increases) in measures of crystallized intelligence (e.g., vocabulary tests).

Materials. We selected all study and test items from the same pool of trivia statements used by Simons et al. (2004). For instance, one statement was “Al Capone’s business card said he was a used furniture dealer.”

Procedure. The study phase consisted of 80 statements and 8 filler statements—4 at the beginning and 4 at the end of the list—that were included to control for primacy and recency effects. Each statement appeared visually on a computer screen for 7 s and was also simultaneously spoken once by either a male or a female speaker over headphones. Each speaker presented half of the statements. The statements appeared in a random order with the constraint that not more than two statements from a particular speaker could occur consecutively. Above each statement was a picture of the speaker, below which was printed the speaker’s name. Participants received instructions to pay attention to who presented each statement as there would be a subsequent memory test. To match source memory between the young–delay and older–1Rep adult groups, the older adults began the source monitoring task after a 5-min study–test delay, whereas the young–delay adults departed following the study phase and returned 24 hr later to complete the test phase.

The test phase consisted of 120 statements: 40 statements presented by the male speaker during the study phase, 40 statements presented by the female speaker during the study phase, and 40 new statements. We included an additional 4 statements at the beginning of the test to demonstrate the task to participants. Test statements appeared visually. Below each statement were pictures of both speakers with their corresponding names printed just below each image.

We asked participants to make a sequence of responses for each test statement. First, they were to indicate whether the test statement was originally spoken by the male voice, was originally spoken by the female voice, or was new. Then, they were to express their confidence in the accuracy of their response using a 6-point scale from 50 to 100 (i.e., 50, 60, 70, 80, 90, and 100), where 50 indicated they were guessing and 100 indicated they were certain. We instructed participants to base their confidence ratings on their source memory (rather than on their old–new memory) for each test statement, and we encouraged them to use the entire scale of confidence ratings across the entire test.[1]

Results

Matching source identification performance. Table 2 displays overall memory performance. We measured source memory performance with conditional source identification scores. These scores were based on the proportion of studied items recognized as old that were attributed to the correct source. There were no significant differences between the older–1Rep (.66) and young–delay (.64) groups, $t(34) = 0.66$.

We analyzed item memory by examining hit rates to studied items, false alarm rates to new items, and recognition d' scores (derived from the hit rates and false alarm rates).[2] Not surprisingly, the older–1Rep group exhibited higher hit rates, $t(34) = 1.99$, $p = .055$; lower false alarm rates, $t(34) = 4.29$, $p < .01$; and higher recognition d' scores, $t(34) = 4.34$, $p < .01$, than the young–delay group. Thus, the steps taken to match the age groups on source memory performance—that is, introducing a 1-day delay for younger adults—had the effect of producing lower recognition rates among the younger participants in comparison with their older counterparts.

Testing signal detection models of source memory. Testing the misrecollection account involves estimating source memory ROC and z-ROC parameters associated with fitting the standard signal detection model and our misrecollection model to the young and older adults’ data. We constructed source ROCs based on the confidence rating data, as shown in the Appendix, that participants gave for their source judgments. We constructed z-ROCs by taking the z transformation of the corresponding ROCs (i.e., converting probabilities into standard deviation units; see Macmillan & Creelman, 1991).

If older adults are prone to experiencing false recollections, then our misrecollection model should provide a significantly better fit to their data than will standard signal detection models of source

---

[1] Additionally, after making their confidence judgment, participants saw their initial response and decided either to submit or withhold it for scoring. They were told that only responses they submitted would be scored for accuracy. In both experiments, we observed that younger adults significantly improved their recognition and source scores by withholding items, in contrast to the lack of significant improvement on these measures by older adults. Younger adults, however, withheld more items than did older adults. In the interest of brevity and to maintain focus on the central predictions of the model, we do not consider the submit–withhold manipulation. That is, we based all analyses on all of the responses, regardless of the submit–withhold decision.

[2] Following the recommendation of Snodgrass and Corwin (1988), d' scores were computed from hit rates and false alarm rates that had been transformed by adding .5 to the raw score and dividing by $N + 1$.

Table 1
Demographic Characteristics and Psychometric Scores for Younger and Older Adults

Group	Age	Years of ed.	Logical memory[a]	Word-list recall[a]	Vocabulary[b]
Experiment 1					
Young–delay	21	16	49.7 (2.4)	37.1 (1.5)	55.5 (1.4)
Older–1Rep	66	18	41.5 (2.1)	34.2 (1.0)	56.3 (1.5)
Experiment 2					
Young–1Rep	20	14	47.9 (1.7)	38.3 (1.0)	52.3 (1.0)
Older–2Rep	70	16	44.7 (2.1)	30.3 (1.6)	50.8 (2.5)
Older–3Rep	68	16	41.2 (2.6)	34.0 (1.3)	54.7 (1.2)

Note. Values are means, with standard errors in parentheses. Ed. = education.
[a] Subset of the Wechsler Memory Scale. [b] Subset of the Wechsler Adult Intelligence Scale—Third edition.

memory. To test this prediction, we used maximum likelihood procedures to fit both standard UEV signal detection and IR models to the source ROC data. Specifically, model parameters were adjusted to minimize the difference between empirically observed source hit rates and model-estimated source hit rates, with these differences used to compute chi-square (χ^2) values (and parameter standard errors are estimated as part of the model-fitting procedure; Press, Teukolsky, Vetterling, & Flannery, 1992). As mentioned earlier and as shown in Table 3, each of these models has a source d' parameter that measures source discrimination and a σ_{s1}/σ_{s2} parameter that measures the ratio of the standard deviations for the two distributions. In addition, the misrecollection model contains a parameter (ζ) that measures the proportion of items from a particular source that are misrecollected as originating from the other source. In other words, the standard signal detection model is nested within the misrecollection model because when the misrecollection parameter equals zero, both models are formally identical to each other (i.e., Equations 3 and 4 reduce to Equations 1 and 2). This nesting between the two models means that they can be directly compared with each other in terms of how well they fit the data. Finally, we also fit a two-parameter two-high-threshold model to the data in order to assess its validity and to allow for comparisons with previous research (for a review, see Slotnick & Dodson, 2005). Degrees of freedom associated with each model are equal to the number of ROC points minus the number of model parameters.

Table 2
Source Scores, Hit Rates, False Alarm Rates, and Recognition d′ Scores

Experiment	Source score	Hit rate	False alarm rate	Recognition d'
Experiment 1				
Young–delay	.64 (.02)	.83 (.02)	.11 (.02)	2.29 (0.17)
Older–1Rep	.66 (.03)	.89 (.02)	.02 (.01)	3.23 (0.14)
Experiment 2				
Young–1Rep	.79 (.02)	.88 (.02)	.03 (.01)	3.04 (0.16)
Older–2Rep	.76 (.03)	.96 (.01)	.02 (.01)	3.70 (0.14)
Older–3Rep	.81 (.03)	.99 (.01)	.01 (.003)	4.46 (0.10)

Note. Parentheses contain the standard error of the mean.

Table 3 shows the parameter values of the standard UEV and IR models. For the young–delay adults' data, the standard signal detection model provided an adequate fit, $\chi^2(9) = 15.26$, and the misrecollection model did not provide a fit that was any better than the standard signal detection model (i.e., there was no significant difference between the two models in the chi-square values associated with how well each model fits the data; $\chi^2[1] < 1$). This is not surprising because the misrecollection parameter (.02) for the young–delay adults' data was not significantly different from zero, $t(8) < 1$, which means that the misrecollection model is formally equivalent to the standard signal detection model. Thus, young–delay adults show few—if any—misrecollections in this paradigm, as measured by the misrecollection model.

By contrast, the misrecollection model fit the older adults' data to a significantly better extent than did the standard signal detection model, $\chi^2(1) = 14.04$, $p < .001$, with both models providing an adequate fit of the data: UEV, $\chi^2(9) = 15.70$, $p = .073$; IR, $\chi^2(8) = 1.66$, $p > .20$. In addition, the value of .23 for the misrecollection parameter indicates that older adults misrecollected 23% of the items—a value that is significantly greater than zero, $t(8) = 9.24$, $p < .001$. Finally, although not shown in Table 3, the two-high-threshold model did not fit either the younger or older adults' source data, both $\chi^2(9) > 340$, $ps < .001$, which is consistent with previous findings in the literature (for a review, see Slotnick & Dodson, 2005).

We also conducted an individual-participant chi-square analysis to ensure the results were not due to averaging effects, and the

Table 3
Parameter Values From the Standard Unequal Variance (UEV) and the Illusory Recollection (IR) Signal Detection Models for the Young and Older Adults' Data in Experiment 1

Condition	Source d'	σ_{s1}/σ_{s2}	Misrec
Young–delay			
Standard UEV model	0.77 (0.05)	1.04 (0.03)	
IR model	0.81 (0.25)	1.05 (0.09)	.02 (.15)
Older–1Rep			
Standard UEV model	0.70 (0.05)	1.04 (0.03)	
IR model	1.50 (0.10)	1.04 (0.06)	.23 (.03)

Note. Parentheses contain the standard error of the mean. Misrec = Misrecollection parameter, which refers to the proportion of items that are misremembered as originating from the incorrect source.

identical pattern of results was observed. Participants with source d' values within two standard errors of zero were excluded from the analysis, as data close to chance cannot be used to distinguish between the models (4 participants were excluded from each group). For the young–delay adults' data the standard signal detection model provided an adequate fit, $\chi^2(126) = 118.79$, $p > .20$, and the misrecollection model did not provide a fit that was any better than the standard signal detection model, $\chi^2(14) < 14$. However, the misrecollection model fit the older adults' data to a significantly better extent than did the standard signal detection model, $\chi^2(14) = 27.21$, $p < .05$, with both models providing an adequate fit of the data: UEV, $\chi^2(126) = 129.95$, $p > .20$; IR, $\chi^2(112) = 102.74$, $p > .20$. The two-high-threshold model did not fit either the younger or the older adults' source data: both $\chi^2(126) > 1{,}110$, $ps < .001$.

Examining the shape of the source z-ROC. Figure 3 shows the *z*-ROCs from the young–delay (Figure 3, top panel) and older–1Rep (Figure 3, bottom panel) adults' data. If the hypothesis under investigation is correct, misrecollections will produce a negative curvature of the *z*-ROC. Testing this prediction requires a linearity analysis, which involves fitting a line to the *z*-ROC and then, separately, fitting a line plus a quadratic to the *z*-ROC and determining whether the quadratic significantly improves the fit (i.e., assessing whether the *z*-ROC has significant curvature). Regression is conducted in the *x*- and *y*-direction where parameters correspond to the direction with the lowest sum-of-squares error between the model and the data. It is important to note that testing for quadratic *z*-ROC curvature is an indirect measure (as compared with chi-square analysis, which directly compares model estimates and observed values), as the *z*-ROC curvature associated with illusory recollections is not necessarily quadratic (e.g., Figure 2, right panel; Slotnick & Dodson, 2005). Table 4 displays the linear analysis values for the *z*-ROCs and ROCs. With regard to our prediction, curvature is measured by the c_{quad} parameter and is shown on the right side of Table 4, with negative values corresponding to negative curvature. As evident in Figure 3 (top panel), for the young–delay group the linearity analysis showed that there was significant *positive* curvature of the *z*-ROC ($c_{\text{quad}} = .09$, $p < .001$). Critically, as seen in Figure 3 (bottom panel), the *z*-ROCs for the older adults' data showed a significant *negative* curvature ($c_{\text{quad}} = -.10$, $p < .001$), consistent with the prediction of our IR model.

Finally, we analyzed the shape of the ROC, with the results shown in the bottom half of Table 4. As with the *z*-ROC analysis, this analysis involves fitting a line as well as a line plus quadratic to the ROC. A ROC is considered linear if the fit of the line and quadratic is not significantly different from the fit of the line alone (i.e., adding a quadratic does nothing to improve the fit). By contrast, the ROC is considered curvilinear if the addition of the quadratic significantly improves the fit. This analysis of the shape of the ROCs showed that for both younger and older adults, source ROCs were curvilinear (i.e., all c_{quad} values < 0, $ps < .05$), which is inconsistent with threshold models of source memory but is consistent with extant signal detection models (e.g., Glanzer et al., 2004; Qin et al., 2001; Slotnick & Dodson, 2005).

Overall, the data fit the predictions of our misrecollection account. It is difficult to attribute the greater occurrence of misrecollections on the part of older adults than young–delay adults to differences in overall source memory, as both groups were well matched on their source memory for the different items. Instead, it appears that cognitive aging is associated with a vulnerability to experience illusory recollections. These illusory recollections contributed to the poor fit of the standard signal detection model but were well captured by our misrecollection model.

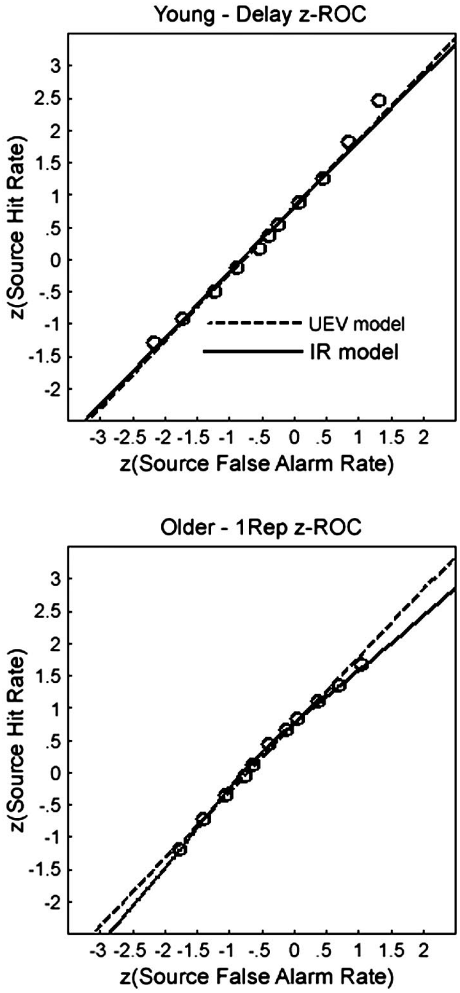

Figure 3. Source memory *z*-ROC with best fit unequal variance (UEV) model and illusory recollection (IR) model for the young–delay group (top panel) and the older–1Rep group (bottom panel) from Experiment 1. *z*-ROC = *z*-transformed receiver operating characteristics.

Experiment 2

Experiment 1 shows that when older and younger adults have relatively poor source memory for past events, older adults are prone to misrecollecting who said what. Experiment 2 uses the same procedure and materials as Experiment 1 and asks how improving source memory affects the occurrence of misrecollections in older adults. This experiment compared younger adults with an immediate test with older adults who also received an immediate test. However, the older adults encountered the study items multiple times so as to improve and equate their overall source memory with that of younger adults.

Table 4

Analysis of the Receiver Operating Characteristics (ROCs) of the Young and Older Adults' Data

Group	R^2_{line}	F_{line}	MSE_{line}	p_{line}	F_{quad}	MSE_{quad}	C_{quad}	p_{quad}
z-ROC linearity analysis results								
Experiment 1								
Young–delay	.991	1,008.94	1.3×10^{-2}	.000	154.19	6.2×10^{-4}	0.09	.000
Older–1Rep	.992	1,132.17	6.9×10^{-3}	.000	53.88	1.0×10^{-3}	−0.10	.000
Experiment 2								
Young–1Rep	.979	428.87	2.7×10^{-2}	.000	1.63	2.5×10^{-2}	0.05	.238
Older–2Rep	.964	241.81	1.7×10^{-2}	.000	143.6	9.9×10^{-4}	−0.27	.000
Older–3Rep	.993	1,302.07	3.9×10^{-3}	.000	1.59	3.7×10^{-3}	0.04	.243
ROC linearity analysis results								
Experiment 1								
Young–delay	.933	125.9	7.2×10^{-3}	.000	196.65	3.2×10^{-4}	−1.03	.000
Older–1Rep	.913	94.55	7.7×10^{-3}	.000	118.09	5.5×10^{-4}	−1.30	.000
Experiment 2								
Young–1Rep	.678	18.96	2.7×10^{-2}	.002	25.56	7.3×10^{-3}	−2.15	.001
Older–2Rep	.775	30.96	1.1×10^{-2}	.000	39.92	2.1×10^{-3}	−2.51	.000
Older–3Rep	.756	27.94	9.9×10^{-3}	.001	34.24	2.1×10^{-3}	−2.32	.000

Method

Participants. Eighteen undergraduates (range of 18–23 years) participated in the young–1Rep condition, and 36 older adults (range of 60–80 years) participated in the older–2Rep and older–3Rep conditions (i.e., 18 in each group). Table 1 contains demographic information for all participants. One individual's demographic and psychometric data from the older–3Rep group were lost owing to human error. Both groups of older adults performed worse than the younger adults on the word-list recall test (both ts > 2.62, ps $< .05$), as well as on the logical memory test: old–3Rep versus young, $t(33) = 2.05$, $p < .05$; old–2Rep versus young, $t(34) = 1.13$. Finally, neither group of older adults performed significantly differently from the younger adults on the vocabulary test (both ts < 1.19). As in Experiment 1, these findings are consistent with the usual age-related patterns of performance on these psychometric tests.

Materials and procedure. The materials were identical to those used in Experiment 1. Individuals in the young–1Rep condition encountered each study statement once, whereas individuals in the older–2Rep and older–3Rep conditions encountered each study statement two and three times, respectively. For these latter two groups, at least three trials separated repetitions of study statements. After the encoding phase, all groups completed a 5-min distractor task before proceeding to the test phase. In all other respects, the procedure was identical to the one we used in Experiment 1.

Results

Matching source identification performance. Table 2 displays overall memory performance. There were no significant differences between the groups in conditional source scores, all ts$(34) < 1.08$. With respect to item memory, false alarm rates were lower for the older–3Rep individuals than for the young–1Rep or the older–2Rep individuals, both ts$(34) > 1.98$, $p < .055$; the latter two groups were no different. However, each group's hit rates and recognition d' scores were significantly different from each of the others, all ts$(34) > 3.18$, ps $< .01$.[3] As in Experiment 1, these recognition memory differences are an expected consequence of our procedures for equating source memory performance between younger and older adults.

Testing signal detection models of source memory. We next fit our misrecollection model and the standard signal detection model to the data from the young–1Rep, older–2Rep, and older–3Rep groups. Table 5 shows the corresponding parameter values associated with the IR and standard UEV models. Consistent with Experiment 1, the misrecollection model fit the young–1Rep group's data no better than the standard signal detection model, $\chi^2(1) < 1.22$, $p > .20$, although neither model adequately fit the data: UEV, $\chi^2(9) = 69.81$; IR, $\chi^2(8) = 68.59$, ps $< .001$. This lack of fit is not unusual for the standard signal detection model and has been observed on occasion in other articles (e.g., Hilford et al., 2002; Slotnick & Dodson, 2005; Slotnick et al., 2000). The misrecollection parameter (ζ) was not significantly different from zero for the young–1Rep group's data, $t(8) < 1$, indicating that the misrecollection model is formally equivalent to the standard signal detection model and also indicating that the younger adults experienced few misrecollections. For the older–2Rep data, the IR model provided a significantly better fit of the data than did the UEV model, $\chi^2(1) = 32.31$, $p < .001$. In fact, only the IR model provided an adequate fit of the older–2Rep data: UEV, $\chi^2(9) = 38.75$, $p < .001$; IR, $\chi^2(8) = 6.44$, $p > .20$. In addition, the value of .12 for ζ indicates that 12% of the older–2Rep group's source identification responses were based on illusory recollections—a value that is significantly different from zero, $t(8) = 3.60$, $p < .01$. Of interest, for the older–3Rep group, the UEV model provided an adequate fit to the data, $\chi^2(9) = 8.04$, and the IR model provided a fit that was no better than the standard UEV model, $\chi^2(1) < 1$. Moreover, the misrecollection parameter for this group was not significantly different from zero, $t(8) < 1$, which indicates that for older adults, misrecollections decrease with improvements in source memory. Finally, as in Experiment 1, the two-high-threshold model did not provide an adequate fit of any group's data (all χ^2 values > 403, ps $< .001$).

---

[3] As in Experiment 1, d' scores were computed from transformed hit rates and false alarm rates.

Table 5
Parameter Values From the Standard Unequal Variance (UEV) and the Illusory Recollection (IR) Signal Detection Models for the Young and Older Adults' Data in Experiment 2

Condition	Source d'	σ_{s1}/σ_{s2}	Misrec
Young–1Rep			
Standard UEV model	1.61 (0.12)	0.99 (0.06)	
IR model	1.62 (0.19)	1.01 (0.12)	0 (.06)
Older–2Rep			
Standard UEV model	1.34 (0.08)	1.00 (0.06)	
IR model	2.22 (0.11)	0.89 (0.07)	.12 (.03)
Older–3Rep			
Standard UEV model	1.73 (0.05)	0.96 (0.03)	
IR model	1.73 (0.05)	0.96 (0.03)	0 (.07)

Note. Parentheses contain the standard error of the mean. Misrec = Misrecollection parameter, which refers to the proportion of items that are misremembered as originating from the incorrect source.

An individual-participant chi-square analysis produced the same pattern of results. Three participants were excluded from the older–3Rep condition owing to source d' values of less than two standard errors from zero (no participants were excluded from the other two conditions). As in the group analysis, the misrecollection model fit the young–1Rep group's data no better than the standard signal detection model, $\chi^2(18) < 18$, $p > .20$, and neither model adequately fit the data: UEV, $\chi^2(162) = 344.88$; IR, $\chi^2(144) = 351.64$, $ps < .001$. For the older–2Rep data, the IR model provided a numerically better fit of the data than did the UEV model, $\chi^2(18) = 14.61$, with both models providing an adequate fit of the data: UEV, $\chi^2(162) = 185.42$, $p = .10$; IR, $\chi^2(144) = 170.81$, $p = .063$. For the older–3Rep group, neither model provided an adequate fit to the data: UEV, $\chi^2(135) = 216.71$, $p < .001$; IR, $\chi^2(120) = 290.91$, $p < .001$; and as in the group analysis, the IR model provided a fit that was no better than the standard UEV model, $\chi^2(15) < 15$. The two-high-threshold model did not provide an adequate fit of any data (all χ^2 values > 1,853, $ps < .001$).

Examining the shape of the source z-ROC. Lastly, the IR model predicts that misrecollections will cause a negative curvature of the slope of the *z*-ROC for the source data. Figure 4 shows the *z*-ROCs for the source data from the young–1Rep, older–2Rep, and older–3Rep conditions. The notable pattern in this figure is the negative curvature of the *z*-ROC for the older–2Rep group (Figure 4, middle panel) but the straight *z*-ROCs for the young–1Rep group (Figure 4, top panel) and the older–3Rep group (Figure 4, bottom panel). Table 4 displays the linearity analysis of the *z*-ROC and ROC. Younger adults' data showed a slight positive curvature ($c_{quad} = .05$) of the *z*-ROC that was not significantly different from zero ($p > .20$). By contrast, the older–2Reps' data showed a significantly negative curvature ($c_{quad} = -.27$, $p < .001$) of the *z*-ROC, which is consistent with the results from Experiment 1. For the older–3Rep group, however, there was a slight positive curvature ($c_{quad} = .04$) of the *z*-ROC that was not different from zero ($p > .20$). Finally, the bottom half of Table 4 presents the linearity analysis of the ROCs. As in Experiment 1, this analysis indicated that all groups showed significantly curvilinear source ROCs ($ps < .001$).

Overall, Experiments 1 and 2 indicate that regardless of whether young adults show good source memory or poor source memory

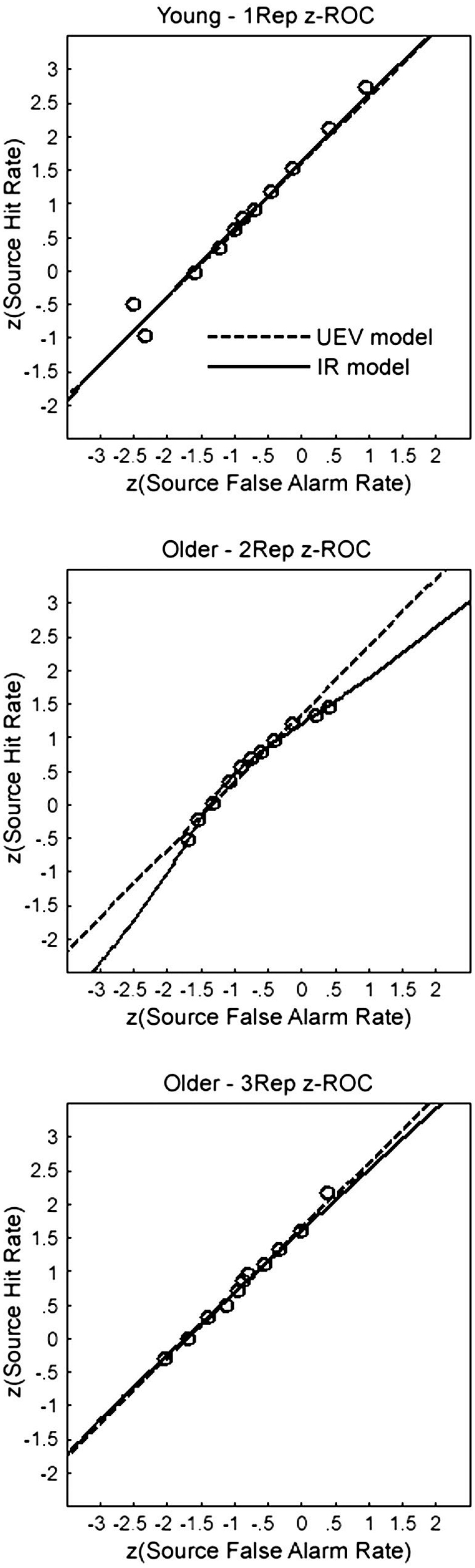

Figure 4. Source memory *z*-ROC with best fit unequal variance (UEV) model and illusory recollection (IR) model for the young–1Rep group (top panel), the older–2Rep group (middle panel), and the older–3 Rep group (bottom panel) from Experiment 2. *z*-ROC = *z*-transformed receiver operating characteristics.

for past events, they experience very few misrecollections, at least as we measured them in this paradigm. By contrast, older adults were prone to misrecollecting past events. Of interest, this vulnerability steadily diminished with improvements in source memory. Older adults who encountered each statement once were more likely to misrecollect past events than were older adults who encountered each statement twice or thrice. In fact, older adults who studied each statement three times showed very few misrecollections as measured by our IR model. Thus, for older adults, well-learned events appear less vulnerable to being misremembered.

General Discussion

We have sought to answer the question of why source memory, such as remembering who said what, worsens with old age. Several findings indicate that older adults' increased susceptibility to illusory recollections primarily caused age-related differences in source memory (i.e., remembering who said what). Both experiments show that even when younger and older adults are matched on their overall accuracy at remembering source information, older adults are nevertheless prone to misrecollections, as illustrated by the negative curvature of their *z*-ROCs in Figure 3 (bottom panel) and Figure 4 (middle panel). We propose a new illusory recollection signal detection model that is better able to fit older adults' data than is the standard unequal-variance signal detection model. For younger adults the fits of both models did not differ significantly. Of importance, these results indicate that illusory recollections in this paradigm are a consequence of cognitive aging, as they are not observed in younger adults whose source memory is just as poor as older adults.

As we foreshadowed in the introduction, one of the exciting findings from these experiments is that age-related differences in source accuracy may be substantially if not completely caused by the occurrence of misrecollections. This result is visible by comparing younger and older adults who experienced identical study and test conditions, such as the young–1Rep and older–1Rep groups. An examination of their conditional source scores in Table 2 confirms that we replicated the well-established finding that older adults (.66) show worse source memory than younger adults (.79) when both have encountered the identical study and test conditions, $t(34) = 4.46$, $p < .001$. This difference in performance is also clear by examining the source d' scores from the standard signal detection model for older–1Rep adults (UEV $d' = 0.70$) and young–1Rep adults (UEV $d' = 1.61$). However, worse performance on the part of older adults could be caused either by misrecollections, by guessing due to an inability to remember diagnostic source information, or by some combination of these two factors.

To answer the question of why older–1Rep adults show worse source memory than young–1Rep adults, it is necessary to examine the source d' and misrecollection parameters from our IR model, shown in Tables 3 and 5. Strikingly, when we factor misrecollections out of performance, the young–1Rep (IR source $d' = 1.62$) and older–1Rep (IR source $d' = 1.50$) adults show no significant differences in their source d' scores, $t(34) < 1$. Older–1Rep adults, however, have a much higher misrecollection score (.23) than do young–1Rep adults (0). These results suggest, then, that in this paradigm the increased occurrence of misrecollections on the part of older adults primarily cause age-related differences in remembering who said what. Once we account for these misrecollections, older and younger adults show no significant differences in source accuracy.

Some prevailing theories of source monitoring explain source errors in terms of response biases or guessing strategies, but not in terms of illusory memories (e.g., Batchelder & Riefer, 1990, 1999).[4] Yet the assumption that illusory memories contribute to source errors is a key aspect of Johnson and colleagues' source monitoring framework (Johnson et al., 1993). This framework characterizes source misattributions as occurring, in part, when a memory contains features that are typical of memories from another source, such as when a memory for an imagined event elicits a sufficient amount of visual detail that individuals mistakenly judge it as having been seen. There are other prominent memory theories, such as fuzzy trace theory, that also assume the occurrence of illusory recollections (e.g., Brainerd, Payne, Wright, & Reyna, 2003). Although we have used our IR model to examine age-related differences in source memory, the potential applications of this model are much broader and include investigating any factor that impairs source identification performance, such as divided attention or similarity between the sources.

The illusory recollection and reduced memory accounts offer different, but not exclusive, explanations of the effects of cognitive aging on episodic memory. Nor are they the sole explanations, as age-related memory deficits can also involve changes in the way memories are evaluated (e.g., Multhaup, 1995). Although we did not observe age-related evidence for the reduced memory account in our particular source memory paradigm, from our reading of the episodic memory literature there is support for both the illusory recollection and reduced memory accounts. As stated earlier, the reduced memory account argues for an increase in guessing because of the inaccessibility/unavailability of diagnostic information. Consistent with the reduced recollection account, on typical old–new recognition tests older adults are less likely to justify recognition judgments with "remember" responses than are younger adults (e.g., Bunce & Macready, 2005; Parkin & Walter, 1992). By contrast, studies that have used a variety of false memory paradigms show that there are age-related increases in "remember" responses associated with false memories (e.g., Norman & Schacter, 1997; Schacter et al., 1997). For instance, Schacter et al. (1997) observed that when individuals incorrectly responded that an event had been seen earlier in a video when in fact it actually appeared in a photograph, older adults were much more likely than younger adults to justify these source errors by claiming to recollect specific details about the event. Similarly, Norman and Schacter (1997) observed that older adults were more likely to report that they recollected—as indexed by both "remember" responses and memory characteristics questionnaire ratings—specific details about lure items that were semantically related to what was actually studied in the Deese–Roediger–McDermott false memory paradigm (Deese, 1959; Roediger & McDermott, 1995). It may be that older adults' reduced ability to recollect detailed information about past events is indeed partly responsible for the age-related increase in false recollections in these foregoing

---

[4] We thank David Riefer for this and some of the other points in this paragraph.

paradigms, particularly because older adults remembered fewer sensory details than younger adults about what was studied in the initial videotape or in the initial list of words (Schacter et al., 1998). Alternatively, our misrecollection account provides a new explanation for both older adults' worse source memory performance and their increased susceptibility to memory distortions in false memory paradigms: Older adults have higher rates of illusory recollection.

A variety of different binding accounts have been proposed to explain both age-related memory deficits and the occurrence of false recollections (e.g., Chalfonte & Johnson, 1996; Henkel et al., 1998; Koutstaal et al., 2001; Kroll et al., 1996; Light, Patterson, Chung, & Healy, 2004; Lyle, Bloise, & Johnson, 2006; Mitchell, Johnson, Raye, & D'Esposito, 2000; Naveh-Benjamin, Hussain, Guez, & Bar-On, 2003; Schacter et al., 1998). For instance, Naveh-Benjamin and colleagues have proposed that older adults' episodic memory impairments are due, in part, to a reduced ability to encode and/or retrieve associations between units of information, such as associations between different items or between an item and a speaker (Naveh-Benjamin et al., 2003). This kind of binding deficit corresponds to what we refer to as the reduced recollection account, because it explains memory failures in terms of a reduced ability to retrieve the different features of an episode. This account does not appear to explain the occurrence of misrecollections on the part of older adults.

However, there are a variety of binding accounts that can explain the occurrence of misrecollections. For example, Kroll et al. (1996) suggested that during the encoding stage, features from one event may become miscombined with features from other events. Kroll et al. observed that patients with hippocampal damage were especially vulnerable to falsely recognize conjunction lures (e.g., test items, such as *barley*, that consist of syllabic features from different studied items, such as *barter* and *valley*), particularly when the conjunction lures were based on items that occurred close together during the study phase. They proposed that hippocampal damage causes excessive or disinhibited binding so that features from events close in temporal proximity have a high probability of being miscombined.

Alternatively, Henkel et al. (1998) have proposed a binding account to explain the occurrence of false recollections in their paradigm (see also Lyle et al., 2006). Henkel et al. observed that older adults were especially likely to misremember that an imagined item, such as a lollipop or banana, had been seen when these items were either physically or conceptually similar to other items that were actually seen, such as a magnifying glass or an apple (Lyle, Bloise, & Johnson, 2006; Lyle & Johnson, 2006). To explain these kinds of errors, they suggested that during retrieval, memory for features from other items may become activated because of their similarity to the tested item. Lyle and Johnson (2006) use the term *importation* to refer to this process of inadvertent, similarity-based activation and misattribution of features at retrieval (see similar ideas in Koutstaal et al.'s [2001] notion of *misplaced recollection* and Lampinen and colleagues' view of *content borrowing*; e.g., Lampinen, Meier, Arnal, & Leding, 2005). The likelihood of importing features at retrieval depends on binding processes at encoding. For example, the overall amount of features that are bound together is critical because, as Lyle et al. (2006) suggested, if older adults have degraded memories and fail to encode particular features, then they will have fewer opportunities to inadvertently import features, simply because there are fewer features available to possibly be imported (see also their related view about the effects of features that are not tightly bound together).

We suggest two different mechanisms that contribute to the occurrence of misrecollections, and we draw on the ideas of Kroll et al. (1996) and the source monitoring framework of Johnson and colleagues (Henkel et al., 1998; Lyle & Johnson, 2006; Lyle et al., 2006) and the constructive memory framework of Schacter and colleagues (Schacter, Norman, & Koutstaal, 1998). First, because the hippocampus shows substantial and accelerating age-related shrinkage (e.g., Raz et al., 2005), following Kroll et al. (1996) we suggest that older adults' misrecollections are due, in part, to disinhibited binding processes that operate across an expanded temporal window. That is, aging is associated with an increased probability that features from different events become bound together so that they form an engram that contributes to phenomenal experiences that are indistinguishable from compelling "true" memories. Thus, we predict that older adults are prone to miscombining features from studied events that occur in close temporal proximity to each other. It is important to note that these temporal-proximity binding errors that occur during the encoding phase will not necessarily be similarity based.

By contrast, building on the ideas within the source monitoring and constructive memory frameworks, the second mechanism that contributes to misrecollections involves pattern completion processes (McClelland, McNaughton, & O'Reilly, 1995), whereby retrieval cues activate features of the desired memory and sometimes features of nondesired memories. Here, similarity-based processes will contribute to misrecollections by activating features from nondesired but similar memories (Henkel et al., 1998). In older adults, however, these misrecollections will tend to be associated with high confidence.

Finally, the illusory recollection account predicts that illusory recollections drive age-related memory impairments on *any* task that requires remembering specific information as a basis for a response, such as on tests of word-list recall, cued recall, and prospective memory—in other words, tasks that individuals cannot answer on the basis of familiarity. In support of this prediction, we have observed in an eyewitness suggestibility paradigm that when younger and older adults were matched on their overall memory for experienced events, older adults most often made suggestibility errors (i.e., claiming to have seen objects that were actually only suggested) when they were certain about the correctness of their response. By contrast, their younger, accuracy-matched counterparts most often made these errors when they were uncertain about the correctness of their response (Dodson & Krueger, in press). We have observed similar age-related increases in the occurrence of high-confidence errors and an associated lack of calibration on cued-recall tests, even when younger and older adults are matched on their overall levels of cued-recall accuracy (Dodson, Bawa, & Krueger, in press). Moreover, in support of our account, we have observed no age-related differences in the occurrence of high-confidence errors (i.e., misrecollections) on primarily familiarity-based tasks, such as old–new recognition.

In conclusion, we have proposed an illusory recollection account of why cognitive aging is associated with episodic memory deficits. We have shown that older adults are prone to misrecollecting past events. We have instantiated this account in a new illusory recollection signal detection model that is better able to fit older adults' data than is the standard signal detection model. Moreover, we observed that age-related memory impairments virtually disappear once these illusory recollections are accounted for. Finally, then, it appears that Maurice Chevalier's elderly character in the film *Gigi* may be the appropriate model for aging memory when he confidently misremembers the details of a love affair, in a song that perhaps should have been titled "I Misremember It Well."

References

Anderson, N. D., & Craik, F. I. M. (2000). Memory in the aging brain. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 411–425). Oxford, England: Oxford University Press.

Balota, D. A., Dolan, P. O., & Duchek, J. M. (2000). Memory changes in healthy older adults. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 395–410). Oxford, England: Oxford University Press.

Banks, W. P. (2000). Recognition and source memory as multivariate decision processes. *Psychological Science, 11,* 267–273.

Batchelder, W. H., & Riefer, D. M. (1990). Multinomial processing models of source monitoring. *Psychological Review, 97,* 548–564.

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin & Review, 6,* 57–86.

Brainerd, C. J., Payne, D. G., Wright, R., & Reyna, V. F. (2003). Phantom recall. *Journal of Memory and Language, 48,* 445–467.

Bunce, D., & Macready, A. (2005). Processing speed, executive function, and age differences in remembering and knowing. *Quarterly Journal of Experimental Psychology, 58,* 155–168.

Castel, A. D., & Craik, F. I. M. (2003). The effects of aging and divided attention on memory for item and associative information. *Psychology and Aging, 18,* 873–885.

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

DeCarlo, L. T. (2002). Signal detection theory with finite mixture distributions: Developments with applications to recognition memory. *Psychological Review, 109,* 710–721.

DeCarlo, L. T. (2003). Source monitoring and multivariate signal detection theory, with a model for selection. *Journal of Mathematical Psychology, 47,* 292–303.

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

Dodson, C. S., Bawa, S., & Krueger, L. E. (in press). Aging, metamemory, and high confidence errors: A misrecollection account. *Psychology and Aging.*

Dodson, C. S., Koutstaal, W., & Schacter, D. L. (2000). Escape from illusion: Reducing false memories. *Trends in Cognitive Science, 4,* 391–397.

Dodson, C. S., & Krueger, L. E. (in press). I misremember it well: Why older adults are unreliable eyewitnesses. *Psychonomic Bulletin & Review.*

Dodson, C. S., & Schacter, D. L. (2002). Aging and strategic retrieval processes: Reducing false memories with a distinctiveness heuristic. *Psychology & Aging, 17,* 405–415.

Gardiner, J. M., & Richardson-Klavehn, A. (2000). Remembering and knowing. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 229–244). Oxford, England: Oxford University Press.

Glanzer, M., Hilford, A., & Kim, K. (2004). Six regularities of source recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 1176–1195.

Glisky, E. L., Rubin, S. R., & Davidson, P. S. R. (2001). Source memory in older adults: An encoding or retrieval problem? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 1131–1146.

Hashtroudi, S., Johnson, M. K., & Chrosniak, L. D. (1989). Aging and source monitoring. *Psychology and Aging, 4,* 106–112.

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

Hilford, A., Glanzer, M., Kim, K., & DeCarlo, L. T. (2002). Regularities of source recognition: ROC analysis. *Journal of Experimental Psychology: General, 131,* 494–510.

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

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

Kausler, D. H., & Puckett, J. M. (1980). Adult age differences in recognition memory for a nonsemantic attribute. *Experimental Aging Research, 6,* 349–355.

Koutstaal, W., Schacter, D. L., & Brenner, C. (2001). Dual task demands and gist-based false recognition of pictures in younger and older adults. *Journal of Memory and Language, 44,* 399–426.

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

Lampinen, J. M., Meier, C. R., Arnal, J. D., & Leding, J. K. (2005). Compelling untruths: Content borrowing and vivid false memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 954–963.

Light, L. L., Patterson, M. M., Chung, C., & Healy, M. R. (2004). Effects of repetition and response deadline on associative recognition in young and older adults. *Memory & Cognition, 32,* 1182–1193.

Lyle, K. B., Bloise, S. M., & Johnson, M. K. (2006). Age-related binding deficits and the content of false memories. *Psychology and Aging, 21,* 86–95.

Lyle, K. B., & Johnson, M. K. (2006). Importing perceived features into false memories. *Memory, 14,* 197–213.

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

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review, 102,* 419–457.

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

Memon, A., Bartlett, J., Rose, R., & Gray, C. (2003). The aging eyewitness: Effects of age on face, delay and source-memory. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 6,* 338–345.

Mitchell, K. J., Johnson, M. K., Raye, C. L., & D'Esposito, M. D. (2000). fMRI evidence of age-related hippocampal dysfunction in feature binding in working memory. *Cognitive Brain Research, 10,* 197–206.

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

Naveh-Benjamin, M., Hussain, Z., Guez, J., & Bar-On, M. (2003). Adult age differences in episodic memory: Further support for an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 826–837.

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

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

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

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (1992). *Numerical recipes in C* (2nd ed.). New York: Cambridge University Press.

Qin, J., Raye, C. L., Johnson, M. K., & Mitchell, K. J. (2001). Source memory ROCs are (typically) curvilinear: Comment on Yonelinas (1999). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 1110–1115.

Raz, N., Lindenberger, U., Rodrigue, K. M., Kennedy, K. M., Head, D., Williamson, A., et al. (2005). Regional brain changes in aging healthy adults: General trends, individual differences and modifiers. *Cerebral Cortex, 15,* 1676–1689.

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,* 803–814.

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

Schacter, D. L., Koutstaal, W., Johnson, M. K., Gross, M. S., & Angell, K. A. (1997). False recollection induced by photographs: A comparison of older and younger adults. *Psychology and Aging, 12,* 203–215.

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

Siedlecki, K. L., Salthouse, T. A., & Berish, D. E. (2005). Is there anything special about the aging of source memory? *Psychology and Aging, 20,* 19–32.

Simons, J. S., Dodson, C. S., Bell, D., & Schacter, D. L. (2004). Specific and partial source memory: Effects of aging. *Psychology and Aging, 19,* 689–694.

Slotnick, S. D., & Dodson, C. S. (2005). Support for a continuous (single-process) model of recognition memory and source memory. *Memory & Cognition, 33,* 151–170.

Slotnick, S. D., Klein, S. A., Dodson, C. S., & Shimamura, A. P. (2000). An analysis of signal detection and threshold models of source memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1499–1517.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117,* 34–50.

Wechsler, D. (1997a). *Wechsler Adult Intelligence Scale—Third edition*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (1997b). *Wechsler Memory Scale—Third edition*. San Antonio, TX: Psychological Corporation.

Appendix

Confidence Rating Data for Source Judgments

Experiment 1

Young–delay

	← “Female”						“Male” →					
	100	90	80	70	60	50	50	60	70	80	90	100
Male items	4	17	42	48	67	38	42	69	85	79	47	58
Female items	56	62	78	85	75	32	33	64	47	41	16	9

Older–1Rep

	← “Female”						“Male” →					
	100	90	80	70	60	50	50	60	70	80	90	100
Male items	30	26	30	42	32	48	77	42	76	82	73	74
Female items	95	59	73	79	48	67	52	28	48	41	28	25

Experiment 2

Young–1Rep

	← “Female”						“Male” →					
	100	90	80	70	60	50	50	60	70	80	90	100
Male items	2	9	30	35	37	25	35	60	90	111	92	104
Female items	105	109	131	82	51	30	22	30	36	31	4	0

Older–2Rep

	← “Female”						“Male” →					
	100	90	80	70	60	50	50	60	70	80	90	100
Male items	50	13	16	41	29	21	29	59	83	68	78	209
Female items	230	55	90	77	46	29	30	31	32	23	11	31

Older–3Rep

	← “Female”						“Male” →					
	100	90	80	70	60	50	50	60	70	80	90	100
Male items	11	28	26	33	22	18	35	49	50	90	81	271
Female items	250	105	93	60	52	18	10	32	35	26	17	15

Note. These matrices contain the overall frequency of *female* and *male* responses and the associated confidence rating on a scale from 100 (*certain correct*) to 50 (*guessing*) for each group of participants in each of the experiments.

Received March 14, 2006
Revision received July 5, 2006
Accepted July 6, 2006 ■