Neural Basis for Recognition Confidence in Younger and Older Adults

Elizabeth F. Chua and Daniel L. Schacter
Harvard University

Reisa A. Sperling
Brigham and Women's Hospital and
Massachusetts General Hospital

Although several studies have examined the neural basis for age-related changes in objective memory performance, less is known about how the process of memory monitoring changes with aging. The authors used functional magnetic resonance imaging to examine retrospective confidence in memory performance in aging. During low confidence, both younger and older adults showed behavioral evidence that they were guessing during recognition and that they were aware they were guessing when making confidence judgments. Similarly, both younger and older adults showed increased neural activity during low- compared to high-confidence responses in the lateral prefrontal cortex, anterior cingulate cortex, and left intraparietal sulcus. In contrast, older adults showed more high-confidence errors than younger adults. Younger adults showed greater activity for high compared to low confidence in medial temporal lobe structures, but older adults did not show this pattern. Taken together, these findings may suggest that impairments in the confidence–accuracy relationship for memory in older adults, which are often driven by high-confidence errors, may be primarily related to altered neural signals associated with greater activity for high-confidence responses.

Keywords: aging, memory, metamemory, monitoring, fMRI

It is well known that younger and older adults show differences in remembering previously acquired information across a variety of tasks and conditions (for review, see Zacks, Hasher, & Li, 2000). However, memory performance depends on both remembering stored information and metamemory monitoring of that information (Koriat & Goldsmith, 1996). Although the effects of aging on memory monitoring are less well studied than the effects of aging on remembering, experimental evidence points toward age-related differences in memory monitoring (C. M. Kelley & Sahakyan, 2003). In particular, many studies have shown that older adults are more likely than younger adults to make memory errors with high confidence (Dodson, Bawa, & Krueger, 2007; Dodson, Bawa, & Slotnick, 2007; Dodson & Krueger, 2006; C. M. Kelley

& Sahakyan, 2003; Norman & Schacter, 1997). It is striking, however, that relatively little is known about potential age-related differences in the neural basis of metamemorial monitoring, although numerous studies have examined the neural basis for differences in memory in aging during retrieval (Cabeza, Anderson, Houle, Mangels, & Nyberg, 2000; Cabeza, Anderson, Locantore, & McIntosh, 2002; Cabeza et al., 1997; Daselaar, Fleck, Dobbins, Madden, & Cabeza, 2006; Daselaar, Veltman, Rombouts, Raaijmakers, & Jonker, 2003; Duverne, Habibi, & Rugg, in press; Grady, 1996; Grady & Craik, 2000; Hazlett et al., 1998; Schacter, Savage, Alpert, Rauch, & Albert, 1996), and recent work has also examined age-related differences in false recognition (Dennis, Kim, & Cabeza, 2008). To help fill this gap, in the current study we used functional magnetic resonance imaging (fMRI) to examine confidence-related neural activity during recognition and confidence assessment tasks in healthy younger and older adults.

Evidence from the behavioral literature is mixed concerning age-related differences in the relationship between confidence and accuracy in memory (Dodson, Bawa, & Krueger, 2007; Dodson, Bawa, & Slotnick, 2007; Dodson & Krueger, 2006; Dunlosky & Hertzog, 2000; C. M. Kelley & Sahakyan, 2003; Matvey, Dunlosky, Shaw, Parks, & Hertzog, 2002; Pliske & Mutter, 1996). Some studies have documented that older adults show less correspondence between confidence and accuracy (Dodson, Bawa, & Krueger, 2007; Dodson, Bawa, & Slotnick, 2007; Dodson & Krueger, 2006; C. M. Kelley & Sahakyan, 2003), whereas other studies have not found age differences (Dunlosky & Hertzog, 2000; Matvey et al., 2002) or have even shown more accurate confidence judgments in older adults (Pliske & Mutter, 1996). However, there is growing evidence that older adults exhibit higher rates of false recognition (Dodson & Schacter, 2002; Jacoby, Bishara, Hessels, & Toth, 2005; Koutstaal & Schacter, 1997; Norman & Schacter, 1997; Schacter, Koutstaal, & Norman,

1997) and also show higher confidence in their false memories (Dodson, Bawa, & Krueger, 2007; Dodson, Bawa, & Slotnick, 2007; Dodson & Krueger, 2006; Karpel, Hoyer, & Toglia, 2001; C. M. Kelley & Sahakyan, 2003). Taken together, these findings could indicate that monitoring effectiveness may be dependent on the level of confidence expressed because older adults make more high-confidence than low-confidence errors compared to younger adults, and therefore their monitoring accuracy is different for high and low confidence ratings. This pattern could explain some of the discrepancies seen in the literature, because many studies examine the confidence–accuracy relationship overall and often do not account for differences in monitoring accuracy for high or low confidence ratings by examining them separately (Dodson, Bawa, & Krueger, 2007).

Given that there are differences in the confidence–accuracy relationship between younger and older adults in episodic tasks, it may be that older adults base their confidence judgments on different sources of information than younger adults do. Overall, confidence judgments are thought to be based on a combination of memory strength and additional analytic factors (Busey, Tunnicliff, Loftus, & Loftus, 2000; C. M. Kelley & Jacoby, 1996; C. M. Kelley & Lindsay, 1993; Stretch & Wixted, 1998; Wells, Olson, & Charman, 2003). One possible explanation for an altered confidence–accuracy relationship is that older adults have a lower criterion for high-confidence decisions compared to younger adults. For example, older adults have been shown to make recognition decisions more on the basis of familiarity, whereas younger adults will often make recognition decisions on the basis of recollection (Daselaar et al., 2006; Parkin & Walter, 1992). Given this difference in memory, older adults may make high-confidence decisions on the basis of familiarity, whereas younger adults may make high-confidence decisions on the basis of recollection (C. M. Kelley & Sahakyan, 2003), and this could lead to differences in confidence-related activity (Kim & Cabeza, 2007).

Research has shown that confidence assessment also involves factors in addition to memory strength (Busey et al., 2000; C. M. Kelley & Jacoby, 1996; C. M. Kelley & Lindsay, 1993; Wells et al., 2003), which raises the possibility that these additional factors may also vary with aging. This point may be especially relevant because older adults may show reduced memory strength compared with younger adults. It is possible that if older adults are aware that weaker memories are not as useful as stronger memories in making memory decisions, they would rely more on other information (C. M. Kelley & Jacoby, 1996). Experimental evidence has also shown that confidence decisions can be based, in part, on cue-related information, such as cue familiarity and ease of processing the cue (Busey et al., 2000). If older adults lack strong cue–target associative signals, which is likely given the deficits in associative memory in aging (e.g., Chalfonte & Johnson, 1996; Krause et al., 2000; Mitchell, Johnson, Raye, & D'Esposito, 2000; Naveh-Benjamin, Hussain, Guez, & Bar-On, 2003), they may instead base their decisions on information related to the cue. Basing confidence judgments on cue-related factors will often increase confidence without a corresponding increase in accuracy (Busey et al., 2000) and could thus be a source for high-confidence errors in older adults.

Metamemory in older adults has often been studied using questionnaires (e.g., Chaffin & Herrmann, 1983; Hultsch, Hertzog, & Dixon, 1987), and answers to these metamemory questions have been shown to be based in part on prior beliefs and in part on processes occurring at the time the judgment is made (McDonald-Miszczak, Hertzog, & Hultsch, 1995; McFarland, Ross, & Giltrow, 1992). It is widely believed that memory declines with age (Magnussen et al., 2006); indeed, older adults generally rate their memory as having gotten worse over time (Dixon & Hultsch, 1983; Gilewski, Zelinski, & Schaie, 1990; Hultsch et al., 1987), even though they will also make contradictory claims that they have not experienced any changes in how often they forget things (Chaffin & Herrmann, 1983). These kinds of general beliefs or heuristics about their own memory could influence older adults' confidence judgments in either direction. For example, older adults who believe their memory has gotten worse may give many low-confidence responses. However, older adults who believe that they do not forget things may make more high-confidence responses. Although the metamemory items on questionnaires, which tend to ask about memory function more generally, are different from trial-by-trial metamemory judgments in a confidence rating task, the findings from questionnaires raise the possibility that older adults might also make trial-by-trial confidence judgments on the basis of both memory monitoring and general beliefs or heuristics about their own memory.

Findings from functional neuroimaging studies raise the possibility that older adults may show differences in the neural underpinnings of memory monitoring compared with younger adults. We previously documented confidence-related neural activity in medial temporal lobe, dorsal and ventral medial prefrontal, medial parietal (including posterior cingulate, precuneus, and retrosplenial cortex), and lateral parietal (including the intraparietal sulcus and inferior parietal lobule) regions during a memory task (Chua, Schacter, Rand-Giovannetti, & Sperling, 2006). Similar regions have been implicated in the *default network* (Gusnard & Raichle, 2001; Raichle et al., 2001), a group of regions that have shown task-induced deactivations in a wide variety of cognitive tasks (Shulman et al., 1997) and are thought to be involved in cognitive operations that occur during rest and passive states (Gusnard & Raichle, 2001; Raichle et al., 2001). Similar default network regions have shown alterations in both normal and pathological aging (Andrews-Hanna et al., 2007; Celone et al., 2006; Grady, Springer, Hongwanishkul, McIntosh, & Winocur, 2006; Greicius, Srivastava, Reiss, & Menon, 2004; Herholz et al., 2002; Lustig et al., 2003; Miller et al., 2008), both in terms of resting metabolism (Herholz et al., 2002) and in functional tasks (Lustig et al., 2003; Miller et al., 2008). Furthermore, these regions show disrupted functional correlations at rest and white matter integrity with aging (Andrews-Hanna et al., 2007). Given the structural and functional alterations in these regions, the question of whether older adults show different confidence-related activity than younger adults arises.

Despite evidence of structural and functional changes within default network regions associated with aging, it still remains possible that older adults would also show confidence-related magnetic resonance (MR) signals in these regions that were similar to young adults. One study that examined recollection and familiarity in aging used confidence ratings to measure familiarity and showed that younger and older adults exhibited similar confidence effects in posterior cingulate and left lateral parietal regions (Daselaar et al., 2006). However, this study examined correct responses only; it could be that this activation reflects memory

strength, as the authors suggest, but it leaves open the possibility that these regions modulate on the basis of subjective confidence level in both younger and older adults. Another study investigating age-related differences in true and false recognition using a categorized word-list paradigm reported that both younger and older adults showed differences between high- and low-confidence responses for both true and false retrieval in the left inferior parietal cortex and in the left superior frontal cortex (Dennis et al., 2008). If the parietal and frontal regions overlapped between true and false recognition, it may be that activity in these regions tracks confidence rather than retrieval of studied words (true recognition) or retrieval of semantically related words (false recognition). However, these regions were not directly compared, so it is difficult to tell if they overlapped. Nevertheless, these studies indicate that younger and older adults can show similar modulation in medial temporal, medial prefrontal, medial parietal, and lateral parietal regions and raise the possibility that they may modulate on the basis of the subjective confidence level expressed.

In this study, we examine neural activity related to confidence and accuracy in younger and older adults during a face–name associative memory task. We separately examined trials in which participants made a three-alternative forced-choice recognition decision and trials in which participants made a postrecognition confidence assessment about the accuracy of their recognition decision. Two aims of this study were to explore the consistency of the confidence-related activity (Chua et al., 2006) and to examine differences in confidence-related activity associated with age-related performance changes in memory. There were three alternative possible patterns of confidence-related neural activity: (a) If younger and older adults show similar patterns of confidence-related activity, despite performance differences, then the neural signatures of confidence in recognition memory are robust and invariant across age and performance; (b) if younger and older adults show different patterns of confidence-related activity, then the neural signatures of confidence may be dependent on performance or age; and (c) if younger and older adults show similar confidence-related activity in one level of confidence (high or low) but not the other, then differences in confidence-related activity are dependent on the rating given, and this may help explain why older adults exhibit differences in confidence–accuracy calibration.

Method

Participants

Sixteen healthy, older adults (10 women, 6 men; ages 66–81) were recruited from a larger longitudinal study on cognition and aging to participate in this study. These older adults were given the Clinical Dementia Rating (CDR; Morris, 1993) scale as part of the longitudinal study; each participant had a CDR score of 0.0 with a sum of box scores of 0.0 and performed within 1.5 *SD* of age and education matched norms on standard neuropsychological tests, indicating that the participant was not showing signs of cognitive impairment. Data from the 16 younger participants with complete data from Chua et al. (2006) were used as the comparison group in this study (9 women, 7 men; ages 21–29). All participants were free from psychiatric and neurologic disorders and were screened for contraindications to MRI. Each participant provided written informed consent in a manner

approved by the Human Research Committee at Massachusetts General Hospital, Boston, Massachusetts.

Cognitive Activation Task

Participants were scanned during encoding, recognition, and retrospective confidence judgment tasks using a face–name associative memory paradigm (Figure 1; Chua et al., 2006). These tasks were presented in a mixed event-related and block design. Prior to each block, participants viewed instructions that specified whether the block was an encoding or recognition–confidence block and gave specific task instructions. Each stimulus was presented for a duration of 3.5 s. The stimuli were intermixed with brief periods of visual fixation (white cross presented on a black background), ranging from 0.5 to 6.0 s. Fixation trials were intermixed with stimuli using a jittered design and optimized using OptSeq (n.d.). Although the paradigm was presented in a mixed event-related and block design, the jittering was optimized for separating trials within each block because the primary reason for a mixed design was to minimize task switching, not to assess state effects. The paradigm was presented using MacStim 2.5 (Darby, 1993) on a Macintosh Powerbook and viewed by participants in the scanner using a rear projection system.

Encoding blocks consisted of four encoding stimuli intermixed with fixation crosses. Encoding stimuli were digital photographs of unfamiliar faces presented on a black background paired with a fictional first name printed in white underneath. In addition to instructions at the beginning of each block, each encoding trial included the word *learn* printed in white above each stimulus to remind participants of the task. During encoding, participants were instructed to try to remember the name associated with the face for later testing and also to make a purely subjective decision about whether or not the name fit the face. Participants indicated their decision with a button press: *yes*, the name fits the face; *no*, the name does not fit the face; or *don't know* if the name fits the face.

Recognition–confidence blocks consisted of alternating recognition stimuli and confidence judgment stimuli intermixed with fixation, with four trials of each type. During recognition trials, participants viewed each face seen during encoding with three names (one correct, two incorrect) printed underneath the face. One of the incorrect names was one that was paired with a different face; this was done to ensure that the recognition decision was based on the face–name association and not based solely on familiarity of the name. The other incorrect name had not been presented during encoding. The word *recall* was printed in white above the stimulus to remind participants to try to recall the name associated with the face to make their recognition decision. The task was a recognition task in which participants indicated via button press whether the correct name for the face was on the *left*, *middle*, or *right*.

After a varying interstimulus interval following fixation, ranging from 0.5 to 6.0 s, participants viewed confidence judgment stimuli, which were similar to the recognition stimuli, but the names were presented in grey and above the face with the words *high* and *low* printed in white below the face. The name choices were presented to avoid working memory components that could be associated with holding the name choices online. The words *confidence rating* were printed above the stimuli to remind participants of the task. Participants indicated via button press whether they had *high* or *low* confidence that they had chosen the correct name.
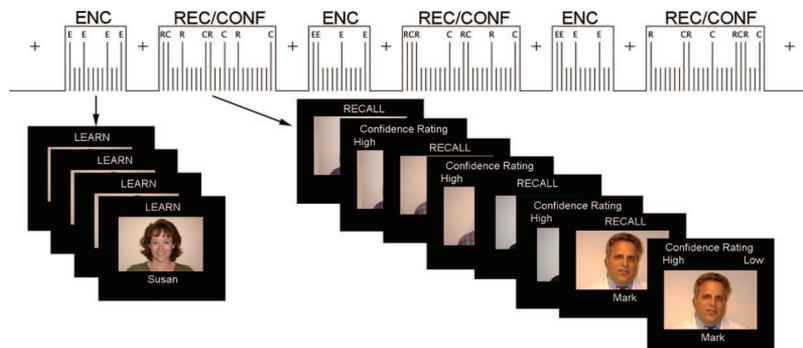
Figure 1. Schematic of the face–name associative paradigm used to study confidence in recognition memory. Stimuli were presented in a mixed block and event-related design with each run consisting of three encoding (ENC) blocks, three recognition/confidence assessment blocks (REC/CONF), and alternating blocks of visual fixation (+). Each large tick mark represents a single presentation of a 3.5-s face–name stimulus followed by 0.5 s of fixation. Each smaller tick mark represents 1 s of visual fixation. In each encoding block, there were four novel face–name pair encoding trials (E) with intermixed fixation trials. In each recognition/confidence assessment block, subjects viewed four faces that were encoded approximately 5 min earlier in the previous run. In the recognition trials (R), subjects chose the correct name among three names. Then, after a varying interstimulus interval, in the confidence assessment trials (C) subjects indicated whether they had high or low confidence that they chose the correct name. The individuals whose faces appear here were aware that they might be used in academic publications and agreed to this at the time the photographs were taken. Adapted from "Understanding metamemory: Neural correlates of the cognitive process and subjective level of confidence in recognition memory," by E. F. Chua, D. L. Schacter, E. Rand-Giovannetti, and R. A. Sperling, *NeuroImage, 29,* p. 1151. Copyright 2006 by Elsevier.

Participants viewed 120 stimuli for each task across 10 runs that each lasted 4 min 24 s. Each recognition–confidence block occurred in the subsequent run to the corresponding encoding blocks, approximately 5 min later. The faces that were retrieved in the first run were encoded in a practice run that was done in the scanner prior to functional image acquisition.

Younger and older adults received different prescan training outside of the scanner. Younger adults completed two practice runs prior to scanning (one outside the scanner and one inside the scanner) to familiarize them with the task. Older participants, on the other hand, had much more extensive training to ensure that they understood the task and that they performed the task adequately and responded within the given time period. Older participants completed two shorter instructional runs with longer stimulus durations and fewer stimuli to ensure that they understood the task outside the scanner. They then completed another four longer runs to familiarize them with the timing of the task outside the scanner, plus one additional practice run in the scanner.

Behavioral Analyses

Trials were categorized on the basis of subjective confidence level and objective recognition accuracy, yielding four main conditions for recognition and four main conditions for confidence judgments: high confidence–correct (HC-Correct), high confidence–incorrect (HC-Incorrect), low confidence–correct (LC-Correct), and low confidence–incorrect (LC-Incorrect). Proportion of responses and reaction time were compared for both younger and older adults using analysis of variance (ANOVA) and *t* tests in SPSS (SPSS Inc., Chicago, IL) and were considered significant at $p < .05$, two-tailed.

To quantify metamemory accuracy, which is the congruency between subjective confidence and objective recognition accuracy, we used the Hamann index (Schraw, 1995). There are two kinds of measurement that are typically used to assess metamemory accuracy: absolute measures and relative measures (Nelson, 1996). Absolute measures, such as the Hamann index and calibration curves, measure whether the subjective value (in this case, confidence) given to a trial is followed by the occurrence of that value on the criterion test (in this case, the recognition test). In contrast, relative measures, such as the gamma statistic, assess whether the subjective value given to one trial compared to another trial is followed by the same ordering of these items on the criterion test. We chose the Hamann index because it can be used with binary measures of confidence, whereas gamma cannot, and because absolute measures may be more sensitive to detecting differences related to a particular confidence level (Dodson, Bawa, & Krueger, 2007). The Hamann index is calculated by subtracting the number of nonmatches from the number of matches and then dividing this by the total number of cases; it ranges from $+1$ to -1; $+1$ represents perfect concordance, -1 represents perfect discordance, and 0 represents chance. HC-Correct and LC-Incorrect were considered matches, and HC-Incorrect and LC-Correct were considered nonmatches. This index was calculated for metamemory accuracy overall ([(HC-Correct + LC-Incorrect) − (HC-Incorrect + LC-Correct)]/all) for each participant, using the number of trials in each condition. We calculated a similar index within each confidence rating and calculated within high-confidence [(HC-Correct − HC-Incorrect)/all HC] and low-confidence responses [(LC-Incorrect − LC-Correct)/all LC]. For each group, a one-sample *t* test was used to compare the indices to chance and

was considered significant at $p < .05$, two-tailed. The two groups were compared using independent samples t tests and considered significant at $p < .05$, two-tailed.

Imaging Parameters

Scanning was performed on a Siemens 3T Trio scanner (Siemens Medical Systems, Iselin, NJ) with a three-axis gradient head coil. Functional data were collected using a gradient-echo echo-planar pulse sequence (return time = 2,000 ms, echo time = 30 ms, flip angle = 90). Twenty-eight oblique coronal slices were acquired perpendicular to the anterior commissure–posterior commissure line (5 mm slices, skip 1 mm). Each functional run consisted of 132 time points and lasted 4 min 24 s. Ten functional runs were collected from each participant. For one of the older adults, two functional runs were unusable due to scanner malfunction.

Imaging Data Analyses

The fMRI data were preprocessed and analyzed using SPM2 (Wellcome Trust Centre for Neuroimaging, 2003) for MATLAB (Mathworks, Natick, MA). Images were motion corrected using INRIAlign, a motion correction algorithm unbiased by local signal changes. Next, the data were spatially normalized to the standard SPM2 echo planar imaging template and resliced into $3 \times 3 \times 3$ mm resolution in Montreal Neorological Institute space. The data were then spatially smoothed using an 8-mm full-width half-maximum isotropic Gaussian kernel. First, individual participants' data were entered into a general linear model with each run concatenated in time and treated as a single time series. In addition to the cognitive conditions of interest, additional regressors were used to appropriately implement a high pass filter of 70 s. The cognitive conditions that were modeled included HC-Correct, HC-Incorrect, LC-Correct, and LC-Incorrect trials during encoding, recognition, and confidence assessment. Trial onsets were based on the stimulus onsets, and trials were modeled as events (i.e., as a stick function with a duration of 0 s) using the canonical hemodynamic response function.

At the next level, contrasts for each individual participant's data were modeled in a mixed ANOVA, treating each participant as a random effect. This model allowed us to examine both within-group effects and Group $\times$ Condition interactions. For all models the between-groups factor was age (younger vs. older adults). Two separate mixed ANOVAs were created, one for trials during confidence assessment and one for trials during recognition. The conditions in the ANOVA were contrasts of (a) HC-Correct, (b) HC-Incorrect, (c) LC-Correct, and (d) LC-Incorrect compared to baseline fixation.

Functional Region of Interest Analyses

Regions of interest (ROIs) were generated separately for confidence and recognition on the basis of the statistical parametric map for the F test for all effects of interest thresholded at $p < .001$, uncorrected at the voxel level and corrected for multiple comparisons at the cluster level with $p < .05$ (46 contiguous voxels = $p < .05$, corrected at the cluster level for t contrasts). These regions were then subjected to post hoc analyses for modulation based on confidence level and recognition accuracy. Finite impulse response

function time courses were extracted from suprathreshold clusters using MarsBar (Brett, Anton, Valabregue, & Poline, 2002). The average percentage of signal change from 2 to 8 s within the clusters were entered into mixed ANOVAs where the between-groups factor was age (younger, older) and the within-group factors were confidence level (high, low) and recognition accuracy (correct, incorrect). Regions that showed significant interactions in the mixed ANOVA were also entered into subsequent ANOVAs and paired t tests to determine the nature of the interactions. These analyses were performed using SPSS, and results were considered significant at $p < .05$, two-tailed.

Within-Group Performance Analyses

For both younger and older adults, within-group analyses were conducted to examine effects of recognition accuracy and metamemory accuracy on fMRI activity. Within each group, we performed a median split based on overall recognition accuracy and another median split based on the overall Hamann index, such that there were high-performing younger adults, low-performing younger adults, high-performing older adults, and low-performing older adults. We then performed between-groups analyses using the two-sample t test in SPM at the whole-brain level and also tested for between-groups differences in the ROIs that showed confidence-related neural activity. We also performed whole-brain correlation analyses within the younger and older adult groups to examine which brain regions showed brain activity significantly correlated with recognition accuracy and the Hamann index. Results were considered significant at $p < .001$ uncorrected at the voxel level and corrected at $p < .05$ at the cluster level (46 voxels).

Results

Behavioral Results

Younger and older adults showed significantly different behavioral responses with respect to both recognition accuracy and confidence judgments. Two-tailed, two-sample t tests showed significant differences between younger and older adults in the overall proportion of correct responses ($M \pm SEM$): younger, 0.79 ± 0.10; older, 0.42 ± 0.04; $t(30) = 14.37$, $p < .00001$; the proportion of HC-Correct out of all HC responses: younger, 0.88 ± 0.09; older, 0.48 ± 0.05; $t(30) = 15.15$, $p < .00001$; and the number of HC-Correct responses out of all Correct responses: younger, 0.92 ± 0.05; older, 0.65 ± 0.05; $t(30) = 5.82$, $p < .00001$; with younger adults performing better than older adults. One-sample t tests showed that both younger, $t(15) = 19.00$, $p < .00001$, and older, $t(15) = 10.16$, $p < .00001$, adults attained above-chance performance (33%). Younger and older adults did show similar performance for low-confidence responses, with nearly identical proportions of LC-Correct responses out of all LC responses: younger, 0.35 ± 0.03; older: 0.33 ± 0.02; $t(30) = 0.68$, $p < .51$; and these proportions were not different from chance in either group. Thus, assuming that the low-confidence response option reflected guessing, both younger and older adults appear to be well calibrated during low confidence responses because the proportion correct for these responses did not differ from chance. The number of trials in each condition for each age group is shown in Table 1.

Table 1

Number of Trials and Reaction Time (Mean ± SEM) for Both Younger and Older Adults Based on Confidence and Accuracy During Both Recognition and Confidence Assessment Trials

	Correct		Incorrect	
Variable	Younger M ± SEM	Older M ± SEM	Younger M ± SEM	Older M ± SEM
Number of trials				
High confidence	85.4 ± 3.0	30.9 ± 2.2	7.1 ± 1.1	17.4 ± 2.5
Low confidence	12.1 ± 2.2	34.3 ± 2.9	12.8 ± 1.4	32.6 ± 2.8
Confidence RT (s)				
High confidence	0.89 ± 0.05	0.96 ± 0.06	0.99 ± 0.08	0.95 ± 0.05
Low confidence	1.18 ± 0.09	1.00 ± 0.07	1.12 ± 0.09	1.02 ± 0.07
Recognition RT (s)				
High confidence	2.10 ± 0.05	2.50 ± 0.07	2.26 ± 0.10	2.61 ± 0.08
Low confidence	2.83 ± 0.06	2.80 ± 0.11	2.82 ± 0.10	2.77 ± 0.08

Note. RT = reaction time.

Metamemory accuracy was quantified using the Hamann index for overall metamemory accuracy and assessed within high-confidence and low-confidence responses. Younger adults exhibited greater overall metamemory accuracy compared to older participants: younger, $0.67 ± 0.04$; older: $0.10 ± 0.02$; $t(30) = 12.62$, $p < .00001$; although both the younger and older adults had metamemory accuracy that was greater than chance ($p < .0005$). Within high-confidence responses, younger adults showed metamemory accuracy that was greater than chance, $t(15) = 16.60$, $p < .00001$, but older adults did not; thus, younger adults had greater overall metamemory accuracy compared to older adults: younger, $0.75 ± 0.05$; older, $-0.04 ± 0.03$; $t(30) = 15.15$, $p < .00001$. This finding is consistent with younger adults showing more correct responses within high-confidence responses and older adults showing approximately equal proportions of correct and incorrect responses within high-confidence responses. Within low-confidence responses, both the younger and older adults showed metamemory accuracy that was greater than chance overall ($p < .0005$), and the two groups did not significantly differ in metamemory accuracy: younger, $0.29 ± 0.06$; older, $0.34 ± 0.04$; $t(30) = -0.678$, *ns*. This finding is consistent with both younger and older adults showing more incorrect responses within low-confidence responses because LC-Incorrect responses were considered matched (congruent) and LC-Correct responses were considered nonmatched (incongruent). The metamemory accuracy data indicate that both younger and older adults used the low-confidence response option similarly but differed in their use of the high-confidence response option. This pattern is consistent with previous research showing that older adults make more high-confidence errors than younger adults (Dodson, Bawa, & Krueger, 2007; Dodson, Bawa, & Slotnick, 2007; Dodson & Krueger, 2006; C. M. Kelley & Sahakyan, 2003; Norman & Schacter, 1997).

Mixed ANOVAs with age as the between-groups factor and confidence and accuracy as within-group factors showed significant differences in reaction time both between and within groups during both confidence assessment and recognition (Table 1). During confidence assessment, there was a main effect of confidence and Confidence × Group and Confidence × Accuracy × Group interactions. The Confidence × Group × Accuracy interaction was driven by a nonsignificant Confidence × Accuracy

effect ($p < .053$, $\eta_p^2 = .23$) in the younger group only, with HC-Correct responses made faster than HC-Incorrect responses, $p < .068$, $\eta_p^2 = .21$; and similar reaction times (RTs) for low-confidence responses. Both the younger and older groups showed a main effect of confidence, with high-confidence responses made faster than low-confidence responses: younger, $p < .008$, $\eta_p^2 = .38$; older, $p < .051$ (*ns*), $\eta_p^2 = .23$. During recognition there was a main effect of confidence and Confidence × Group and Confidence × Accuracy interactions. Subsequent ANOVAs and paired t tests showed that both younger and older adults showed a significant main effect of confidence, with high-confidence responses made faster than low-confidence responses for both correct and incorrect responses ($p < .018$ for each pair-wise comparison). For both groups, the Confidence × Accuracy interaction was driven by faster responses for HC-Correct compared to HC-Incorrect responses: younger, $p < .10$; older, $p < .019$; but not for LC-Correct compared to LC-Incorrect.

We also used two-sample t tests to compare reaction times in younger and older adults on the basis of trial type. Younger adults made HC-Correct and HC-Incorrect trials faster than older adults during recognition ($p < .006$), but did not differ from older adults in the reaction time for low confidence trials during recognition. There were no significant differences in RT for HC-Correct, HC-Incorrect, LC-Correct, or LC-Incorrect responses between younger and older adults during confidence assessment.

Imaging Results

Confidence and accuracy effects during confidence assessment. The following regions showed significant activity in the overall F test for all effects of interest during confidence assessment: left and right medial temporal lobe (MTL), left intraparietal sulcus (IPS), left and right dorsolateral prefrontal cortex (DLPFC), left and right ventrolateral prefrontal cortex (VLPFC), and anterior cingulate cortex. These regions were then interrogated for effects of confidence, accuracy, Confidence × Accuracy interactions, and Condition × Group interactions.

Younger and older adults show similar activity for low greater than high confidence in frontoparietal regions. Several frontoparietal regions showed main effects of confidence level based on

ROI analyses, with greater activity for low- compared to high-confidence responses in both younger and older adults: bilateral DLPFC, VLPFC, anterior cingulate, and left IPS (Figure 2, Table 2). The left DLPFC showed a main effect of confidence and no other significant main effects or interactions. The left IPS, right VLPFC, right DLPFC, and anterior cingulate regions also showed a Confidence × Group interaction, but within-group ANOVAs revealed that both groups showed main effects of confidence, suggesting that the interaction was driven by greater differences between high- and low-confidence responses in the younger adults than the older adults. Overall, the younger and older adults show similar patterns of activity related to low confidence, but may show some differences in magnitude.[1]

The left IPS, left and right VLPFC, and anterior cingulate regions also showed an Accuracy × Group interaction, with greater activity for incorrect than correct trials in the younger adults but not older adults. Thus, these regions show confidence-related activity in both younger and older adults, but also appear to carry additional accuracy-related information in younger adults only.

Because low confidence trials had longer RTs than high confidence trials, we conducted fMRI analyses that covaried overall and task-specific RT effects. Map-wise comparisons of the main effect of confidence showed significantly greater activity during low- compared to high-confidence responses in left and right DLPFC (left: −42, 21, 21; 286 voxels; right: 39, 36, 24; 175 voxels), VLPFC (left: −33, 21, −3; 201 voxels; right: 33, 24, −6; 170 voxels), the anterior cingulate (0, 21, 48; 532 voxels), and left IPS (−42, −33, 45; 523 voxels). Thus, these regions show modulation that cannot be solely explained by time on task.

Only younger adults showed differential activity for high > low confidence in medial temporal lobes. The MTL showed greater activity for high-confidence compared to low-confidence responses during confidence assessment; however, this was primarily seen in the younger adults (Figure 2, Table 3).[2] The right MTL also showed a main effect of accuracy, with greater activity during correct compared to incorrect responses. In the analyses that accounted for reaction time, the MTL did not meet the minimum voxel extent threshold (46 voxels) to be considered significant when corrected for multiple comparisons, but some voxels were significant at $p < .001$ uncorrected at the voxel level (left: −18, −12, −27; 6 voxels; right: 33, −9, −24; 11 voxels).

Confidence and accuracy effects during recognition. Although we were primarily interested in activity during confidence assessment, we also analyzed recognition trials. The following regions showed significant activity in the overall F test for all effects of interest and were examined for effects of confidence and accuracy during recognition: left superior frontal gyrus, medial prefrontal cortex, left basal ganglia, medial parietal cortex, and bilateral posterior inferior parietal lobule (IPL; Figure 3, Table 4).

There were regions that showed modulation based on confidence level during recognition, although most regions showed effects only in young participants (Figure 3, Table 4).[3] All of these regions showed relatively greater MR signal during high- compared to low-confidence recognition trials. The basal ganglia and medial parietal regions showed significant main effects of confidence level, and no significant Confidence × Group interactions. In contrast, the medial prefrontal cortex and left IPL showed significant Group × Confidence interactions, with significant confidence-related activity in younger adults only. The right IPL showed a main effect of accuracy, with relatively greater MR signal during correct compared to incorrect responses. The left superior frontal region showed no significant effects of confidence, accuracy, or their interaction.

These regions, which have been previously implicated in the default network, typically demonstrate greater deactivation or decreased MR signal relative to baseline during more challenging cognitive tasks. During recognition, high-confidence responses were made faster than low-confidence responses, and thus we examined whether this pattern of activity would remain when accounting for RT. The medial parietal region showed significantly greater MR signal for high- compared to low-confidence trials during recognition (0, −72, 30; 147 voxels) when accounting for RT. The other default network regions, namely the IPL and medial prefrontal cortex, did not show significant differences when

---

[1] Whole brain analyses thresholded at $p < .001$ uncorrected at the voxel level and $p < .05$ corrected for multiple comparisons at the cluster level yielded similar findings to the ROI analyses. Statistical maps showed a main effect of confidence with greater activity for low- compared to high-confidence responses in bilateral DLPFC, bilateral VLPFC, anterior cingulate, and left IPS. Whole-brain analyses for a Group × Confidence interaction revealed that a subset of voxels shown in the low > high contrast within the right VLPFC, bilateral DLPFC, anterior cingulate, and left IPS showed a significant interaction effect with greater differences in the younger than older adults. Conjunction analyses with each group map thresholded at $p < .01$ for a joint probability of $p < .001$ showed overlapping activation for both groups in bilateral VLPFC and anterior cingulate regions.

[2] Whole-brain analyses thresholded at $p < .001$ uncorrected at the voxel level and $p < .05$ corrected for multiple comparisons at the cluster level yielded similar findings to the ROI analyses. Statistical maps showed a main effect of confidence with greater activity for high- compared to low-confidence responses in bilateral MTL, right precentral gyrus, medial prefrontal cortex, and left superior temporal gyrus. Whole-brain analyses for a Group × Confidence interaction revealed that a subset of voxels shown in the high greater than low in the left and right MTL showed a significant interaction with greater differences in the younger than older adults. Within group contrasts showed that only the younger adults showed significantly greater activity for high greater than low in the MTL ($p < .001$, 46 voxel extent).

[3] Whole-brain analyses thresholded at $p < .001$ uncorrected at the voxel level and $p < .05$ corrected for multiple comparisons at the cluster level yielded similar findings to the ROI analyses. Statistical maps showed a main effect of confidence with greater activity for high- compared to low-confidence responses during recognition in medial prefrontal, medial parietal, bilateral MTL, bilateral IPL, bilateral frontopolar cortex, and left lateral superior temporal cortex. Whole-brain analyses for a Group × Confidence interaction revealed that a subset of voxels shown in the high greater than low in the anterior left IPL showed a significant interaction with greater differences in the younger than older adults. Conjunction analyses with each group map thresholded at $p < .01$ for a joint probability of $p < .001$ showed overlapping activation for both groups in left posterior IPL and the medial prefrontal cortex. Within-group contrasts showed that only the younger adults showed significantly greater activity for high > low in medial prefrontal, medial parietal, bilateral MTL, bilateral IPL, bilateral frontopolar cortex, and left lateral superior temporal cortex ($p < .001$, 46 voxel extent), whereas the older adults did not show suprathreshold effects.
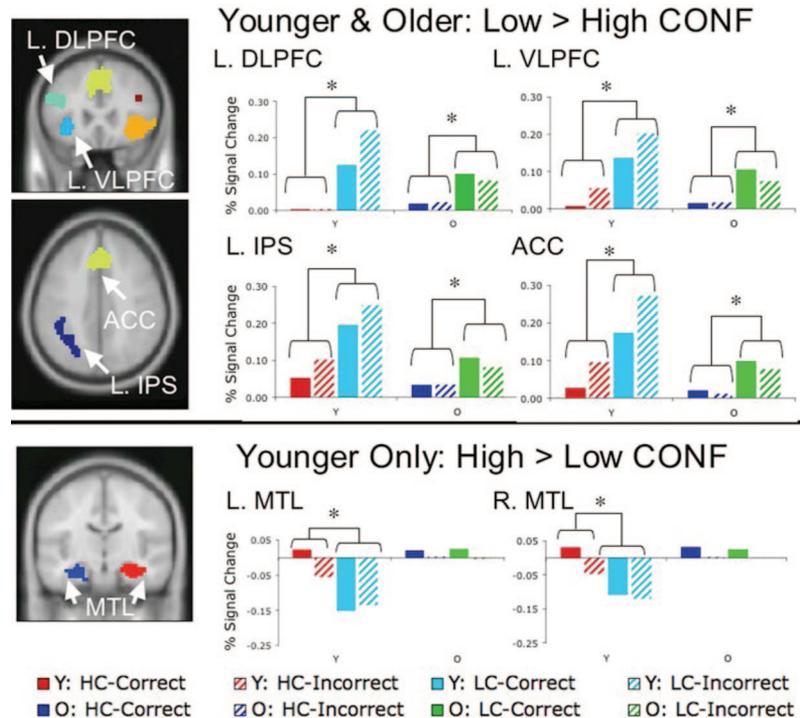
Figure 2. During confidence assessment, several regions of interest showed modulation based on the subjective level of confidence expressed. There was greater activity during low-confidence compared to high-confidence responses in both younger (Y) and older (O) adults (top) in several fronto-parietal regions, including the bilateral dorsolateral prefrontal cortex (DLPFC), bilateral ventrolateral prefrontal cortex (VLPFC), left intraparietal sulcus (IPS), and anterior cingulate (ACC) regions. In contrast, the left and right medial temporal lobe (MTL) showed effects of confidence level with high greater than low confidence in younger adults only (bottom). Bar graphs depict percentage signal change data from regions of interest for effects of confidence and accuracy in both groups. CONF = confidence; HC = high confidence; LC = low confidence. $^{*}p < .05$.

correcting for multiple comparisons, but did show smaller clusters with differential MR signal at $p < .001$ (left IPL: −48, −69, 36; 18 voxels; right IPL: 36, −57, 51; 24 voxels; medial prefrontal cortex: −3, 48, 9; 12 voxels). Thus, RT differences did not seem to account for all of the differences between high- and low-confidence responses, but may have made some contribution to the effects.

Within-group performance analyses. Neither whole-brain nor ROI analyses that compared younger adults with higher recognition accuracy to those with lower recognition accuracy or higher Hamann indices to lower Hamann indices showed any significant results; the same was true for comparisons within the older adult group. Neither whole-brain nor ROI analyses that correlated recognition accuracy or the Hamann index with fMRI activity within younger adults or within older adults showed any significant results.

Discussion

The primary goal of this study was to investigate confidence-related neural activity in younger and older adults during an associative memory paradigm. Behaviorally, both younger and older adults showed evidence that they were guessing during some trials of the recognition task, and their confidence judg-

ments suggested that they were aware that they had been guessing. Similarly, both younger and older adults showed greater activity for low- than high-confidence decisions in a frontoparietal network, and we hypothesize that these differences in activity are related to feelings of uncertainty. For high-confidence decisions, older adults showed more errors than younger adults. Younger adults showed greater MR signal for high- compared to low-confidence responses in the medial temporal lobes, whereas older adults did not. This indicates that confidence judgments in young participants are based on at least two kinds of neural signals: one where there is greater activity for high compared to low confidence and one where there is greater activity for low compared to high confidence. Our findings suggest that impairments in the confidence–accuracy relationship in older adults, which are often driven by greater high-confidence errors in older adults, may be related to the failure to show greater activity for high-confidence responses in the MTL. We also suggest that more accurate confidence judgments, such as those seen in younger adults, incorporate both greater activity for low-confidence responses in fronto-parietal regions and greater activity for high-confidence responses in the MTL, which is what we observed in younger adults.

Table 2

Regions of Interest Used to Examine Effects of Confidence Level (CONF: High or Low) and Recognition Accuracy (ACC: Correct or Incorrect) That Showed Greater Activity During Low-Confidence Compared to High-Confidence Trials

Region	Coordinates	Voxels	Main effects	Interactions	Effects in younger adults related to interaction	Effects in older adults related to interaction
Left DLPFC	−39, 21, 24	190	CONF: $F(1, 30) = 18.72$, $p < .0002$, $\eta_p^2 = .38$	*ns*		
Right DLPFC	39, 36, 27	49	CONF: $F(1, 30) = 20.97$, $p < .00008$, $\eta_p^2 = .41$	CONF × Age: $F(1, 30) = 8.15$, $p < .008$, $\eta_p^2 = .21$	CONF: $F(1, 15) = 15.40$, $p < .001$, $\eta_p^2 = .51$	CONF: $F(1, 15) = 7.23$, $p < .017$, $\eta_p^2 = .33$
Left VLPFC	−33, 24, −6	134	CONF: $F(1, 30) = 24.14$, $p < .00003$, $\eta_p^2 = .45$	ACC × Age: $F(1, 30) = 6.10$, $p < .019$, $\eta_p^2 = .17$	ACC: $F(1, 15) = 5.75$, $p < .030$, $\eta_p^2 = .28$	ACC: *ns*, $\eta_p^2 = .051$
Right VLPFC	36, 24, −9	229	CONF: $F(1, 30) = 29.07$, $p < .00001$, $\eta_p^2 = .49$	CONF × Age: $F(1, 30) = 4.93$, $p < .034$, $\eta_p^2 = .14$ ACC × Age: $F(1, 30) = 5.50$, $p < .025$, $\eta_p^2 = .16$	CONF: $F(1, 16) = 18.90$, $p < .001$, $\eta_p^2 = .56$ ACC: $F(1, 15) = 5.53$, $p < .033$, $\eta_p^2 = .27$	CONF: $F(1, 16) = 10.77$, $p < .005$, $\eta_p^2 = .42$ ACC: *ns*, $\eta_p^2 = .04$
Anterior cingulate cortex	0, 21, 48	465	CONF: $F(1, 30) = 46.76$, $p < .00001$, $\eta_p^2 = .70$ ACC: $F(1, 30) = 4.19$, $p < .049$, $\eta_p^2 = .12$	CONF × Age: $F(1, 30) = 6.68$, $p < .015$, $\eta_p^2 = .18$ ACC × Age: $F(1, 30) = 8.98$, $p < .005$, $\eta_p^2 = .23$	CONF: $F(1, 15) = 30.23$, $p < .00007$, $\eta_p^2 = .67$ ACC: $F(1, 15) = 8.53$, $p < .011$, $\eta_p^2 = .36$	CONF: $F(1, 15) = 17.03$, $p < .001$, $\eta_p^2 = .53$ ACC: *ns*, $\eta_p^2 = .06$
Left inferior parietal lobule	−33, −57, 48	333	CONF: $F(1, 30) = 40.92$, $p < .00001$, $\eta_p^2 = .58$	CONF × Age: $F(1, 30) = 7.15$, $p < .012$, $\eta_p^2 = .19$ ACC × Age: $F(1, 30) = 4.43$, $p < .044$, $\eta_p^2 = .13$	CONF: $F(1, 16) = 31.90$, $p < .00005$, $\eta_p^2 = .68$ ACC: $F(1, 15) = 4.27$, $p < .056$, $\eta_p^2 = .22$	CONF: $F(1, 16) = 9.76$, $p < .007$, $\eta_p^2 = .39$ ACC: *ns*, $\eta_p^2 = .031$

Note. Main effects of condition (CONF, ACC) and interactions between conditions or groups (age: younger or older) from a mixed analysis of variance (ANOVA) are reported. For significant interactions, results from relevant within-group ANOVAs are reported. Partial eta-squared (η_p^2) values are reported as a measure of effect size. DLPFC = dorsolateral prefrontal cortex; VLPFC = ventrolateral prefrontal cortex.

Similar Neural Signals Related to Uncertainty in Younger and Older Adults

Both younger and older adults showed greater activity in bilateral DLPFC, VLPFC, anterior cingulate, and left IPS for low-compared to high-confidence trials during confidence assessment. These findings are consistent with previous fMRI studies in younger adults that showed greater activity in right dorsolateral prefrontal (Henson, Rugg, Shallice, & Dolan, 2000; Kim & Cabeza, 2007) and parietal regions (Kim & Cabeza, 2007; Moritz, Glascher, Sommer, Buchel, & Braus, 2006) for low compared to high confidence during recognition memory in younger adults, and they expand on the previous findings in two ways. First, these studies required participants to make a one-step recognition decision that incorporated both the confidence and the old–new recognition judgment. In our task, the confidence trial was examined separately, and we observed confidence-related activity during confidence assessment trials, but not during recognition trials, indicating that the prefrontal and parietal activity is related to the confidence decision. Second, younger and older adults showed similar patterns of neural activity for low- greater than high-confidence responses, and showed similar metamemory accuracy during low-confidence responses, which suggests that this pattern of activity remains despite age-related changes in recognition performance.

Greater activity in these prefrontal and parietal regions for low-compared to high-confidence decisions may reflect neural signals of uncertainty. Behaviorally, younger and older adults showed recognition performance consistent with guessing during low-confidence responses, which should occur in a forced choice recognition paradigm if participants are unsure about the correct decision. In fMRI studies of decision-making, these prefrontal and parietal regions have been implicated in decision-making under uncertainty (Huettel, Song, & McCarthy, 2005; Paulus et al., 2001) and in particular, internally attributed uncertainty (Volz, Schubotz, & von Cramon, 2004, 2005). Based on the pattern of results in this study, it appears that both younger and older adults show similar signals of uncertainty when making low-confidence responses.

The frontal and parietal regions have been implicated in a wide variety of executive functions (Bunge, Hazeltine, Scanlon, Rosen, & Gabrieli, 2002; Cohen, Botvinick, & Carter, 2000; Fincham, Carter, van Veen, Stenger, & Anderson, 2002; MacDonald, Cohen, Stenger, & Carter, 2000; Paulus et al., 2001) that may subserve assessing and signaling uncertainty. Other memory paradigms have implicated these regions in increased monitoring and memory (Henson et al., 2000; Maril, Wagner, & Schacter, 2001), and increased monitoring and searching may be signs of feeling uncertain. In memory paradigms, two potential manifestations of uncertainty include signals that are close to the old–new response

Table 3

Regions of Interest Used to Examine Effects of Confidence Level (CONF: High or Low) and Recognition Accuracy (ACC: Correct or Incorrect) During Confidence Assessment Trials That Showed Greater Activity During High-Confidence Compared to Low-Confidence Trials in Younger but not Older Adults

Region	Coordinates	Voxels	Main effects	Interactions	Effects in younger adults related to interaction	Effects in older adults related to interaction
Left medial temporal lobe	−27, −18, −21	96	CONF: $F(1, 30) = 14.67$, $p < .001$, $\eta_p^2 = .33$	CONF × AGE: $F(1, 30) = 13.99$, $p < .001$, $\eta_p^2 = .32$	CONF: $F(1, 15) = 22.85$, $p < .0003$, $\eta_p^2 = .60$	CONF: *ns*, $\eta_p^2 = .0003$
Right medial temporal lobe	39, −18, −21	108	CONF: $F(1, 30) = 15.16$, $p < .001$, $\eta_p^2 = .34$ ACC: $F(1, 30) = 5.57$, $p < .025$, $\eta_p^2 = .16$	CONF × AGE: $F(1, 30) = 13.01$, $p < .001$, $\eta_p^2 = .30$	CONF: $F(1, 15) = 26.26$, $p < .0002$, $\eta_p^2 = .64$	CONF: *ns*, $\eta_p^2 = .003$

Note. Main effects of condition (CONF, ACC) and interactions between conditions or group (age: younger or older) from a mixed analysis of variance (ANOVA) are reported. For regions with significant interactions, results from relevant within-group ANOVAs are reported. Partial eta-squared (η_p^2) values are reported as a measure of effect size.

criterion, or in a forced choice or recall situation, alternative choices that had equivalent memory strength; in both of these cases, there may be response competition or conflict. The lateral prefrontal and anterior cingulate regions have shown differential activity related to both conflict monitoring and response competition, and therefore differences in activation in these regions may signal uncertainty (Kerns et al., 2004).

Even though younger and older adults showed similar confidence-related activity in these prefrontal and parietal regions, only younger adults also showed accuracy-related differences. Younger adults showed greater activity during incorrect responses than correct responses in these prefrontal and parietal regions. This finding may seem surprising because

fMRI studies have shown increased activity in frontal-parietal regions for familiarity-based correct responses during recognition (Yonelinas, Otten, Shaw, & Rugg, 2005). However, a recent study by Kim and Cabeza (2007) also showed greater activity in frontal-parietal regions for low- compared to high-confidence responses for true recognition, but greater activity for high- than low-confidence responses for false recognition, which indicates that the same fronto-parietal regions may have roles in both top-down monitoring and control and in familiarity signals. We observed these differences during confidence assessment, not recognition, so our findings may be more likely to reflect top-down monitoring and control processes. One possibility is that younger adults experienced greater uncertainty
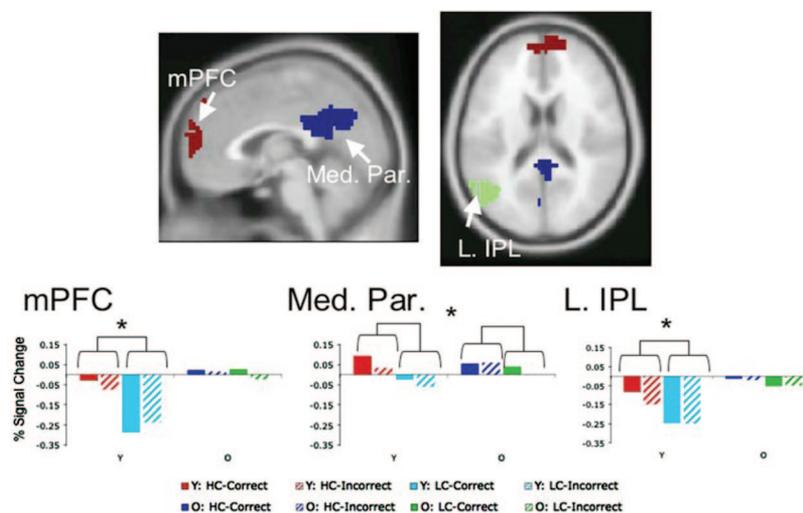
Figure 3. During recognition, several regions of interest associated with the default network showed greater activity during high-confidence compared to low-confidence responses. The left inferior parietal lobule (L. IPL) and the medial prefrontal cortex (mPFC) showed significant effects of confidence in the younger group only, whereas the medial parietal region (Med. Par.) showed a significant main effect of confidence. Bar graphs depict percentage of signal change data from regions of interest for effects on confidence and accuracy in both groups. Y = younger; O = older; HC = high confidence; LC = low confidence. *$p < .05$.

Table 4
Regions of Interest Used to Examine Effects of Confidence Level (CONF: High or Low) and Recognition Accuracy (ACC: Correct or Incorrect) During Recognition Trials

Region	Coordinates	Voxels	Main effects	Interactions	Effects driving interaction
Medial parietal cortex	0, −69, 33	546	CONF: $F(1, 30) = 5.27$, $p < .029$, $\eta_p^2 = .15$	*ns*	
Basal ganglia	−33, −9, −3	92	CONF: $F(1, 30) = 4.68$, $p < .039$, $\eta_p^2 = .14$	*ns*	
Left lateral parietal	−48, −60, 27	525	CONF: $F(1, 30) = 12.54$, $p < .001$, $\eta_p^2 = .30$	CONF × AGE: $F(1, 30) = 4.23$, $p < .048$, $\eta_p^2 = .12$	Younger CONF: $F(1, 15) = 12.50$, $p < .003$, $\eta_p^2 = .46$ Older CONF: *ns*, $\eta_p^2 = .09$
Medial prefrontal cortex	0, 57, 12	309	CONF: $F(1, 30) = 15.81$, $p < .0005$, $\eta_p^2 = .35$	CONF × AGE: $F(1, 30) = 11.44$, $p < .002$, $\eta_p^2 = .28$	Younger CONF: $F(1, 15) = 21.37$, $p < .0004$, $\eta_p^2 = .59$ Older CONF: *ns*, $\eta_p^2 = .016$
Right lateral parietal	54, −63, 30	71	ACC: $F(1, 30) = 4.65$, $p < .039$, $\eta_p^2 = .13$	*ns*	

Note. Main effects of condition (CONF, ACC) and interactions between conditions or group (age: younger or older) from a mixed analysis of variance (ANOVA) are reported. For regions with significant interactions, results from relevant within-group ANOVAs are reported. Partial eta-squared (η_p^2) values are reported as a measure of effect size.

when they chose the incorrect response, but were limited by the binary response scale in choosing high or low confidence.

An alternative explanation to uncertainty is that activity in these prefrontal and parietal regions is based on task difficulty (Duncan & Owen, 2000; Satterthwaite et al., 2007). Although this remains a possibility, given that low-confidence responses were slower and these regions have previously shown modulation based on difficulty, we believe that this idea cannot fully explain the data, because there were also reaction time differences during recognition but no corresponding differences in neural activity within these regions. Furthermore, when reaction time was entered as a parameter, there was still greater activity in these fronto-parietal regions for low- compared to high-confidence responses. Thus, we believe that neural activity in these regions may signal uncertainty and that both younger and older adults use these signals of uncertainty to make accurate low-confidence decisions.

One may be concerned about drawing strong conclusions about patterns of activity based on low-confidence responses because both younger and older adults showed recognition performance at floor in this condition. However, it is important to remember that although recognition performance was at floor, metamemory accuracy was not. The low-confidence category should reflect uncertain decisions and guesses, in which case performance at chance indicates that the participants are well-calibrated at this level of confidence. We used a Hamann-like index to test this possibility more formally and showed that in terms of calibration, both younger and older adults were significantly greater than chance and matched in performance. Nevertheless, there may still be floor effects related to recognition performance, so these findings may need to be interpreted with some caution.

Different Signals for High Greater Than Low Confidence in Younger and Older Adults

The finding that younger and older adults showed similar patterns of neural activity for low- greater than high-confidence responses, but not for high- greater than low-confidence responses, during confidence assessment suggests that signals of

uncertainty are preserved in aging and any impairments in monitoring effectiveness may be specific to high-confidence decisions. This hypothesis is consistent with our own findings that older adults made more high-confidence incorrect decisions and with reports from other studies that older adults exhibit high confidence in false memories (Dodson, Bawa, & Krueger, 2007; Dodson, Bawa, & Slotnick, 2007; Dodson & Krueger, 2006; Karpel et al., 2001; C. M. Kelley & Sahakyan, 2003). This idea may also be consistent with work using other metamemory tasks, namely the feeling-of-knowing, which has suggested that there are distinct cognitive and neural processes associated with feelings-of-knowing and feelings-of-not-knowing (Liu, Su, Xu, & Chan, 2007).

In young participants only, the MTL showed greater MR signal during high- compared to low-confidence decisions, and we have discussed these findings in young participants in greater depth in our previous report (Chua et al., 2006). Although the medial temporal lobes are thought to play a role in more objective aspects of memory, there is evidence that different regions within the MTL, such as the parahippocampal cortex and amygdala, may contribute to confidence in memory either through retrieved content (Eldridge, Knowlton, Furmanski, Bookheimer, & Engel, 2000; Yonelinas et al., 2005), or independent of retrieved content (Kensinger & Schacter, 2005, 2006; Sharot, Delgado, & Phelps, 2004), respectively.

The differential MR signal pattern seen in younger, but not older, adults for high- greater than low-confidence responses during recognition in the medial parietal, medial prefrontal, MTL, and IPL regions shows a striking resemblance to default network regions. Recent work has related the function of the default network regions to both memory and self-related processing. The default network also shows overlaps with brain regions that have shown greater MR signal for retrieval success, autobiographical memory effects, thinking about the future, and theory of mind (Buckner, Andrews-Hanna, & Schacter, 2008), and may be involved in internally directed attention. Medial prefrontal activity, in particular, has previously been linked to self-related processing

(Frith & Frith, 1999; Johnson et al., 2002; W. M. Kelley et al., 2002; Levine et al., 2004), perhaps explaining its involvement in metamemory, which engages self-reflective processes. The analyses for high- greater than low-confidence responses during recognition revealed a set of regions that resemble the default network and are consistent with the idea that default network regions are involved in internally directed cognition, because confident recognition involves self-related processing and memory processing, which are internally directed.

It is worth noting that the brain regions that showed greater MR signal for high- compared to low-confidence responses in younger adults tended to show signal decreases below baseline during low-confidence responses. These kinds of deactivations are quite common with passive baseline tasks, such as our own, and would likely change to activations with an active baseline task (Stark & Squire, 2001). Functional MRI comparisons are always relative to another condition, even if a task is labeled as a baseline task. This makes it difficult to assess whether the differences are true deactivations. In the case of fMRI comparisons, it is the relative difference between conditions that is important. However, it is worth noting that these default network regions tend to deactivate during most cognitive tasks and show greater deactivation with more demanding cognitive tasks (McKiernan, Kaufman, Kucera-Thompson, & Binder, 2003; Raichle et al., 2001). Thus, one possible interpretation is that low-confidence responses show greater deactivation because low-confidence responses are more cognitively demanding. This possibility seems less likely, however, because activity in these regions does not vary strictly with reaction time.

One possible explanation for younger but not older adults showing greater activity for high- compared to low-confidence responses relates to the default network. Default network regions have shown alterations in both normal and pathological aging (Andrews-Hanna et al., 2007; Celone et al., 2006; Herholz et al., 2002; Lustig et al., 2003; Miller et al., 2008), which raises the possibility that older adults were not able to modulate these areas in our task. Other studies, however, have shown similar modulation in these regions in younger and older adults during memory tasks (Daselaar et al., 2006). Furthermore, in a recent study investigating the self-reference effect in aging, younger and older adults showed overlapping activation in a medial prefrontal region during self-referencing (Gutchess, Kensinger, & Schacter, 2007). The medial prefrontal region in the Gutchess et al. (2007) study was more ventral than the region that showed greater activity during high- compared to low-confidence responses in younger adults, but nevertheless indicates that older adults are capable of recruiting medial prefrontal cortex in response to some task demands.

The differences between younger and older adults in neural activity for regions showing greater activity for high- compared to low-confidence decisions during both confidence assessment and recognition trials raise questions about the source of the differences associated with high confidence. Although we remain agnostic as to the origins of high-confidence errors in this study, and future research will be needed to determine the basis for the differences in younger and older adults, there are several potential explanations for the age differences associated with high confidence. One possibility is that older adults are effectively monitoring their memory, realize that they have weaker memory signals, but are limited to high- and low-confidence choices. In such a case,

younger and older adults' "high confidence" ratings would have different meanings. If this were the case, then we might expect to see similar but weaker effects in the older adults compared to the younger adults during confidence assessment trial for high- compared to low-confidence responses in the MTL. However, younger adults show significant differences based on confidence level in the MTL during confidence assessment, and if anything, the older adults show more of an effect of accuracy, not confidence, in the MTL ($p < .15$ for accuracy compared to $p > .80$ for confidence). Thus, it appears that older adults are indeed monitoring their memory differently that the younger adults for high-confidence responses.

Another possible explanation for high-confidence errors in this study is that older adults used a different criterion for assigning high confidence to a recognition judgment, which is consistent with poorer overall recognition accuracy in older adults. Data from younger adults suggest that their decisions were based, at least in part, on recollective information, because regions that showed greater activity for high- compared to low-confidence decisions during recognition (i.e., MTL, medial prefrontal, medial parietal, and IPL) have previously shown greater activity associated with recollection (Wheeler & Buckner, 2003, 2004; Yonelinas et al., 2005). Older adults did not show significant differences in MR signal in these regions, suggesting that they may have based their confidence decisions on different information than younger adults (e.g., younger adults may have made high-confidence decisions when they recalled the name, but older adults may have made high-confidence decisions when the face-name pairing was familiar). Indeed, studies of confidence in true and false recognition have suggested that there may be some differences in confidence-related activity when the judgments are based on recollection and familiarity (Kim & Cabeza, 2007). We suggest that the differences in monitoring activity are related to this potential difference in the basis for the confidence judgments.

It has previously been hypothesized that older adults misrecollect information as a result of binding problems and thus make high-confidence errors (Dodson, Bawa, & Krueger, 2007; Dodson, Bawa, & Slotnick, 2007). Behaviorally, older adults made more high-confidence errors in our study, which may be consistent with the misrecollection hypothesis because we used a difficult associative paradigm. However, if high-confidence errors were due solely to misrecollections, then we might expect to see similar activity related to high confidence in younger and older adults during confidence assessment because they would have based their decisions on information that was, at least subjectively, similar. Instead, based on our data, it seems likely that high-confidence errors in older adults are characterized by failing to show greater activity for high- compared to low-confidence decisions.

Overall, this study showed different patterns of activation related to confidence associated with high and low confidence, and it indicates that confidence is not a unitary function. We suggest that confidence judgments are made on the basis of the combination of information from brain regions that show greater activity for high than low confidence and regions that show greater activity for low than high confidence. Older adults who showed performance deficits showed similar activity to young participants for low greater than high confidence in fronto-parietal regions, but did not show the pattern of high greater than low confidence in the MTL or default network regions. Furthermore, older adults did not

show greater MR signal than younger adults for high- compared to low-confidence responses in any brain regions. This observation raises the possibility that older adults in this study primarily based their confidence judgments on a subset of the information that younger adults used; specifically, older adults may have relied on feelings of uncertainty associated with greater activity for low- compared to high-confidence decisions in frontoparietal regions. This possibility does not necessarily mean that older adults first decided whether they had a feeling of low confidence and then chose the high-confidence response option in the absence of a feeling of low confidence. If this were the case, then we would have expected low-confidence decisions to be made faster than high-confidence decisions, and this pattern was not observed. Instead, we suggest that in our study older adults did not use the same combination of information that younger adults used to make their confidence decisions, even though we did not find evidence of additional brain areas recruited by older adults.

A few limitations of our study merit consideration. Aging and performance effects are confounded in this study, making it difficult to disentangle their separate contributions. Our data do, however, indicate that there are differences in neural activity associated with the age-related performance changes in monitoring effectiveness. It is, however, the case that within each group, there were no brain regions that significantly correlated with overall accuracy or the Hamann index during confidence assessment or recognition for either age group. Future research will be needed to determine whether these patterns remain when controlling for performance between groups. Another potential concern is that older adults were given more practice with the task prior to entering the scanner and therefore may have been subject to more proactive interference compared to younger adults, which may have influenced their confidence and accuracy.

Another potential concern in interpreting these data is that there were significant differences in reaction times between high- and low-confidence responses during both recognition and confidence assessment trials, and thus differences in neural activity might be attributable to task difficulty. If regional activity modulated strictly on the basis of RT, then younger and older adults, both of whom show faster RTs for high than low responses, should both show the resulting differences in neural activity; but this pattern was not observed in the majority of regions showing greater activity for high than low responses. Furthermore, the RT differences were present in both recognition and confidence assessment; therefore, if these regions modulated strictly on RT, they should show differences during both tasks, but this is not the case for regions showing differences based on confidence in either direction (high > low or low > high). Separate analyses that included RT as a parameter showed similar effects for low > high confidence effects in fronto-parietal regions during confidence assessment and in medial parietal cortex for high > low confidence effects during recognition. However, the analyses that accounted for RT did minimize some effects of high greater than low confidence during confidence assessment and showed only small clusters that were significant at $p < .001$ that did not meet a corrected cluster extent threshold in the MTL, IPL, and medial prefrontal cortex. Thus, RT remains a correlated variable in this study—in fact, confidence judgments have been shown to be based in part on ease of retrieval (Busey et al., 2000)—but it seems unlikely that it is the only

reason there are age-related differences between high- and low-confidence responses.

In summary, we suggest that the preponderance of high-confidence errors in older adults compared to younger adults occurs not only because of differences in memory accuracy but also because older adults do not effectively monitor memory. This conclusion is based on the finding that older adults did not show greater MR signal for high-confidence responses in the same regions as younger adults during confidence assessment. However, younger and older adults showed similar patterns of neural activity associated with uncertainty, which may explain why monitoring effectiveness in older adults tends to be diminished for high-confidence but not low-confidence responses. Overall, these findings suggest that confidence judgments are not unitary decisions and are based on patterns of activity that show greater activity for high compared to low confidence and also on patterns of activity that show greater activity for low compared to high confidence. Thus, the relationship between neural activity and confidence decisions in fronto-parietal, MTL, and default network regions may be differentially affected by age-related changes in memory performance.

References

Andrews-Hanna, J. R., Snyder, A. Z., Vincent, J. L., Lustig, C., Head, D., Raichle, M. E., et al. (2007). Disruption of large-scale brain systems in advanced aging. *Neuron, 56,* 924–935.

Brett, M., Anton, J., Valabregue, R., & Poline, J. (2002, June 2–6). *Region of interest analysis using an SPM toolbox.* Paper presented at the Eighth International Conference on Functional Mapping of the Human Brain, Sendai, Japan.

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network: Anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences, 1124,* 1–38.

Bunge, S. A., Hazeltine, E., Scanlon, M. D., Rosen, A. C., & Gabrieli, J. D. (2002). Dissociable contributions of prefrontal and parietal cortices to response selection. *NeuroImage, 17,* 1562–1571.

Busey, T. A., Tunnicliff, J., Loftus, G. R., & Loftus, E. F. (2000). Accounts of the confidence–accuracy relation in recognition memory. *Psychonomic Bulletin & Review, 7,* 26–48.

Cabeza, R., Anderson, N. D., Houle, S., Mangels, J. A., & Nyberg, L. (2000). Age-related differences in neural activity during item and temporal-order memory retrieval: A positron emission tomography study. *Journal of Cognitive Neuroscience, 12,* 197–206.

Cabeza, R., Anderson, N. D., Locantore, J. K., & McIntosh, A. R. (2002). Aging gracefully: Compensatory brain activity in high-performing older adults. *NeuroImage, 17,* 1394–1402.

Cabeza, R., Grady, C. L., Nyberg, L., McIntosh, A. R., Tulving, E., Kapur, S., et al. (1997). Age-related differences in neural activity during memory encoding and retrieval: A positron emission tomography study. *Journal of Neuroscience, 17,* 391–400.

Celone, K. A., Calhoun, V. D., Dickerson, B. C., Atri, A., Chua, E. F., Miller, S. L., et al. (2006). Alterations in memory networks in mild cognitive impairment and Alzheimer's disease: An independent component analysis. *Journal of Neuroscience, 26,* 10222–10231.

Chaffin, R., & Herrmann, D. J. (1983). Self reports of memory abilities by old and young adults. *Human Learning, 2,* 17–28.

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

Chua, E. F., Schacter, D. L., Rand-Giovannetti, E., & Sperling, R. A. (2006). Understanding metamemory: Neural correlates of the cognitive process and subjective level of confidence in recognition memory. *NeuroImage, 29,* 1150–1160.

Cohen, J. D., Botvinick, M., & Carter, C. S. (2000). Anterior cingulate and prefrontal cortex: Who's in control? *Nature Neuroscience, 3,* 421–423.

Darby, D. (1993). MacStim (Version 2.5) [Computer software]. West Melbourne, Victoria, Australia: White Ant Occasional Publishing.

Daselaar, S. M., Fleck, M. S., Dobbins, I. G., Madden, D. J., & Cabeza, R. (2006). Effects of healthy aging on hippocampal and rhinal memory functions: An event-related fMRI study. *Cerebral Cortex, 16,* 1771–1782.

Daselaar, S. M., Veltman, D. J., Rombouts, S. A. R. B., Raaijmakers, J. G. W., & Jonker, C. (2003). Neuroanatomical correlates of episodic encoding and retrieval in young and elderly subjects. *Brain, 126,* 43–56.

Dennis, N. A., Kim, H., & Cabeza, R. (2008). Age-related differences in brain activity during true and false memory retrieval. *Journal of Cognitive Neuroscience, 20,* 1390–1402.

Dixon, R. A., & Hultsch, D. F. (1983). Structure and development of metamemory in adulthood. *Journal of Gerontology, 38,* 682–688.

Dodson, C. S., Bawa, S., & Krueger, L. E. (2007). Aging, metamemory, and high-confidence errors: A misrecollection account. *Psychology and Aging, 22,* 122–133.

Dodson, C. S., Bawa, S., & Slotnick, S. D. (2007). Aging, source memory, and misrecollections. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 169–181.

Dodson, C. S., & Krueger, L. E. (2006). I misremember it well: Why older adults are unreliable eyewitnesses. *Psychonomic Bulletin & Review, 13,* 770–775.

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

Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences, 23,* 475–483.

Dunlosky, J., & Hertzog, C. (2000). Updating knowledge about encoding strategies: A componential analysis of learning about strategy effectiveness from task experience. *Psychology and Aging, 15,* 462–474.

Duverne, S., Habibi, A., & Rugg, M. D. (in press). Regional specificity of age effects on the neural correlates of episodic retrieval. *Neurobiology of Aging.*

Eldridge, L. L., Knowlton, B. J., Furmanski, C. S., Bookheimer, S. Y., & Engel, S. A. (2000). Remembering episodes: A selective role for the hippocampus during retrieval. *Nature Neuroscience, 3,* 1149–1152.

Fincham, J. M., Carter, C. S., van Veen, V., Stenger, V. A., & Anderson, J. R. (2002). Neural mechanisms of planning: A computational analysis using event-related fMRI. *Proceedings of the National Academy of Sciences of the USA, 99,* 3346–3351.

Frith, C. D., & Frith, U. (1999, November 26). Interacting minds—A biological basis. *Science, 286,* 1692–1695.

Gilewski, M. J., Zelinski, E. M., & Schaie, K. W. (1990). The Memory Functioning Questionnaire for assessment of memory complaints in adulthood and old age. *Psychology and Aging, 5,* 482–490.

Grady, C. L. (1996). Age-related changes in cortical blood flow activation during perception and memory. *Annals of the New York Academy of Sciences, 777,* 14–21.

Grady, C. L., & Craik, F. I. (2000). Changes in memory processing with age. *Current Opinions in Neurobiology, 10,* 224–231.

Grady, C. L., Springer, M. V., Hongwanishkul, D., McIntosh, A. R., & Winocur, G. (2006). Age-related changes in brain activity across the adult lifespan. *Journal of Cognitive Neuroscience, 18,* 227–241.

Greicius, M. D., Srivastava, G., Reiss, A. L., & Menon, V. (2004). Default-mode network activity distinguishes Alzheimer's disease from healthy aging: Evidence from functional MRI. *Proceedings of the National Academy of Sciences of the USA, 101,* 4637–4642.

Gusnard, D. A., & Raichle, M. E. (2001). Searching for a baseline: Functional imaging and the resting human brain. *Nature Reviews Neuroscience, 2,* 685–694.

Gutchess, A. H., Kensinger, E. A., & Schacter, D. L. (2007). Aging, self-referencing, and the medial prefrontal cortex. *Social Neuroscience, 2,* 117–133.

Hazlett, E. A., Buchsbaum, M. S., Mohs, R. C., Spiegel-Cohen, J., Wei, T. C., Azueta, R., et al. (1998). Age-related shift in brain region activity during successful memory performance. *Neurobiology of Aging, 19,* 437–445.

Henson, R. N., Rugg, M. D., Shallice, T., & Dolan, R. J. (2000). Confidence in recognition memory for words: Dissociating right prefrontal roles in episodic retrieval. *Journal of Cognitive Neuroscience, 12,* 913–923.

Herholz, K., Salmon, E., Perani, D., Baron, J. C., Holthoff, V., Frolich, L., et al. (2002). Discrimination between Alzheimer dementia and controls by automated analysis of multicenter FDG PET. *NeuroImage, 17,* 302–316.

Huettel, S. A., Song, A. W., & McCarthy, G. (2005). Decisions under uncertainty: Probabilistic context influences activation of prefrontal and parietal cortices. *Journal of Neuroscience, 25,* 3304–3311.

Hultsch, D. F., Hertzog, C., & Dixon, R. A. (1987). Age differences in metamemory: Resolving the inconsistencies. *Canadian Journal of Psychology, 41,* 193–208.

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.

Johnson, S. C., Baxter, L. C., Wilder, L. S., Pipe, J. G., Heiserman, J. E., & Prigatano, G. P. (2002). Neural correlates of self-reflection. *Brain, 125,* 1808–1814.

Karpel, M. E., Hoyer, W. J., & Toglia, M. P. (2001). Accuracy and qualities of real and suggested memories: Nonspecific age differences. *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences, 56,* P103–110.

Kelley, C. M., & Jacoby, L. L. (1996). Adult egocentrism: Subjective experience versus analytic bases for judgment. *Journal of Memory and Language, 35,* 157–175.

Kelley, C. M., & Lindsay, D. S. (1993). Remembering mistaken for knowing: Ease of retrieval as a basis for confidence in answers to general knowledge questions. *Journal of Memory and Language, 32,* 1–24.

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

Kelley, W. M., Macrae, C. N., Wyland, C. L., Caglar, S., Inati, S., & Heatherton, T. F. (2002). Finding the self? An event-related fMRI study. *Journal of Cognitive Neuroscience, 14,* 785–794.

Kensinger, E. A., & Schacter, D. L. (2005). Emotional content and reality-monitoring ability: fMRI evidence for the influences of encoding processes. *Neuropsychologia, 43,* 1429–1443.

Kensinger, E. A., & Schacter, D. L. (2006). Reality monitoring and memory distortion: Effects of negative, arousing content. *Memory & Cognition, 34,* 251–260.

Kerns, J. G., Cohen, J. D., MacDonald, A. W., 3rd, Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004, February 13). Anterior cingulate conflict monitoring and adjustments in control. *Science, 303,* 1023–1026.

Kim, H., & Cabeza, R. (2007). Trusting our memories: Dissociating the neural correlates of confidence in veridical versus illusory memories. *Journal of Neuroscience, 27,* 12190–12197.

Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review, 103,* 490–517.

Koutstaal, W., & Schacter, D. L. (1997). Gist-based false recognition of pictures in older and younger adults. *Journal of Memory and Language, 37,* 555–583.

Krause, J. B., Taylor, J. G., Schmidt, D., Hautzel, H., Mottaghy, F. M., &

Muller-Gartner, H. W. (2000). Imaging and neural modelling in episodic and working memory processes. *Neural Networks, 13,* 847–859.

Levine, B., Turner, G. R., Tisserand, D., Hevenor, S. J., Graham, S. J., & McIntosh, A. R. (2004). The functional neuroanatomy of episodic and semantic autobiographical remembering: A prospective functional MRI study. *Journal of Cognitive Neuroscience, 16,* 1633–1646.

Liu, Y., Su, Y., Xu, G., & Chan, R. C. (2007). Two dissociable aspects of feeling-of-knowing: Knowing that you know and knowing that you do not know. *Quarterly Journal of Experimental Psychology, 60,* 672–680.

Lustig, C., Snyder, A. Z., Bhakta, M., O'Brien, K. C., McAvoy, M., Raichle, M. E., et al. (2003). Functional deactivations: Change with age and dementia of the Alzheimer type. *Proceedings of the National Academy of Sciences of the USA, 100,* 14504–14509.

MacDonald, A. W., III, Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000, June 9). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science, 288,* 1835–1838.

Magnussen, S., Andersson, J., Cornoldi, C., De Beni, R., Endestad, T., Goodman, G. S., et al. (2006). What people believe about memory. *Memory, 14,* 595–613.

Maril, A., Wagner, A. D., & Schacter, D. L. (2001). On the tip of the tongue: An event-related fMRI study of semantic retrieval failure and cognitive conflict. *Neuron, 31,* 653–660.

Matvey, G., Dunlosky, J., Shaw, R. J., Parks, C., & Hertzog, C. (2002). Age-related equivalence and deficit in knowledge updating of cue effectiveness. *Psychology and Aging, 17,* 589–597.

McDonald-Miszczak, L., Hertzog, C., & Hultsch, D. F. (1995). Stability and accuracy of metamemory in adulthood and aging: A longitudinal analysis. *Psychology and Aging, 10,* 553–564.

McFarland, C., Ross, M., & Giltrow, M. (1992). Biased recollections in older adults: The role of implicit theories of aging. *Journal of Personality and Social Psychology, 62,* 837–850.

McKiernan, K. A., Kaufman, J. N., Kucera-Thompson, J., & Binder, J. R. (2003). A parametric manipulation of factors affecting task-induced deactivation in functional neuroimaging. *Journal of Cognitive Neuroscience, 15,* 394–408.

Miller, S. L., Celone, K., DePeau, K., Diamond, E., Dickerson, B. C., Rentz, D., et al. (2008). Age-related memory impairment associated with loss of parietal deactivation but preserved hippocampal activation. *Proceedings of the National Academy of Sciences of the USA, 105,* 2181–2186.

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

Moritz, S., Glascher, J., Sommer, T., Buchel, C., & Braus, D. F. (2006). Neural correlates of memory confidence. *NeuroImage, 33,* 1188–1193.

Morris, J. C. (1993). The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology, 43,* 2412–2414.

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.

Nelson, T. O. (1996). Gamma is a measure of the accuracy of predicting performance on one item relative to another item, not of the absolute performance on an individual item. *Applied Cognitive Psychology, 10,* 257–260.

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

OptSeq [Computer software]. (n.d.). Retrieved February 2004 from http://surfer.nmr.mgh.harvard.edu/optseq/

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

Paulus, M. P., Hozack, N., Zauscher, B., McDowell, J. E., Frank, L., Brown, G. G., et al. (2001). Prefrontal, parietal, and temporal cortex networks underlie decision-making in the presence of uncertainty. *NeuroImage, 13,* 91–100.

Pliske, R. M., & Mutter, S. A. (1996). Age differences in the accuracy of confidence judgments. *Experimental Aging Research, 22,* 199–216.

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences of the USA, 98,* 676–682.

Satterthwaite, T. D., Green, L., Myerson, J., Parker, J., Ramaratnam, M., & Buckner, R. L. (2007). Dissociable but inter-related systems of cognitive control and reward during decision making: Evidence from pupillometry and event-related fMRI. *NeuroImage, 37,* 1017–1031.

Schacter, D. L., Koutstaal, W., & Norman, K. A. (1997). False memories and aging. *Trends in Cognitive Sciences, 1,* 229–236.

Schacter, D. L., Savage, C. R., Alpert, N. M., Rauch, S. L., & Albert, M. S. (1996). The role of hippocampus and frontal cortex in age-related memory changes: A PET study. *Neuroreport, 7,* 1165–1169.

Schraw, G. (1995). Measures of feeling-of-knowing accuracy: A new look at an old problem. *Applied Cognitive Psychology, 9,* 321–332.

Sharot, T., Delgado, M. R., & Phelps, E. A. (2004). How emotion enhances the feeling of remembering. *Nature Neuroscience, 7,* 1376–1380. doi:10.1038/nn1353.

Shulman, G. L., Fiez, J. A., Corbetta, M., Buckner, R. L., Meizin, F. M., Raichle, M. E., et al. (1997). Common blood flow changes across visual tasks: II. Decreases in cerebral cortex. *Journal of Cognitive Neuroscience, 9,* 648–663.

SPM2 [Computer software]. (2003). London: Wellcome Trust Centre for Neuroimaging.

Stark, C. E., & Squire, L. R. (2001). When zero is not zero: The problem of ambiguous baseline conditions in fMRI. *Proceedings of the National Academy of Sciences of the USA, 98,* 12760–12766.

Stretch, V., & Wixted, J. T. (1998). Decision rules for recognition memory confidence judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 1397–1410.

Volz, K. G., Schubotz, R. I., & von Cramon, D. Y. (2004). Why am I unsure? Internal and external attributions of uncertainty dissociated by fMRI. *NeuroImage, 21,* 848–857.

Volz, K. G., Schubotz, R. I., & von Cramon, D. Y. (2005). Variants of uncertainty in decision-making and their neural correlates. *Brain Research Bulletin, 67,* 403–412.

Wells, G. L., Olson, E. A., & Charman, S. D. (2003). Distorted retrospective eyewitness reports as functions of feedback and delay. *Journal of Experimental Psychology: Applied, 9,* 42–52.

Wheeler, M. E., & Buckner, R. L. (2003). Functional dissociation among components of remembering: Control, perceived oldness, and content. *Journal of Neuroscience, 23,* 3869–3880.

Wheeler, M. E., & Buckner, R. L. (2004). Functional–anatomic correlates of remembering and knowing. *NeuroImage, 21,* 1337–1349.

Yonelinas, A. P., Otten, L. J., Shaw, K. N., & Rugg, M. D. (2005). Separating the brain regions involved in recollection and familiarity in recognition memory. *Journal of Neuroscience, 25,* 3002–3008.

Zacks, R. T., Hasher, L., & Li, K. Z. H. (2000). Human memory. In F. I. Craik & T. A. Salthouse (Eds.), *Handbook of aging and cognition* (pp. 1–90). Mahwah, NJ: Erlbaum.