Psychology and Aging
2009, Vol. 24, No. 2, 494–500

0882-7974/09/$12.00 DOI: 10.1037/a0015269

Aging and Recollection in the Accuracy of Judgments of Learning

Karen A. Daniels and Jeffrey P. Toth
University of North Carolina, Wilmington

Christopher Hertzog
Georgia Institute of Technology

Dual-process theories propose that episodic memory performance reflects both recollection of prior details as well as more automatic influences of the past. The authors explored the idea that recollection mediates the accuracy of judgments of learning (JOLs) and may also help explain age differences in JOL accuracy. Young and older adults made immediate JOLs at study and then completed recognition or recall tests that included a recollect/familiar judgment. JOLs were found to be strongly related to recollected items but not to items remembered on the basis of familiarity. The pattern was weaker in older adults, consistent with age-related declines in recollection.

Keywords: judgments of learning, metacognition, recollection, recognition, cued recall

A fundamental issue in the study of metacognition is what drives metacognitive accuracy. One area in which this issue has received considerable attention is *judgments of learning* (JOLs) in which a person rates how well they will remember something in the future (Arbuckle & Cuddy, 1969). As noted by a number of researchers (e.g., Hertzog, Kidder, Powell-Moman, & Dunlosky, 2002; Kelemen & Weaver, 1997), the relative accuracy of immediate JOLs—that is, the ability of people to predict which items they will remember—tends to be comparatively modest. Moreover, a number of studies have shown little or no age-related declines in relative JOL accuracy (e.g., Connor, Dunlosky, & Hertzog, 1997; Hertzog et al., 2002; Rabinowitz, Craik, Ackerman, & Hinchley, 1982; but see Shaw & Craik, 1989). One way to interpret this pattern is to assume that JOLs are primarily based on automatic, age-invariant factors such as processing fluency (Begg, Duft, Lalonde, & Melnick, 1989; Benjamin, Bjork, & Schwartz, 1998). Indeed, effects of processing fluency, coupled with demonstrations that JOLs are often insensitive to powerful extrinsic influences on memory (e.g., retention interval), have led some to suggest that "online JOLs are based predominantly—perhaps exclusively—on the subjective experience associated with processing fluency" (Koriat, Bjork, Sheffer, & Bar, 2004, p. 653).

While not disputing the importance of indirect cues like fluency in affecting JOLs and JOL accuracy, our goal was to explore another mnemonic cue—contextual detail—that may offer a more reliable basis for making JOLs and mediating JOL accuracy. In doing so, we hoped to raise the issue of whether privileged access to contextual, recollective details—as conceptualized in dual-process theories of memory—should play a larger role in current theories of metamemory.

Despite findings of age invariance in relative JOL accuracy, a number of studies have suggested that age-related declines may be observable under some conditions. For example, older adults consistently show deficits in remember/know (e.g., Perfect & Dasgupta, 1997) and source memory tasks (e.g., McIntyre & Craik, 1987), both of which arguably tap memory-monitoring mechanisms. In addition, Kelley and Sahakyan (2003) found older adults to have lower correspondence between memory accuracy and confidence, a pattern mimicked by young adults whose attention had been divided at encoding. These results suggest that recollection of contextual details may play an important role in a variety of memory-monitoring situations (see also Carroll & Shanahan, 1997; Gallo, Cotel, Moore, & Schacter, 2007; Hicks & Marsh, 2002).

Dual-process theories of memory distinguish between recollection and more automatic influences of memory such as familiarity (Jacoby, 1991; Mandler, 1980; Yonelinas, 2002). While familiarity is viewed as a relatively undifferentiated feeling of "oldness" (Whittlesea, 1993), recollection involves the conscious retrieval of contextual details associated with a prior event. Such details may include perceptual features (e.g., color, location), sources of information (e.g., who said something), and the private thoughts, images, or feelings experienced during the initial event. Recollection is thus a central component of many high-level forms of memory, including source memory (Johnson, Hashtroudi, & Lindsay, 1993), explicit memory (Richardson-Klavehn & Bjork, 1988), episodic memory (Tulving, 1985), and autobiographical memory (Rubin, 2005). In the extreme, episodic recollection can be viewed as a form of "mental time travel," whereby a person consciously relives past events (Wheeler, Stuss, & Tulving, 1997).

The goal of the present study was to examine the role of recollection in mediating JOL accuracy. In particular, we explored the idea that JOL accuracy is mediated by the recollection at test of contextual, item-specific details that were generated and used as cues to drive JOLs at study. To examine this idea, we had young and older adults make immediate, item-by-item JOLs to a list of study words. They then took a test of either recognition memory or cued recall. For recognition, intact test words were classified as

Karen A. Daniels and Jeffrey P. Toth, Department of Psychology, University of North Carolina, Wilmington; Christopher Hertzog, Department of Psychology, Georgia Institute of Technology.

Portions of this research were supported by a grant from the Georgia Gerontology Consortium to Karen A. Daniels. We thank Colleen Kelley for constructive comments on an earlier version of this article and Amanda Lash for assistance in preparing the manuscript.

Correspondence concerning this article should be directed to Karen A. Daniels, Department of Psychology, University of North Carolina, 601 South College Road, Wilmington, NC 28403. E-mail: danielsk@uncw.edu

recollect (*R*), *familiar* (*F*), or *no memory* (*N*). For cued recall, participants completed word stems, classifying their completions as *R*, *F*, or *N*. On the basis of research showing that recollection involves the retrieval of contextual details generated at encoding (Gardiner, Ramponi, & Richardson-Klavehn, 1998), we predicted that JOLs for items subsequently classified as *R* would be significantly higher than those rated *F* or *N*. We expected both age groups to show this effect. However, given age-related declines in recollection (e.g., Jacoby, 1999), we predicted a reduced effect for older adults. A final prediction concerned recognition versus cued recall. Although recall may require greater effort than recognition (Craik, 1986), research has shown equivalent levels of recollection in matched versions of these tests (Hamilton & Rajaram, 2003; Toth, 2005). We therefore predicted that the increased JOL accuracy for R items, compared with F and N items, would be roughly equivalent in the recognition and cued recall tests.

Method

Participants

Seventy-six young adults (42 men, 34 women; mean age = 19.6 years) and 78 older adults (25 men, 53 women; mean age = 71.4 years) participated.[1] Young adults were undergraduates at the Georgia Institute of Technology and participated for course credit. Older adults were from the Atlanta community and were paid $10 per hour. Participants from the two age groups were randomly assigned to the recognition and cued-recall test conditions.

Materials

Stimuli consisted of 120 five-letter nouns and adjectives of moderate frequency (e.g., *glass*, *clown*) that were divided into three sets of 40. Each participant studied words from two of the sets, with the remaining set acting as unstudied items during test. Word sets were counterbalanced across participants such that each word served equally often as studied and unstudied. The same words and word sets were used in the recognition and cued-recall tests. For cued recall, we created word-stem cues by replacing the final two letters from each word with underlines. Each stem could be completed by at least two words but only one word from the stimulus set. Testing was done on Windows-based (Microsoft Corp., Redmond, WA) PCs with 11 keys across the top of the keyboard (from the symbol ~ to the number *0*) relabeled *0* to *100* in intervals of 10.

Procedure

At study, participants were told they would see a list of words that they should remember for a later memory test and that they "may want to associate each word with something that is personally meaningful" or "generate a mental image." They were also asked to rate how likely they were to remember each word on the later memory test. Words were presented one at a time in the center of the screen with the JOL scale (0 . . . 100) appearing at the bottom of the screen 3 s after the word was presented. Participants were told a rating of 0 meant they were "absolutely certain" they would not remember the word, a rating of 100 meant they were "absolutely certain" they would remember the word, and ratings from 10 to 90 indicated intermediate levels of certainty. Participants were encouraged to use the entire scale. Words remained visible until the JOL rating was entered. One second after the rating, the next word was presented.

Following study, participants completed two distractor tasks, resulting in a study-test delay of 10–12 min. For recognition, participants saw words one at a time in the center of the screen and were asked to rate the quality of their memory for each using as ratings the words *recollect*, *familiar*, or *no memory*. Definitions of these responses were similar to those used in remember/know studies (e.g., Gardiner, 1988). Briefly, participants were told to respond *recollect* when they could clearly remember specific details associated with studying a word; *familiar* when they felt the word was from the study list but could not remember specific details; and *no memory* when the word neither felt familiar nor triggered any memory of details about its earlier presentation. The response options were presented at the bottom of the screen as "Recollect (R), Familiar (F), No Memory (N)." Participants responded using the *R*, *F*, and *N* keys.

For cued recall, participants were presented with word stems and told to complete each with a word from the study list. If they could not remember a study word, they were to complete the stem with the first word that came to mind. Immediately after a word was entered, the R/F/N response options, defined as in the recognition condition, appeared at the bottom of the screen.

Results

We begin by comparing memory performance in young and older adults. We then analyze JOL ratings and their relation to observed memory performance (i.e., JOL accuracy).

Memory Performance

Table 1 contains the proportion of R, F, and N responses given by young and older adults on the recognition test. Table 2 contains the corresponding data for cued recall, presented as both observed and rescaled proportions. For the observed proportions (top panel), the R, F, and N responses are expressed as a proportion of the total number of target words (80 studied, 40 unstudied); thus, the observed R, F, and N proportions sum to the total proportion of stems completed with target words. For the rescaled proportions (bottom panel), the R, F, and N responses are expressed as a proportion of the number of test stems completed with target words; thus, the adjusted R, F, and N proportions sum to 1.0, as in recognition. Presenting the cued-recall data as adjusted proportions allows a better comparison across conditions (tests or age groups) that differ in the total number of target words for which a participant can make an R/F/N judgment (see Hamilton & Rajaram, 2003; Toth, 2005).

[1] Data from 4 additional young participants were excluded, 1 due to ceiling performance (one error response), 1 due to computer malfunction, 1 due to experimenter error, and 1 for failing to follow instructions. Data from 10 additional older participants were excluded, 3 due to preexisting health conditions (e.g., recent stroke), 3 due to abnormal (negative memory) performance, 2 due to experimenter error, 1 for failing to follow instructions, and 1 for falling outside the acceptable age range (age 47 years).

Table 1
Mean Observed Proportions of Recollect (R), Familiar (F), and No Memory (N) Responses by Young and Older Participants On The Recognition Memory Test

	Studied items				Unstudied items			
Group	R + F	R	F	N	R + F	R	F	N
Young								
M	.91	.73	.18	.09	.14	.02	.12	.86
SD	.09	.16	.11	.09	.10	.02	.10	.10
Older								
M	.83	.68	.15	.17	.22	.08	.15	.78
SD	.14	.17	.10	.14	.14	.07	.10	.14

For recognition, older adults were less likely than the young to correctly classify studied words as "old" (R + F: .83 vs. .91; $t(75) = 2.85$, $p < .01$) but more likely to incorrectly classify unstudied words as old (.22 vs. .14; $t(75) = 2.95$, $p < .01$). There were no age differences in R and F responses to studied items, but older adults were more likely than the young to make false R responses to unstudied items (.08 vs. .02; $t(75) = 4.95$, $p < .01$); the age difference in F responses to new items was not reliable. Overall then, there was an age-related decrease in recognition accuracy, driven primarily by an age-related increase in false recollection.

For cued recall, target output was higher for young, compared with older, adults for stems corresponding to studied (.60 vs. .46; $t(75) = 5.78$, $p < .001$) and unstudied items (.32 vs. .27; $t(75) = 2.35$, $p < .05$). Given these differences, all subsequent analyses were performed on the rescaled proportions. The rescaled data showed older adults to be less likely than the young to correctly classify recalled studied words as old (R + F: .95 vs. .90; $t(75) = 2.50$, $p < .05$), a difference driven mainly by older adults making fewer correct R responses (.68 vs. .82; $t(75) = 3.76$, $p < .01$). And, as in recognition, older adults also made more incorrect R responses to unstudied items (.26 vs. .09: $t(75) = 3.77$, $p < .01$).

A final set of analyses compared performance in the recognition and cued-recall tests. Compared with recognition, cued recall showed a sharp increase in the proportion of false alarms (R + F to unstudied targets) for both the young (.55 vs. .14; $t(75) = 8.85$, $p < .001$) and older (.63 vs. .22; $t(75) = 7.98$, $p < .001$) adults. This elevation suggests that relative to recognition, self-generation of items in cued recall increases their acceptance as old (Toth, 2005). However, this *false-memory via self-generation* effect was notably different as a function of age. For the young, the increase in false alarms from recognition to recall was mainly in their F responses (from .12 to .46), with R responses showing a much smaller increase (from .02 to .09). This makes sense given that the items triggering these responses were unstudied and thus could not give rise to true recollective detail. Older adults, by contrast, showed a much larger increase in false R responses (from .08 to .26) along with an increase in false F responses (from .15 to .37). A 2 × 2 analysis of variance (ANOVA) on false R responses revealed main effects of age, $F(1, 150) = 24.88$, MSE = .021, $p <$

Table 2
Mean Observed and Adjusted Proportions of Recollect (R), Familiar (F), and No Memory (N) Responses by Young and Older Participants on the Cued-Recall Test

	Studied items				Unstudied items			
Group	R + F	R	F	N	R + F	R	F	N
				Observed proportions				
Young								
M	.57	.50	.08	.03	.18	.03	.15	.14
SD	.12	.14	.05	.04	.11	.04	.09	.09
Older								
M	.42	.32	.09	.04	.17	.07	.11	.10
SD	.13	.14	.05	.05	.10	.07	.07	.08
				Rescaled proportions				
Young								
M	.95	.82	.13	.05	.55	.09	.46	.45
SD	.08	.14	.09	.08	.27	.12	.25	.27
Older								
M	.90	.68	.22	.11	.63	.26	.37	.39
SD	.11	.18	.13	.11	.28	.24	.23	.27

Note. Rescaled proportions were obtained by dividing observed proportions by the proportion of trials for which a target word was output. See text for details. In some cases, *R + F* scores differ from the separate *R* and *F* values due to rounding error.

.001, and test, $F(1, 150) = 31.15$, MSE = .021, $p < .001$, as well as the Age × Test interaction, $F(1, 150) = 5.14$, MSE = .021, $p < .05$.

In sum, older adults showed lower accuracy in both recognition and cued recall, as well as higher levels of false recollection. Both age groups showed higher false memory in cued recall, relative to recognition, suggesting that generating potential target items in cued recall increases feelings of oldness for those items. Older adults were particularly vulnerable to this effect, showing very high levels of false recollection in cued recall. This pattern suggests that not only do older adults have impaired recollection, they also have difficulty discriminating between recollection and familiarity.

JOLs and JOL Accuracy

Older adults' mean JOL ratings were numerically higher than those by the young adults in both recognition (51.97 vs. 45.31) and cued-recall (53.51 vs. 43.15), but ANOVA revealed no significant differences.

We analyzed JOL accuracy in three ways. The first was a back-sorting procedure similar to those used in neuroimaging studies (e.g., Wagner et al., 1998). Figures 1 (recognition) and 2 (cued recall) show mean JOL ratings for items subsequently rated R, F, or N.[2] As can be seen, JOLs at encoding were higher for items that would later receive a *recollect* judgment at test, a pattern that held for both age groups and both tests; however, JOL differences between R items and the other two categories were generally greater for young, compared with the older, adults. These impressions were confirmed in two ANOVAs with age and test judgment (R/F/N) as factors. The ANOVA for recognition revealed effects of age, $F(1, 69) = 4.57$, MSE = 620.00, $p < .05$, and test judgment, $F(2, 138) = 59.66$, MSE = 3979.68, $p < .001$, both qualified by a significant interaction, $F(2, 138) = 5.75$, MSE = 383.22, $p < .005$. Planned comparisons revealed that R items garnered reliably higher JOLs than F or N items in both the young (ps $< .001$) and old (ps $< .001$). By contrast, JOLs did not differ for F and N items in either age group (ts < 1). The Age × Judgment interaction occurred because while R items showed no age difference ($t < 1$), older adults gave higher JOL ratings than the young adults for F and N items, $t(73) = 3.03$, $p < .005$, and $t(71) = 2.05$, $p < .05$, respectively. Stated differently, older adults' JOLs showed less discrimination between items that were later recollected and those that were not.

The ANOVA for cued recall also revealed main effects of age, $F(1, 49) = 7.18$, MSE = 4554.86, $p < .05$, and test judgment, $F(2, 98) = 24.86$, MSE = 4151.37, $p < .001$, although here the interaction was not reliable, $F(2, 98) = 2.12$, MSE = 353.24, $p = .13$. Planned comparisons showed that, as in recognition, test items given R judgments received higher JOLs at study than those judged F or N, for both the young (ps $< .001$) and older (ps $< .005$) adults. Although JOLs were numerically higher for F items compared with N items, neither difference was reliable ($p = .08$ for both age groups). Age comparisons showed the same pattern found in recognition: No difference in JOLs for items judged R, $t(75) = 1.29$; $p = .20$, but higher JOL ratings by older adults for items later judged F, $t(71) = 2.93$, $p < .01$, and N, $t(51) = 2.56$, $p < .05$.

In sum, JOLs were reliably higher for recollected items compared with those that were only familiar or not remembered. This pattern obtained for both tests and both age groups, although older adults showed less JOL discrimination between recollected and nonrecollected items.

A second set of JOL analyses used Goodman-Kruskal gamma correlations, which provide an index of how well participants' JOLs predicted their memory performance on an item-by-item basis. Gammas were computed in two ways. The first was designed to mimic standard old/new recognition by coding both R and F (old) responses as 2 and N (new) responses as 1. For the second "recollect only" method, only R responses were coded as 2, with both F and N responses coded as 1. This coding was designed to assess whether JOL accuracy would be enhanced when memory performance was restricted to instances of recollection. A 2 (age) × 2 (test) ANOVA on the standard-recognition gammas revealed only a main effect of test: recognition = .24, and recall = .09; $F(1, 144) = 13.70$, MSE = .060, $p < .001$. Notably, the effect of age was not significant ($p = .31$). By contrast, the corresponding recollect-only analysis showed an effect of age—young = .32, and old = .22; $F(1, 150) = 7.35$, MSE = .048, $p < .01$—along with an effect of test—recognition = .34, and recall = .20; $F(1, 150) = 14.27$, MSE = .048, $p < .001$. Thus, whereas age differences were not observed with standard-recognition gammas, such differences did emerge when memory performance was exclusively defined in terms of recollection.

A final set of analyses examined Pearson correlations between listwide JOLs (the mean JOL for each individual) and memory accuracy (hits–false alarms). Note that these correlations were computed across subjects and thus provide an index of how well a person's aggregate JOLs predicted his or her performance relative to others. If JOLs are driven by details that later contribute to recollection, then listwide JOLs should show the strongest relations with items given R responses at test. Consistent with this, young adults' listwide JOLs were related to the accuracy of their R judgments—recognition: $r(36) = .41$, $p < .05$; cued-recall: $r(36) = .46$, $p < .01$. It is interesting that significant negative relations were found between young adults' listwide JOLs and their F judgments—recognition: $r(36) = -.36$, $p < .05$); cued-recall: $r(36) = -.49$, $p < .01$. On the basis of the assumption that recollection and familiarity have independent influences on memory, we computed a more process-pure measure of familiarity (F/[1 − R]; see Jacoby, Yonelinas, & Jennings, 1997). Correlations between JOLs and this purified measure of familiarity were much smaller and nonsignificant (rs = .13 & .20). Note that this same general pattern obtained for the older adults although none of the individual correlations achieved significance.

Discussion

The present study produced four main findings. First, items recollected at test garnered higher immediate JOL ratings at study than did items that were either just familiar or not remembered at all. This result indicates an important role for recollection in mediating JOL accuracy. Second, although older adults showed the same pattern of

[2] Data from 4 young and 2 older adults were not included in the recognition analysis and data from 15 young and 11 older adults were not included in the cued-recall analysis because they failed to utilize all three response types during testing.

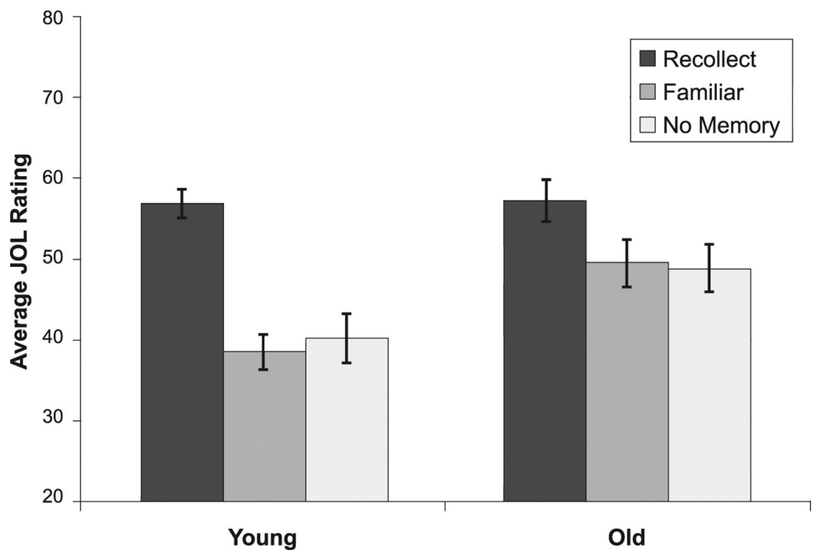

Figure 1. Average judgment of learning (JOL) rating for items rated *recollect* (*R*), *familiar* (*F*), or *no memory* (*N*) in the recognition test for young and older adults. Error bars show standard errors of the means.

JOL accuracy as the young ($R > F$ & N), the difference was weaker, suggesting an age-related deficit in the ability to monitor initial learning. Third, JOL–memory correlations computed across subjects were higher for recollected versus familiar items, suggesting that subjective memory judgments more accurately track true memory differences among people when only recollection is considered. Finally, the impact of recollection on JOL accuracy was similar in the recognition and cued-recall tests for both young and older adults.

The present study thus adds to a growing body of research suggesting that recollection plays an important role in memory monitoring. Indeed, across studies, recollection has been shown to play a major role in mediating the accuracy of immediate JOLs (current study), delayed JOLs (Carroll & Shanahan, 1997), feelings of knowing (FOKs; Hicks & Marsh, 2002), and confidence ratings (Kelley & Sahakyan, 2003). The current study also shows this monitoring–recollection link to be operating similarly in recognition and cued recall, and, along with Kelley and Sahakyan's (2003) findings, indicates recollection as an important factor mediating age differences in memory monitoring as well as aspects of performance that depend on such monitoring (Gallo et al., 2007).

Aging and JOL Accuracy

Prior research has sometimes failed to observe age-related deficits in JOL accuracy (e.g., Connor et al., 1997; Hertzog et al., 2002). In contrast, age deficits in JOL accuracy were clearly

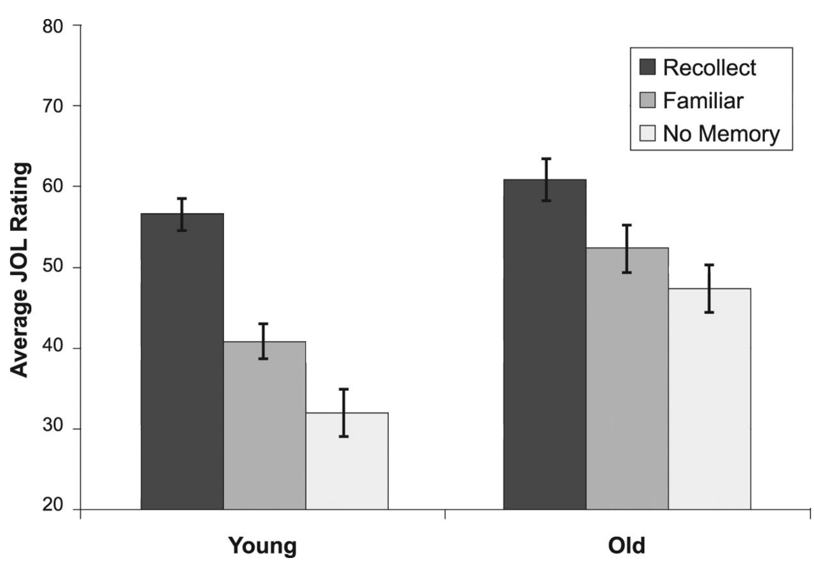

Figure 2. Average judgment of learning (JOL) rating for items rated *recollect* (*R*), *familiar* (*F*), or *no memory* (*N*) in the cued-recall test for young and older adults. Error bars show standard errors of the means.

observed in the present study in both the back-sorting data and gamma correlations. Part of the discrepancy may reflect the degree to which recollection drives overall memory performance. In the present study, participants were explicitly instructed to generate images or personally meaningful associations at study; and the requirement to make recollect/familiar judgments at test likely increased the use of these details. Given age-related deficits in recollection and evidence that memory monitoring is tied to memory for detail (Kelley & Sahakyan, 2003), age deficits in JOL accuracy may only be observed in tasks that emphasize recollection. Conversely, age deficits are unlikely to be observed when either JOLs at study or memory responses at test are driven by more automatic factors. Evidence supporting this account can be found in the present study in that age differences in the gamma correlations only emerged in the recollect-only analysis; that is, when only R responses were coded as hits. When F responses were also coded as hits (standard recognition), no significant age differences were observed.

There are at least two ways in which age-related deficits in recollection could undermine memory monitoring. One is that, relative to the young, older adults may fail to recollect details that influenced their JOL ratings at study. If an item evokes sufficient detail to garner a high JOL but that detail is not recollected at test, it would likely get classified as F or N. This would explain why older adults' JOLs for F and N judgments were higher than those for the young (see Figures 1 and 2). A second possibility concerns false memory. That is, items that do not garner high JOLs at study may nevertheless be falsely "recollected" at test, and thus get classified as R. Support for this explanation is provided by our finding that, relative to the young, older adults made significantly more false R responses to unstudied items in both recognition and cued recall. Given such high rates of false recollection, it is likely that some of the older adults' "correct" R responses to old items were not based on details that were part of the initial encoding.

Recollection, JOL Accuracy, and Theories of Monitoring

Although we agree with others that metacognitive judgments are driven by many factors, including heuristic factors such as fluency, the findings of the present study encourage more research on the use of contextual details in mediating the accuracy of memory judgments. In the case of JOLs, people are likely capable of generating a variety of details that may influence their judgments including intra- or extralist associations, images, idiosyncratic elaborations, or personal experiences. An interesting question is whether different detail types have systematic influences on JOLs. Personal experiences, for example, might naturally engender high JOLs. Alternatively, the type of detail may be less important than its distinctiveness. Regardless, the use of contextual details in making JOLs and the recollection of those details at test can be viewed as a form of "privileged access" that drives metacognitive accuracy (Lovelace, 1984).

Of course, recollection is not infallible as shown by numerous demonstrations of false recollection in both young and older adults (e.g., Jacoby, Bishara, Hessels, & Toth, 2005; Roediger & McDermott, 1995). Nevertheless, a strong case can be made that, compared with more automatic forms of memory, recollection is likely to be more reliable across situations and may be especially important for memory monitoring (Gallo et al., 2007; Kelley & Sahakyan, 2003). Indeed, although JOLs are influenced by a number of factors, whether these factors affect JOL accuracy depends on the degree to which people notice and correctly interpret their influence and whether they are operating at both study and test. Contextual, idiosyncratic detail is likely to fair well on both accounts. That is, the more distinctive a cue, the more likely it will be used in forming a JOL and the higher the probability that it will be recollected at test.

An early contrast in the memory monitoring literature was between direct-access (trace-strength) and inferential (cue-utilization) views. Along with others (e.g., Begg et al., 1989), we find the inferential view most compelling. However, it is interesting to note how recollection, perhaps more than any other factor, shows some compatibility with the direct-access view. For example, in Jacoby's (1998) direct-retrieval model, recollection is thought to provide relatively error-free (albeit limited) access to the past. Similarly, episodic memory has been characterized as a form of mental time travel whereby a person "consciously recollects some prior episode or state as it was previously experienced" (Wheeler et al., 2001, p. 333). These notions try to capture the remarkable ability of humans to relive or recover highly specific details associated with prior events. We believe such details may play a critical role in the monitoring of memory.

References

Arbuckle, T., & Cuddy, L. (1969). Discrimination of item strength at time of presentation. *Journal of Experimental Psychology, 81,* 126–131.

Begg, I., Duft, S., Lalonde, P., & Melnick, R. (1989). Memory predictions are based on ease of processing. *Journal of Memory and Language, 28,* 610–632.

Benjamin, A., Bjork, R., & Schwartz, B. (1998). The mismeasure of memory: When retrieval fluency is misleading as a metamnemonic index. *Journal of Experimental Psychology: General, 127,* 55–68.

Carroll, M., & Shanahan, C. (1997). The effect of context and metamemory judgements on automatic processes in memory. *Acta Psychologica, 97,* 219–234.

Connor, L., Dunlosky, J., & Hertzog, C. (1997). Age-related differences in absolute but not relative metamemory accuracy. *Psychology and Aging, 12,* 50–71.

Craik, F. I. M. (1986). A functional account of age differences in memory. In F. Klix & H. Hagendorf (Eds.), *Human memory and cognitive capabilities, mechanisms, and performances* (pp. 409–422). Amsterdam: Elsevier.

Gallo, D. A., Cotel, S. C., Moore, C. D., & Schacter, D. L. (2007). Aging can spare recollection-based retrieval monitoring: The importance of event distinctiveness. *Psychology and Aging, 22,* 209–213.

Gardiner, J., Ramponi, C., & Richardson-Klavehn, A. (1998). Experiences of remembering, knowing, and guessing. *Consciousness and Cognition, 7,* 1–26.

Hamilton, M., & Rajaram, S. (2003). States of awareness across multiple memory tasks: Obtaining a "pure" measure of conscious recollection. *Acta Psychologica, 112,* 43–69.

Hertzog, C., Dixon, R., & Hultsch, D. (1990). Relationships between metamemory, memory predictions, and memory task performance in adults. *Psychology and Aging, 5,* 215–227.

Hertzog, C., Kidder, D., Powell-Moman, A., & Dunlosky, J. (2002). Aging and monitoring associative learning: Is monitoring accuracy spared or impaired?. *Psychology and Aging, 17,* 209–225.

Hicks, J., & Marsh, R. (2002). On predicting the future states of awareness for recognition of unrecallable items. *Memory & Cognition, 30,* 60–66.

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

Jacoby, L. L. (1998). Invariance in automatic influences of memory: Toward a user's guide for the process–dissociation procedure. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 3–20.

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.

Jacoby, L. L., Bishara, A. J., Hessels, S., & Toth, J. P. (2005). Aging, subjective experience, and cognitive control: Dramatic false remembering by older adults. *Journal of Experimental Psychology: General, 134,* 131–148.

Jacoby, L. L., Yonelinas, A. P., & Jennings, J. M. (1997). The relation between conscious and unconscious (automatic) influences: A declaration of independence. In J. D. Cohen & J. W. Schooler (Eds.), *Scientific approaches to the question of consciousness* (pp. 13–47). Hillsdale, NJ: Erlbaum.

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

Kelemen, W., & Weaver, C. (1997). Enhanced memory at delays: Why do judgments of learning improve over time?. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 1394–1409.

Kelley, C., & Sahakyan, L. (2003). Memory, monitoring, and control in the attainment of memory accuracy. *Journal of Memory and Language, 48,* 704–721.

Koriat, A., Bjork, R., Sheffer, L., & Bar, S. (2004). Predicting one's own forgetting: The role of experience-based and theory-based processes. *Journal of Experimental Psychology: General, 133,* 643–656.

Lovelace, E. A. (1984). Metamemory: Monitoring future recallability during study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 756–766.

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review, 87,* 252–271.

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

Perfect, T. J., & Dasgupta, Z. R. R. (1997). What underlies the deficit in reported recollective experience in old age? *Memory & Cognition, 25,* 849–858.

Rabinowitz, J., Craik, F., Ackerman, B., & Hinchley, J. L. (1982). Aging and metamemory: The roles of relatedness and imagery. *Journal of Gerontology, 37,* 688–695.

Richardson-Klavehn, A., & Bjork, R. (1988). Measures of memory. *Annual review of Psychology, 39,* 475–543.

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.

Rubin, D. C. (2005). A basic-systems approach to autobiographical memory. *Current Directions in Psychological Science, 14,* 79–83.

Shaw, R., & Craik, F. (1989). Age differences in predictions and performance on a cued recall task. *Psychology and Aging, 4,* 131–135.

Toth, J. P. (2005). *States of awareness in memory: A remarkable correspondence between recognition and cued recall*. Unpublished manuscript.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology, 26,* 1–12.

Wagner, A. D., Schacter, D. L., Rotte, M., Koutstaal, W., Maril, A., Dale, A. M., et al. (1998, August 21). Building memories: Remembering and forgetting of verbal experiences as predicted by brain activity. *Science, 281,* 1188–1191.

Wheeler, M., Stuss, D., & Tulving, E. (1997). Toward a theory of episodic memory: The frontal lobes and autonoetic consciousness. *Psychological Bulletin, 121,* 331–354.

Whittlesea, B. (1993). Illusions of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 1235–1253.

Yonelinas, A. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46,* 441–517.

Received April 18, 2008
Revision received December 24, 2008
Accepted December 31, 2008 ■