Published in final edited form as:
Neuropsychologia. 2012 July ; 50(9): 2356–2364. doi:10.1016/j.neuropsychologia.2012.06.005.

Alzheimer's Disease Can Spare Local Metacognition Despite Global Anosognosia: Revisiting the Confidence-Accuracy Relationship in Episodic Memory

David A. Gallo[1,2], **Stefanie J. Cramer**[1], **Jessica T. Wong**[1], and **David A. Bennett**[3,4]

[1]University of Chicago, Department of Psychology

[2]University of Chicago, Center for Cognitive and Social Neuroscience

[3]Rush University Medical Center, Department of Neurological Sciences

[4]Rush Alzheimer's Disease Center, Chicago IL

Abstract

Alzheimer's Disease (AD) can impair metacognition in addition to more basic cognitive functions like memory. However, while global metacognitive inaccuracies are well documented (i.e., low deficit awareness, or anosognosia), the evidence is mixed regarding the effects of AD on local or task-based metacognitive judgments. Here we investigated local metacognition with respect to the confidence-accuracy relationship in episodic memory (i.e., metamemory). AD and control participants studied pictures of common objects and their verbal labels, and then took forced-choice picture recollection tests using the verbal labels as retrieval cues. We found that item-based confidence judgments discriminated between accurate and inaccurate recollection responses in both groups, implicating relatively spared metamemory in AD. By contrast, there was evidence for global metacognitive deficiencies, as AD participants underestimated the severity of their everyday problems compared to an informant's assessment. Within the AD group, individual differences in global metacognition were related to recollection accuracy, and global metacognition for everyday memory problems was related to task-based metacognitive accuracy. These findings suggest that AD can spare the confidence-accuracy relationship in recollection tasks, and that global and local metacognition measures tap overlapping neuropsychological processes.

Keywords

metamemory; amnesia; recollection; insight; monitoring

It is well established that the early or mild stages of Alzheimer's disease (AD) impair basic cognitive abilities, but less is known about the effects of AD on metacognition. With respect to basic cognitive abilities, mild AD impairs the ability to consciously recall or recognize recently presented information (episodic memory), and also tends to impair the ability to

Corresponding author: David A. Gallo, University of Chicago, Department of Psychology, 5848 S. University Ave. Chicago, IL 60637, dgallo@uchicago.edu, Phone: 773-834-3701, Fax: 773-702-0886.

flexibly control attention in cognitive tasks (executive functions). These are the most frequently studied aspects of cognition in early AD, and significant impairment in these basic cognitive abilities in older adults is the hallmark of AD (Storandt, 2008). In addition to these basic cognitive abilities, AD also can impair metacognition. Metacognition is knowledge about cognition (Flavell, 1979; Nelson & Narens, 1994), including the ability to think about one's own basic cognitive abilities (such as episodic memory or executive functions). It is important to understand how different aspects of metacognition might be affected by AD, because accurate insight into one's own cognitive decline is needed to realistically adjust one's personal goals and avoid risky behaviors (e.g., Starkstein et al., 2007).

The research literature on metacognition in AD can be broadly divided into two different areas of inquiry, depending on the scope of the metacognitive processes under consideration (for reviews, see Pannu & Kaszniak, 2005; Souchay, 2007). One research area is primarily concerned with an individual's awareness of their cognitive abilities and declines at a global level. By the global level, we mean general assessments or beliefs about cognitive abilities as they might apply to various activities of daily living (e.g., "In general, I believe my memory is failing me."). Global metacognitive assessments are thought to involve long-term memory representations that contain information about one's personal abilities or self-efficacy beliefs (Mograbi, Brown, & Morris, 2009). In contrast to research on global metacognition, other research is primarily concerned with an individual's awareness of their cognitive abilities at a more local or task-specific level (Hertzog & Dunlosky, 2011). By local metacognition, we mean the online monitoring of cognitive task performance, as when people are asked to assess their performance across different experimental trials using subjective judgments such as confidence ratings (e.g., "I am not very confident in my recall of this particular item on this memory test."). This distinction between global and local metacognition is based on the object of the metacognitive assessment (i.e., assessing one's general cognitive abilities versus one's task-specific cognitive abilities, respectively), but the memory representations that are used to make these different kinds of self-assessment might overlap. For example, global assessments of cognitive ability might be affected by recent experiences with cognitive tasks, and local assessments of task performance might be affected by general representations of self-efficacy.

Deficits in global metacognition are fairly well documented in AD. Many people with AD are thought to have some degree of anosognosia – a relative lack of awareness of the severity of their everyday cognitive declines – although the extent of awareness varies considerably across individuals (Kaszniak & Edmonds, 2010). Evidence for this kind of global metacognitive impairment has come from a variety of sources, including anecdotal reports, clinically structured interviews, and deviations between self-reports and informant-reports on questionnaires that probe everyday problems (Clare, 2004). This kind of metacognition has been investigated most extensively in the clinical research literature, because personal insight into overall cognitive impairments is likely to affect one's decisions and behaviors in a variety of real-life situations (e.g., avoiding risky driving, Cotrell & Wild, 1999).

The research literature is more mixed with respect to the possible effects of the early stages of AD on more local metacognitive abilities. Some cognitive tasks reveal impaired metacognition in AD relative to controls. For example, Schacter, McLachlan, Moscovitch and Tulving (1986) found that AD participants overestimated their predicted recall abilities prior to a word recall task, and Souchay et al. (2002) found that AD impaired feeling-of-knowing judgments on a word recognition memory task. Moreover, some studies have found that individual variability in local metacognitive measures was related to more global metacognitive measures in AD, suggesting that these two kinds of metacognitive measures

might tap overlapping processes or representations (Cosentino, Metcalfe, Butterfield & Stern, 2007; Schmitter-Edgecombe & Seelye, 2011). In contrast to these studies, several other cognitive tasks demonstrate intact local metacognition in AD participants relative to controls. For example, Bäckman and Lipinska (1993) found that AD did not impair feeling-of-knowing judgments on a general knowledge task, and Moulin, Perfect and Jones (2000) found that AD did not impair judgments-of-learning on a word recall task. AD participants also can accurately recalibrate their metacognitive ratings immediately after taking an episodic memory test (e.g., Schmitter-Edgecombe & Seeyle, 2011; Stewart, McGeown, Shanks & Venneri, 2010), suggesting relatively intact metacognitive performance once there is some degree of exposure to the task. These latter findings are inconsistent with the findings of impaired global metacognition in AD, and they suggest a discontinuity between local and global metacognitive measures.

There are several possible reasons for these discrepancies in the literature on local metacognition in AD, including the kinds of metacognitive judgments used and the aspects of cognition that were assessed by the task. Along these lines, Dodson, Spaniol, O'Connor, Deason, Ally and Budson (2011) recently suggested that local metacognitive measures obtained during episodic memory tasks might be particularly sensitive to AD impairments, because episodic memory decline in the early stages of AD may make monitoring performance on these kinds of tasks particularly challenging (also see Pannu & Kaszniak, 2005). More specifically, Dodson et al. (2011) argued that AD participants have difficulty making accurate confidence judgments on a memory test due to increased susceptibility to false recollections that drive high-confidence errors. As described in the next section, however, relatively few studies have investigated the confidence-accuracy relationship in AD and the results have been mixed.

Episodic Memory Confidence

To our knowledge only three studies have investigated the effects of mild AD on the confidence-accuracy relationship in episodic memory retrieval (i.e., metamemory). The first two studies did not find impaired metamemory in AD. Pappas, Sunderland, Weingartner, Vitiello, Martinson and Putnam (1992) found that AD participants and controls were equally accurate at using confidence judgments to assess their ability to recall the final word from studied sentences. A study by Moulin, James, Perfect and Jones (2003) also found an intact confidence-accuracy relationship in AD participants. In this study, episodic memory was assessed using a two-alternative forced-choice test (2AFC), in which a studied word was paired with a nonstudied word. Unlike the recall test used by Pappas et al. (1992), which was potentially sensitive to group differences in response bias, the 2AFC test eliminated this kind of bias, demonstrating that AD participants can have intact metamemory in a very different kind of testing situation.

In contrast to the two studies showing a preserved confidence-accuracy relationship in episodic memory in AD, Dodson et al. (2011) found impairment in AD participants relative to controls. In this task, participants first studied sentences spoken by a male or a female. On a subsequent test, they were visually presented with studied and nonstudied sentences and needed to make four judgments: an old-new recognition decision, a confidence judgment corresponding to this recognition decision, a male-female source memory decision (for items recognized as "old"), and a confidence judgment corresponding to this source memory decision. Dodson et al. (2011) found a reduced confidence-accuracy relationship in the AD group compared to the control group, for both recognition and source memory measures. This metamemory impairment was obtained when memory accuracy was lower in the AD group compared to controls (as in Pappas et al., 1992, and Moulin et al., 2003), and also when the task procedures were varied so that memory accuracy could be artificially matched

across groups. They also found that the source memory deficit in AD participants was most pronounced at the highest levels of confidence, suggesting that AD participants were especially prone to high-confidence false recollections.

The addition of the source memory decision in Dodson et al. (2011) is a significant advance in metamemory research, because unlike the recognition decision (as well as the tests used in Pappas et al., 1992, or Moulin et al., 2003), accuracy on this source test required the recollection of specific details from the study phase. The targeting of recollection is theoretically important, because many memory impairments in AD outside the laboratory are likely to involve the recollection of specific details for personally important or salient events (e.g., Budson, Simons, Waring, Sullivan, Hussion, & Schacter, 2007). Targeting recollection also avoids ambiguities inherent to recognition memory tasks, which can be influenced by vague feelings of "oldness" towards a retrieval cue (familiarity) in addition to the conscious retrieval of specific details (recollections) that are associated with that retrieval cue (i.e., the dual-process framework, see Yonelinas, 2002). Because AD impairs the recollection of specific details or associated information more than it impairs general feelings of familiarity towards test items (e.g., Budson, Dodson, Daffner & Schacter, 2005; Dalla Barba, 1997; Gallo, Sullivan, Daffner, Schacter & Budson, 2004), the failure to control for these different memory processes can complicate the interpretation of metacognitive judgments. For example, a greater reliance on familiarity in AD participants might impair the confidence-accuracy relationship, even if relying on recollection would otherwise allow more accurate metacognitive monitoring.

The results of Dodson et al. (2011) suggest that AD participants may have particular difficulty with metacognitive judgments when the task encourages participants to recollect specific details. In fact, a greater reliance on recollected details by the control participants in this task may have allowed Dodson et al. (2011) to detect a metamemory impairment in AD that was not found in the two earlier studies (Pappas et al., 1992; Moulin et al., 2003). Nevertheless, another difference is that the memory task used by Dodson et al. (2011) was more complicated than the tasks used in these prior studies, in that participants were required to switch between two memory decisions and two confidence decisions for each test item. These increased response demands may have contributed to the observed metamemory problems in AD, as opposed to a fundamental deficit in the monitoring of recollected information. Given the importance of recollection-based metamemory in AD, additional research is needed that targets the confidence-accuracy relationship for the recollection of specific information.

The Current Study

The experiment reported here investigated the confidence-accuracy relationship in mild AD participants, using a relatively simple task that explicitly required recollection-based responding. This task has recently been used to identify metacognitive impairments in cognitively normal older adults (Wong, Cramer & Gallo, in press). Participants studied objects presented as colored photos or as line drawings, and then took a picture recollection test followed by confidence judgments. This test used a two-alternative forced-choice format (2AFC), whereby two picture labels were presented as retrieval cues in each test pair, one associated with a colored photo and the other with a line drawing. Participants had to choose the test word that was studied with a colored picture, and familiarity could not be used as a basis for this discrimination because each word in the pair had been studied (see Manipulation Check section). Requiring participants to use recollection is important, because otherwise group differences in metamemory accuracy would be ambiguous, potentially owing either to recollection impairments in AD or to an overreliance on familiarity or semantic information (cf. Gallo, Shahid, et al., 2006). This test also avoided

differences between AD participants and controls in the setting of a yes/no response criterion when making the memory judgment (e.g., Budson, Wolk, Chong & Waring, 2006), which can complicate interpretations of confidence judgment accuracy.

We addressed two overarching questions with this study. The first question was whether mild AD would impair the ability to use confidence judgments to track recollection accuracy, due to recollection and/or monitoring impairments. One prediction was that the confidence-accuracy relationship would be impaired in AD participants compared to controls, and that this impairment would persist even when the groups were artificially matched on recollection accuracy. These outcomes would replicate the main findings reported by Dodson et al. (2011). Alternatively, AD might not exacerbate the metamemory declines associated with cognitive aging, which would be consistent with the studies by Pappas et al. (1992) and Moulin et al. (2003). The current study used a simpler measure of recollection than these prior studies, as well as conditions that were calibrated to the level of difficulty for AD participants, thereby providing a more direct test of the confidence-accuracy relationship in a recollection task in AD participants.

The second question that we addressed with this study was whether individual differences in local metacognition (as assessed by the confidence-accuracy relationship) would be correlated with individual differences in global metacognition (as assessed by an anosognosia questionnaire procedure). A positive relationship would suggest that these different metacognitive measures tap common processes or representations. Some researchers have found a relationship between global metacognition and episodic memory task performance in AD (Cosentino et al., 2007; Gallo, Chen, Wiseman, Schacter & Budson, 2007). Nevertheless, correlations between local and global metamemory measures across individuals are not always obtained in the neuropsychology literature (Pannu & Kaszniak, 2005). In the context of the confidence-accuracy relationship, the lack of a correlation between local and global measures of metamemory would suggest that these measures tap fundamentally different neuropsychological processes.

Method

Participants

The memory task included 18 AD participants (mean age = 76.2; SD = 5.9; mean education = 14.4 years, SD = 2.4; 12 females) and 18 cognitively normal controls approximately matched on age, sex, and education (mean age = 76.8 years; SD = 7.4; mean education = 15.4 years, SD = 3.3, 13 females). AD participants were clinically diagnosed with probable AD based on NINCDS-ADRDA criteria (McKhann, Drachman, Folstein et al., 1984), and as expected, scores on the Mini Mental State Examination (Folstein, Folstein, & McHugh, 1975) were significantly lower in the AD participants (mean = 23.2, SD = 3.0, range = 17–29; MMSE data were missing for one AD participant) compared to the matched controls (28.8, SD = 1.4, range = 27–30), $t(33) = 7.05$, SEM = .786, $p < .001$, $d = 2.14$. AD participants were primarily recruited from the Rush Alzheimer's Disease Center (RADC; see Bennett, Schneider, Aggarwal et al., 2006), with 2 AD participants recruited from the University of Chicago Memory Center. We also obtained ratings of each AD participant's global cognitive ability from personal informants that attended the testing session (9 spouses and 9 offspring, all adults). Exclusion criteria included comorbid neurodegenerative disease, stroke, severe head trauma, cerebrovascular disease, alcohol or drug abuse, untreated depression, poor or uncorrected vision, or if English was not a primary language. All participants (and caregivers, where appropriate) gave written informed consent and were paid.

Materials

For the metamemory task, the primary stimuli were 96 verbal labels (i.e., common object words, such as *bucket, dresser, tie*) and corresponding pictures. Pictures were presented as black and white line drawings or colored photos, with objects cropped for presentation. The line drawings were primarily from Snodgrass and Vanderwart (1980) and Szekely et al. (2005), and corresponding photos were collected from the Internet. All images were formatted to be about the same size on the computer screen. The stimuli were screened to avoid a large degree of overlap (e.g. “guitar” was included, but “banjo” was not).

To measure global metacognition we administered the anosognosia questionnaire-dementia (AQD, Migliorelli et al., 1995) at the end of the session. This scale contains 30 questions regarding everyday intellectual functions (e.g., remembering dates and conversations, orienting to new environments, balancing checkbooks) as well as emotional or behavioral problems (e.g., irritability, interest in hobbies, depressed feelings). Each AD participant and a personally knowledgeable informant estimated the frequency that the AD participant experienced problems in each domain on a range from 0 (never) to 3 (always). Under the assumption that the significant other can accurately estimate the participant’s deficits, the participant’s total score (across the 30 items) was subtracted from the significant other’s total score to provide an index of the participant’s awareness of their own deficits (with larger discrepancy scores indicating less awareness). We also had the matched control participants rate themselves using the AQD, to provide an additional comparison to the AD participants.

Design

Based on pilot testing, we used two versions of the same task to ensure that both groups would perform in an intermediate range on the recollection tests (i.e., avoiding ceiling effects in controls or floor effects in AD participants). This kind of task calibration is important for calculating metamemory measures and comparing them across groups (see Dodson et al., 2007; Wong et al., in press). For AD participants, the task was divided into six study/test cycles. Each study block contained 16 of the 96 stimuli (8 line drawings and 8 photographs, each with their verbal label) presented in a mixed order. Across participants, stimuli were rotated through the line drawing and photograph conditions, as well as the different study/test cycles. Each test block contained 8 test pairs for the items studied in the preceding study block, for a total of 48 test pairs across the 6 cycles. Each test pair included one label corresponding to a line drawing and one corresponding to a studied photograph, with the target label occurring in the first position for half the trials and the second position for the other half. Across all test blocks there were a total of 48 test pairs.

Control participants were presented with the same stimuli and task as AD participants, but to make the task more difficult they received all 96 items in one study block followed by an additional 96 filler items in a second study block (for a total of 192 studied items). These filler items were drawn from the same sources as the primary items, and were not distinguishable from the participant’s perspective. Immediately after this study phase participants were tested on all 192 items. In order to maximize the retention interval for the primary items, the filler items were always tested in the first test block and the primary items were always tested in the second test block. Other than these differences all task procedures were the same across the two groups. Only the data from the 96 primary items are reported here, thereby equating the stimuli and the number of observations per group that contributed to each metamemory measure.

Procedure

All stimuli were presented on the computer and the experimenter entered the responses in a self-paced fashion. At the start of the experiment, all participants first received a brief (6-item) study/test practice cycle, to familiarize them with the different kinds of stimuli and the task instructions. They also were required to summarize the instructions back to the experimenter to ensure understanding and answer any questions. In this way, participants knew the nature of the memory test prior to the start of the study phase. They then proceeded to the study phase. On each study trial, a picture appeared in the center of the computer screen with its verbal label presented underneath in a large black font. To ensure that they were attending to the stimuli, participants first read each label aloud. They next indicated whether the picture was a colored photograph or black and white line drawing, and finally they rated the picture for amount of realistic detail (1–3 scale, low-high). Memory was tested using a 2AFC format. Participants were presented with a pair of black words, side by side, one corresponding to a studied line drawing and the other to a studied photograph. Participants had to indicate which one of the two labels was studied as a colored picture, guessing if necessary. After making each memory decision, they were prompted to rate their confidence in their answer using a 5-point scale: “guess”, “low”, “medium”, “high”, “certain”. They were told that confidence is analogous to how much they were willing to bet that they were correct, and were encouraged to use the entire scale.

Similar to Wong et al. (in press), we calculated three kinds of measures to assess confidence judgment accuracy. The first measure was the Somer’s d correlation (symmetrical), which represents the extent that a person’s confidence judgments differentiated accurate from inaccurate 2AFC test trials (maximum accuracy = 1). Much like the Goodman-Kruskal gamma correlation often used in metamemory research (Gonzalez & Nelsons, 1996), this correlation is recommended for ordinal-scaled variables and we report it here because it was the measure used in Dodson et al. (2011). (Note that we found similar results using the gamma correlation.) The second measure was the calibration error score, which represents the absolute difference between (a) the average accuracy for each level of confidence and (b) the corresponding confidence judgment itself, which is assumed to reflect a subjective estimate of accuracy. To calculate this measure we transformed our 5-point confidence scale into evenly divided numerical estimates of accuracy, so that guessing = .50, low = .625, medium = .75, high = .875, certain = 1.0. Perfect calibration occurs when there is a match between actual accuracy and subjectively estimated accuracy (e.g., the items with a “medium” confidence rating have an average accuracy of .75), and larger calibration error scores denote worse metamemory. The final measure was a discrimination score, which represented the average confidence difference between correct and incorrect 2AFC trials. Greater metamemory accuracy should yield higher confidence for correct than incorrect responses.

Manipulation Check

Our procedure assumes that word cues associated with the two different kinds of studied pictures do not reliably differ in familiarity, so that the 2AFC discrimination would need to be based on picture recollection. To check this assumption, a group of 12 younger adults was presented with an identical study phase as the control group in the main experiment. Rather than using a 2AFC test, the studied words were intermixed with an equal number of nonstudied words for a standard yes/no recognition test followed by recollection and familiarity subjective judgments (cf. Tulving, 1985). Results confirmed that there were no differences in overall recognition rates for words associated with either color or line drawings (both means = .84), as well as associated “recollection” judgments (.73 and .71, respectively), “familiarity” judgments (.11 and .12) or familiarity estimates using the process-independence correction (i.e., .41 vs. .43, cf. Yonelinas, 2002), all p’s > .50. These

findings confirm that words associated with the two different kinds of pictures were matched in terms of overall memory strength and familiarity, so that accurate performance on the 2AFC picture test would require picture recollection.

Results

The results are presented in four sections. The first two sections present analyses of recollection accuracy obtained on the 2AFC judgments, and the overall distribution of confidence judgments. The third section reports the overall confidence assigned to correct and incorrect responses, as well as three summary measures of confidence judgment accuracy (Somers' d, calibration error scores, and confidence discrimination). For each of these measures we compared the two original groups ($n = 18$ per group), on which AD-related impairments in 2AFC recollection accuracy were observed, as well as groups that were artificially matched on 2AFC recollection accuracy ($n = 12$ per group, see next section). In the final section we report the results of the global metacognition measure (via the anosognosia questionnaire), as well as individual difference correlations between various measures. All analyses were considered significant at the conventional $p < .05$ (two tailed).

Recollection Accuracy

Analysis of the 2AFC judgments indicated reduced accuracy in AD participants (mean = .69, SEM = .03) compared to controls (.84, SEM = .03), t (34) = 3.56, SEM = .042, p = .001, d = 1.19, demonstrating the typical recollection deficit in AD. Nevertheless, the AD group performed significantly greater than chance (.50), t (17) = 5.98, SEM = .191, p < .001, d = 1.41, indicating that our simplified study/test procedure successfully allowed the AD participants to accurately recollect some of the studied pictures. We also created groups that were matched on recollection accuracy (n = 12 per group) by removing the 6 best performing controls (accuracy > .92) and the 6 worst performing AD participants (accuracy < .58). The resulting groups were matched on recollection accuracy (.78 and .77, respectively), and continued to be comparable on age, sex, and education. Analysis of these matched groups is a useful complement to our analysis of the original groups, because the matched groups avoid extreme accuracy scores (ceiling or floor) as well as group differences in guessing rates, either of which could affect metamemory measures. To anticipate, most of the metamemory comparisons yielded similar results with the original and matched groups.

Confidence Distributions

Table 1 shows the frequency that participants used each level of the confidence judgment, independent from response accuracy. Both groups gravitated towards the higher ends of the confidence scale, most likely because recollection-based responding tends to be associated with relatively high confidence (e.g., Tulving, 1985), and as discussed above, both groups were able to recollect the studied pictures in the current task. Because the proportions within each group are interdependent (i.e., sum to 1), we focused on the high-confidence bin to compare the groups. In the original groups, AD participants overall were less likely to make high-confidence responses than were controls (means = .25 and .55, respectively), t (34) = 3.10, SEM = .097, p = .004, d = 1.03, consistent with the recollection impairment in this group (cf. Moulin et al., 2003). In fact, when the groups were matched on recollection accuracy, the difference in high-confidence responses was minimized and no longer significant (.44 and .30), t (22) = 1.14, SEM = .119, p = .268, d = .46), and there was no group difference when collapsing across the two highest levels of confidence (both group means = .64). Overall these findings reveal a similar pattern of confidence judgment distributions across the groups, especially when matched on recollection accuracy, indicating that the AD participants and controls understood and used the confidence judgments in a similar way.

Confidence Judgment Accuracy

Figure 1 presents the average confidence judgments for correct and incorrect 2AFC responses. The first point to take from the figure is that the average confidence for correct responses was greater than incorrect responses in all groups (all p's < .05), implicating accurate metamemory at this level. The second point to take is that, overall, the original group of AD participants had marginally lower confidence compared to controls, F (1, 34) = 3.87, MSE = 1.12, p = .07, η_p^2 = .09, consistent with their reduction in recollection accuracy and again implicating accurate metamemory at this level. This group difference was eliminated when the groups were matched on recollection accuracy. The final point to take is that the average confidence judgment in each condition was higher than the midpoint of the scale (medium confidence = 3), even for incorrect responses. If it is assumed that higher-confidence responding was most likely to reflect the subjective experience of recollection (cf. Dodson et al., 2007), then this pattern suggests that participants had based many of their incorrect decisions on the retrieval of false recollections as opposed to low-confidence guessing.

The three summary measures of confidence judgment accuracy are presented in Table 2. A comparison of the two original groups revealed group differences in only one of these measures, providing little evidence for a robust metamemory impairment in AD. With respect to the Somers' d correlation, there was no difference between controls and AD participants (means = .18 and .19), t (34) < 1, even though both groups showed a correlation that was significantly greater than zero (both p's < .001). There also was no group difference in the discrimination score (means = 0.49 and 0.48), t (34) < 1, even though both groups showed significantly greater confidence judgments for correct relative to incorrect responses (described previously). In contrast to these two measures, calibration error scores were significantly greater in the AD group compared to the control group (.09 and .15), t (34) = 2.67, SEM = .023, p = .012, d = .90. This calibration difference potentially implicates a metamemory impairment in AD, but it must be interpreted with caution because this measure may be sensitive to artifacts caused by group differences in recollection accuracy. Specifically, because recollection was impaired in AD participants, encouraging them to use the full confidence scale (as is typically done) may have caused them to give artificially higher confidence judgments to a disproportionately larger number of weak memories. Analysis of the groups that were matched on recollection accuracy revealed no significant difference on any of the three measures (all t's < 1), again implicating a relative sparing of metamemory in AD.

As an additional way to characterize confidence judgments, we analyzed the proportion of test trials that were assigned to each confidence level, calculated separately for correct and incorrect responses (Table 3). Because the proportions within each group are inter-dependent (i.e., sum to 1), we again focus on the high-confidence bin in our analysis. For the original groups, correct responses were more likely to be assigned the highest level of confidence in controls than in AD participants (means = .58 and .28, respectively), t (34) = 3.13, SEM = .097, p = .004, d = 1.02. This pattern can be attributed to impaired recollection quality in AD, so that even if metamemory were intact, AD participants were less certain in their accurate recollections compared to controls. A similar pattern was observed for incorrect responses (.38 and .14), t (34) = 2.15, SEM = .111, p = .039, d = .71. Although AD participants made more overall errors, these errors were associated with less confidence in the AD group compared to controls, providing no evidence for an increase in high-confidence false recollections in AD. These same patterns were observed when comparing the two matched groups, but neither effect was statistically significant (both p's < .25).

As a final way to characterize confidence judgment accuracy we plotted each of the original group's 2AFC accuracy as a function of their confidence judgments (Figure 2). To increase

the number of responses at the different levels of confidence, especially at the low end of the scale, we collapsed the five levels of confidence into three bins ("guess/low", "medium/"high", and "certain"). As can be seen from the figure, the two groups differed primarily at the lowest confidence levels, where AD participants performed no better than chance (.50). As confidence level increased, both groups were more accurate and the group differences in recollection accuracy were minimized. A similar analysis of the matched groups revealed no significant group differences in any of the confidence bins (collapsed means = .52, .72, and .93), although there were fewer observations in that analysis due to the exclusion of participants. Overall, these patterns are opposite to those observed in source accuracy by Dodson et al. (2011), where the groups were most likely to diverge at the highest-confidence levels.

Global Metacognition

Results from the anosognosia questionnaire are presented in Figure 3, collapsing across the different intellectual and behavioral domains. According to this metacognitive measure, AD participants demonstrated some accurate insight into their cognitive decline, as their self-ratings of everyday problems were greater than the self-ratings of everyday problems reported by controls (means = 21.2 and 14.4), $t(33) = 2.03$, SEM = 3.34, $d = .68$, $p = .05$. Nevertheless, AD participants underestimated their own everyday problems compared to informant reports (30.1), $t(16) = 2.96$, SEM = 3.02, $p = .009$, $d = .79$, a finding that is typically interpreted as reduced or imperfect metacognitive insight (cf. McGeown et al., 2010). The mean difference between AD participant ratings and informant ratings (i.e., the anosognosia score) was 8.94 (SD = 12.46, range = −11 to 33). Most of these anosognosia scores did not reach the threshold suggested by Migliorelli et al. (1995) for clinical anosognosia (>32), likely because the AD participants in the current study tended to be in the mild AD range (cf. Gallo et al., 2007), but the variability nevertheless suggests different levels of awareness across individuals. In fact, although informant ratings were reliably associated with AD participant MMSE scores, $r(16) = -.55$, $p = .03$, the correlation between AD participant self-ratings and MMSE scores was not significant, $r(16) = -.28$, $p = .30$, suggesting that informants had more insight into the level of overall deficits in the AD participants than did the AD participants themselves.

We next analyzed the extent that anosognosia scores were related to recollection task performance across individuals. There was a reliable association between 2AFC recollection accuracy and anosognosia scores, $r(17) = -.55$, $p = .02$, suggesting that AD participants with greater recollection impairments had less global awareness of their deficits. This effect was specific to the anosognosia scores (i.e., the difference between AD participant self-ratings and informant ratings), as the correlations were weaker and nonsignificant when separately comparing recollection accuracy to informant ratings, $r(17) = -.35$, $p = .16$, or AD participant ratings $r(17) = .26$, $p = .32$, and the effect remained significant after statistically controlling for MMSE, $r(16) = -.53$. $p = .04$. The significant negative relationship between recollection accuracy and anosognosia scores conceptually replicates Gallo et al. (2007), and potentially implicates accurate episodic memory as a critical factor in deficit awareness or global metacognition (see Discussion). In contrast, we found no significant correlations between anosognosia scores and any of our three task-based metamemory measures (confidence discrimination, $r(17) = -.40$, $p = .11$, calibration error, $r(17) = .27$, $p = .30$, Somers' d, $r(17) = -.32$, $p = .21$), even though two of the metamemory measures correlated with recollection accuracy (calibration error, $r(18) = -.72$, $p < .001$, Somers' d, $r(18) = .44$, $p = .07$). Thus, anosognosia scores were significantly correlated with one aspect of task performance (recollection accuracy) but not with another (metamemory measures).

Whereas the aforementioned analyses of global metacognition collapsed across all of the items on the anosognosia questionnaire, in a second analysis we calculated separate

anosognosia scores for the questions within each of 4 subdomains. These domains included 5 items on memory (remembering dates, conversations, household items, appointments, and shopping lists), 5 items on comprehension (understanding conversations, newspaper articles, movies, writing notes or letters, communicating with people), 5 items on daily tasks (keeping belongings in order, handling money, doing hobbies, household chores, balancing accounts), and 8 items on emotional/behavioral problems (rigid decisions, egotistical tendencies, irritability, crying episodes, laughing inappropriately, interest in sex or hobbies, depressed feelings). Analysis of the anosognosia scores in each of these domains and our metamemory measures revealed two significant correlations: anosognosia in the memory domain was related to confidence discrimination, $r(17) = -.56$, $p = .019$, and it also was related to Sommers' d, $r(17) = -.50$, $p = .04$. Given our relatively small sample size, these correlations should only be considered exploratory, but the specific relationship between local metamemory measures and more global awareness of memory problems suggests that there may be a domain-specific link between local and global measures of metacognition.

Discussion

One of the primary findings from this study was that participants with mild AD were able to use confidence judgments to track the accuracy of their responses on a recollection test. Not only were our measures of metamemory significantly greater than chance in the AD group, but most of them did not differ significantly from those observed in a group of cognitively normal controls. These effects are consistent with the results of Pappas et al. (1992) and Moulin et al. (2003), and they extend these previous findings to a memory test that targeted specific recollections. Our results further indicate that metamemory is spared in early AD when the two groups differed on recollection accuracy as well as when the two groups were matched on recollection accuracy, thereby avoiding group differences in guessing rates and other issues that might otherwise complicate the interpretation of metamemory measures (see Dodson et al., 2011; Wong et al., in press). Demonstrating preserved metamemory in mild AD participants on a recollection test is theoretically important, because prior work suggests that the early stages of AD should impair metamemory on this kind of task (see Dodson et al., 2011; Pannu & Kaszniak, 2005).

In addition to showing an intact confidence-accuracy relationship in mild AD, our study provided little evidence that mild AD participants were more prone to high-confidence errors than controls, as was found in Dodson et al. (2011). It seems unlikely that these different results were due to differences in the participants, because our AD participants were similar to those of Dodson et al. (2011) in terms of average age (76.2 and 77.0), education (14.4 and 15.0), and MMSE (23.2 and 23.9), and we used similar exclusion criteria. A lack of statistical power to detect group differences in our study also seems unlikely, given that we had more AD participants than each of the conditions of Dodson et al. ($n = 18$ and $n = 12$, respectively), we had more test observations in the studied item condition that was most relevant to recollection performance (48 and 30, respectively), and we had sufficient power to detect above-chance metamemory discrimination in the AD group. In fact, although our sample of AD participants was relatively small, it was larger than any of the three previously published confidence-accuracy studies in AD.

Given these similarities with prior work, our ability to detect a significant confidence-accuracy relationship in AD participants was more likely to be due to differences in task procedures. Perhaps the most critical difference was in our use of a relatively simplified recollection test. Dodson et al. (2011) had participants make an old-new recognition decision (with confidence) immediately prior to each source recollection decision (with confidence). Given that AD is more likely to impair recollection than familiarity in old-new recognition (e.g., Dalla Barba, 1997), this procedure may have encouraged AD participants

to use familiarity on their initial confidence judgments, and this tendency may have contaminated the confidence judgments for the source decision (i.e., a task-switching deficit, cf. Hutchison, Balota & Duchek, 2010). Our study used a single 2AFC judgment (with confidence), and because each member of the 2AFC pair had been associated with a picture at study, our participants were specifically focused on picture recollection. Our study also used a verbally presented confidence scale (guess, low, medium, high, certain) that may have been easier for AD participants than the numerical scale (50, 60, 70, 80, 90, 100) used in Dodson et al. (2011). Either of these differences may have made our task easier for AD participants, and in fact, the overall levels of recollection accuracy in AD participants were greater in our study.

The other major difference between our memory task and the one used by Dodson et al. (2011) was in the to-be-recollected stimuli. In Dodson et al. (2011) participants needed to recollect the voice that had presented a statement, with the same two voices used for multiple statements. This kind of source memory may have been more sensitive to false recollection than the picture recollections tested in the current study, to the extent that voice information is relatively less distinctive across different items relative to the perceptual features that characterize different pictures. Given the large literature showing that false recognition is more likely for less distinctive information in older adults (Dodson et al., 2002; Gallo et al., 2007), this difference in materials may be critical when looking at the confidence-accuracy relationship. In fact, using a similar task as the present study, Wong et al. (in press) found that metamemory accuracy was enhanced when the to-be-recollected stimuli were more distinctive (i.e., pictures relative to word font colors). In addition, Wong et al. (in press) replicated Dodson et al. (2007) by finding an impaired confidence-accuracy relationship in cognitively normal older adults, but unlike Dodson et al. (2007), Wong et al. (in press) found little evidence that cognitively normal older adults were more susceptible to high-confidence false recognition errors than younger adults. This pattern further suggests that different metamemory tasks may differ in their propensity to detect group differences in high-confidence errors.

How did participants use confidence judgments to track recollection accuracy across items in our study? We found a graded relationship between confidence judgments and recollection accuracy across items (Figure 2, see also Dodson et al., 2011), consistent with the idea that the retrieval of recollected information could be described as a continuous process (Mickes, Wais & Wixted, 2009). For example, it is likely that a given test word could have cued the recollection of a variety of associated features from study (e.g., the various colors from a studied picture, its orientation, etc.). While it is unclear if each feature itself was recollected in a continuous or all-or-none fashion, participants may have used between-item variability in the strength and/or amount of these recollected features to assign a relative level of confidence. Because our task explicitly targeted recollection, it may have been better positioned to tap into this kind of recollection variability than more standard tests of recognition memory, in which case recollection tends to be associated with more accurate and confident responding compared to familiarity (cf. Yonelinas, 2002). From this feature-based perspective, even though AD participants had an overall recollection deficit compared to control participants, participants in both groups had variability in the quality and/or quantity of recollected features across items, and both groups were able to use this variability when assigning confidence judgments.

Despite the relative sparing of local metamemory accuracy in our AD group, we replicated the typical finding of imperfect global metacognition, as measured with self-report and informant questionnaires (i.e., the anosognosia measure). This global metacognitive measure correlated with recollection accuracy across AD individuals, and this relationship was maintained after controlling for overall cognitive decline (via MMSE). This effect is

consistent with the idea that memory tasks that place a high demand on monitoring specific recollections are particularly sensitive to the same kinds of memory monitoring processes that are involved in global metacognitive assessments (cf. Cosentino et al., 2007; Gallo et al., 2007). For example, according to Mograbi et al. (2009), the memory deficits in AD make it difficult to update and maintain an accurate sense of self in long-term memory, thereby leading to underestimates of global cognitive decline (see also Agnew & Morris, 1998; Stewart et al., 2010). The recollection accuracy measure used here may have been particularly sensitive to these kinds of memory dysfunction, including the associative memory processes of the hippocampus and the retrieval monitoring processes of the prefrontal cortex.

We also found a significant relationship between metamemory measures and the questions on anosognosia measure that were specific to everyday memory problems (e.g., remembering appointments or shopping lists). This finding is consistent with other studies that have linked more global measures of metacognition in AD to task-based measures of metamemory accuracy, including judgments-of-learning in episodic memory tasks (Cosentino et al., 2007) and post-retrieval estimates of recent performance (Schmitter-Edgecombe & Seelye, 2011). Our results further suggest that the relationship between local and global metacognition may be domain-specific, because only responses to questions about everyday memory problems were related to task-based metamemory accuracy. It seems likely that some of the AD participants that underestimated the severity of their global memory problems also used the confidence judgments in a suboptimal way, making them less accurate at differentiating between correct and incorrect recollection responses.

We found that global metacognition was not perfect in AD, but it is important to note that we did not measure global metacognition in our control participants. Thus, we do not know if the relationship between local and global metacognition also would have been obtained in cognitively normal older adults. It also should be reiterated that the anosognosia measure that we used assumes that each AD participant has a relatively accurate informant. Although this kind of measure is commonly used in clinical research, it is potentially susceptible to biased or inaccurate reporting from the informant. In order to more fully understand how global metacognition is affected across the aging population more generally, future studies should use larger and more comprehensive aging samples as well as measures of metacognition that can be objectively anchored.

Considered as a whole, we found that the confidence-accuracy relationship was not impaired in the early stages of AD, extending previous findings to a recollection memory task. These results indicate that AD participants have relatively preserved metacognitive monitoring abilities on episodic memory tests, at least when the task is relatively well constrained. We also found that individual differences in this local metacognitive assessment were related to more global assessments of cognitive ability in the AD group, indicating that these two kinds of metacognitive assessment tap overlapping representations or neuropsychological processes. This relationship suggests that it may be possible to harness the preserved local metacognitive abilities in AD to enhance the accuracy of more global metacognitive assessments. To the extent that metacognition guides one's behaviors and plans, enhancing metacognitive abilities in AD might mitigate the psychological and behavioral consequences of the disease.

Acknowledgments

We are grateful to Barbara Eubeler, Gina Freed, James Mastrianni, and Raj Shah for assistance recruiting and characterizing AD participants. We also thank Chad Dodson for the calibration error measure. This work was supported by National Institute of Aging grants AG30345 and P30AG10161.

References

Agnew SK, Morris RG. The heterogeneity of anosognosia for memory impairment in Alzheimer's disease: A review of the literature and a proposed model. Aging & Mental Health. 1998; 2:7–19.

Bäckman L, Lipinska B. Monitoring of general knowledge: Evidence for preservation in early Alzheimer's disease. Neuropsychologia. 1993; 31:335–345. [PubMed: 8502369]

Bennett DA, Schneider JA, Aggarwal NT, Arvanitakis Z, Shah RC, Kelly JF, Fox JH, Cochran EJ, Arends D, Treinkman AD, Wilson RS. Decision rules guiding the clinical diagnosis of Alzheimer's disease in two community-based cohort studies compared to standard practice in a clinic-based cohort study. Neuroepidemiology. 2006; 27:169–176. [PubMed: 17035694]

Budson AE, Dodson CS, Daffner KR, Schacter DL. Metacognition and false recognition in Alzheimer's disease: Further exploration of the distinctiveness heuristic. Neuropsychology. 2005; 19:253–258. [PubMed: 15769209]

Budson AE, Wolk DA, Chong H, Waring JD. Episodic memory in Alzheimer's disease: Separating response bias from discrimination. Neurop- schologia. 2006; 44:2222–2232.

Budson AE, Simons JS, Waring JD, Sullivan AL, Hussion T, Schacter DL. Memory for the September 11, 2001, terrorist attacks one year later in patients with Alzheimer's disease, patients with mild cognitive impairment, and healthy older adults. Cortex. 2007; 43:875–888. [PubMed: 17941346]

Clare L. Awareness in early-stage Alzheimer's disease: A review of methods and evidence. British Journal of Clinical Psychology. 2004; 43:177–196. [PubMed: 15169617]

Cosentino S, Metcalfe J, Butterfield B, Stern Y. Objective metamemory testing captures awareness of deficit in Alzheimer's disease. Cortex. 2007; 43:1004–1019. [PubMed: 17941356]

Cotrell V, Wild K. Longitudinal study of self-imposed driving restrictions and deficits in awareness in patients with Alzheimer disease. Alzheimer Disease and Associated Disorders. 1999; 13:151–156. [PubMed: 10485574]

Dalla Barba G. Recognition memory and recollective experience in Alzheimer's disease. Memory. 1997; 5:657–672. [PubMed: 9497906]

Dodson CS, Bawa S, Krueger LE. Aging, metamemory, and high-confidence errors: A misrecollection account. Psychology & Aging. 2007; 22:122–133. [PubMed: 17385989]

Dodson CS, Schacter DL. Aging and strategic retrieval processes: Reducing false memories with a distinctiveness heuristic. Psychology & Aging. 2002; 17:405–415. [PubMed: 12243382]

Dodson CS, Spaniol M, O'Connor MK, Deason RG, Ally BA, Budson AE. Alzheimer's disease and memory-monitoring impairment: Alzheimer's patients show a monitoring deficit that is greater than their accuracy deficit. Neuropsychologia. 2011; 49:2609–2618. [PubMed: 21620877]

Flavell JH. Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. American Psychologist. 1979; 34:906–911.

Folstein MF, Folstein SE, McHugh PR. "Mini-Mental State": a practical method for grading the mental state of patients for the clinician. Journal of Psychiatric Research. 1975; 12:189–198. [PubMed: 1202204]

Gallo DA, Chen JM, Wiseman AL, Schacter DL, Budson AE. Retrieval monitoring and anosognosia in Alzheimer's disease. Neuropsychology. 2007; 21:559–568. [PubMed: 17784804]

Gallo DA, Shahid KR, Olson MA, Solomon TM, Schacter DL, Budson AE. Overdependence on degraded gist memory in Alzheimer's disease. Neurospsychology. 2006; 20:625–632.

Gallo DA, Sullivan AL, Daffner KR, Schacter DL, Budson AE. Associative recognition in Alzheimer's disease: Evidence for impaired recall-to-reject. Neuropsychology. 2004; 18:556–563. [PubMed: 15291733]

Gonzalez R, Nelson TO. Measuring ordinal association in situations that contain tied scores. Psychological Bulletin. 1996; 119:159–165. [PubMed: 8559859]

Hertzog C, Dunlosky J. Metacognition in later adulthood: Spared monitoring can benefit older adults' self-regulation. Current Directions in Psychological Science. 2011; 20:167–173.

Hutchison KA, Balota DA, Duchek JM. The utility of Stroop task switching as a marker for early-stage Alzheimer's disease. Psychology and Aging. 2010; 25:545–559. [PubMed: 20853964]

Kaszniak, AW.; Edmonds, EC. Anosognosia and Alzheimer’s disease: Behavioral studies. In: Prigatano, GP., editor. The study of anosognosia. Oxford: Oxford University Press; 2010. p. 189-227.

McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer’s disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer’s Disease. Neurology. 1984; 34:939–944. [PubMed: 6610841]

Mickes L, Wais PE, Wixted JT. Recollection is a continuous process: Implications for dual-process theories of recognition memory. Psychological Science. 2009; 20:509–515. [PubMed: 19320859]

Migliorelli R, Teson A, Sabe L, Petracca G, Petracchi M, Leiguarda R, Starkstein SE. Anosognosia in Alzheimer’s disease: A study of associated factors. Journal of Neuropsychiatry and Clinical Neurosciences. 1995; 7:338–344. [PubMed: 7580194]

Mograbi DC, Brown RG, Morris RG. Anosognosia and Alzheimer’s disease: The petrified self. Consciousness & Cognition. 2009; 18:989–1003. [PubMed: 19683461]

Moulin CJA, Perfect TJ, Jones RW. Evidence for intact memory monitoring in Alzheimer’s disease: Metamemory sensitivity at encoding. Neuropsychologia. 2000; 38:1242–1250. [PubMed: 10865100]

Moulin CJA, James N, Perfect TJ, Jones RW. Knowing what you cannot recognise: Further evidence for intact metacognition in Alzheimer’s disease. Aging Neuropsychology and Cognition. 2003; 10:74–82.

Nelson, TO.; Narens, L. Why investigate metacognition?. In: Metcalfe, J.; Shimamura, AP., editors. Metacognition: Knowing about knowing. Cambridge: MIT Press; 1994.

Pannu JK, Kaszniak AW. Metamemory experiments in neurological populations: A review. Neuropsychology Review. 2005; 15:105–130. [PubMed: 16328731]

Pappas BA, Sunderland T, Weingartner HM, Vitiello B, Martinson H, Putnam K. Alzheimer’s disease and feeling of knowing for knowledge and episodic memory. Journal of Gerontology: Psychological Sciences. 1992; 47:159–164.

Schacter DL, McLachlan DR, Moscovitch M, Tulving E. Monitoring and recall performance by memory-disordered patients. Journal of Clinical and Experimental Neuropsychology. 1986; 8:130. (abstract).

Schmitter-Edgecombe M, Seelye AM. Predictions of verbal episodic memory in persons with Alzheimer’s disease. Journal of Clinical and Experimental Neuropsychology. 2011; 33:218–225. [PubMed: 20839133]

Snodgrass JG, Vanderwart M. A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. Journal of Experimental Psychology: Human Learning. 1980; 6:174–215.

Souchay C, Isingrini M, Gil R. Alzheimer’s disease and feeling-of-knowing in episodic memory. Neuropsychologia. 2002; 40:2386–2396. [PubMed: 12417467]

Souchay C. Metamemory in Alzheimer’s disease. Cortex. 2007; 43:987–1003. [PubMed: 17941355]

Starkstein SE, Jorge R, Mizahi R, Adrian J, Robinson RG. Insight and danger in Alzheimer’s disease. European Journal of Neurology. 2007; 14:455–460. [PubMed: 17388998]

Stewart G, McGeown WJ, Shanks MF, Venneri A. Anosognosia for memory impairments in Alzheimer’s disease. Acta Neuropsychiatrica. 2010; 22:180–187.

Storandt M. Cognitive deficits in the early stages of Alzheimer’s disease. Current Directions in Psychological Science. 2008; 17:198–202.

Szekely A, D’Amico S, Devescovi A, Federmeier K, Herron D, Iver G, Jacobsen T, Arevalo AL, Vargha A, Bates E. Timed action and object naming. Cortex. 2005; 41:7–25. [PubMed: 15633703]

Tulving E. Memory and consciousness. Canadian Psychology. 1985; 26:1–12.

Wong JT, Cramer SJ, Gallo DA. Age-related reduction of the confidence-accuracy relationship in episodic memory: Effects of recollection quality and retrieval monitoring. Psychology & Aging. in press.

Yonelinas AP. The nature of recollection and familiarity: A review of 30 years of research. Journal of Memory and Language. 2002; 46:441–517.

Highlights

- Compared Alzheimer's participants and matched controls on metacognitive measures
- At a local level, confidence judgments tracked recollection accuracy in both groups
- At a global level, Alzheimer's participants underestimated their cognitive decline
- Aspects of local and global metacognitive measures correlated across individuals
- These metacognitive measures may tap distinct yet overlapping mental processes

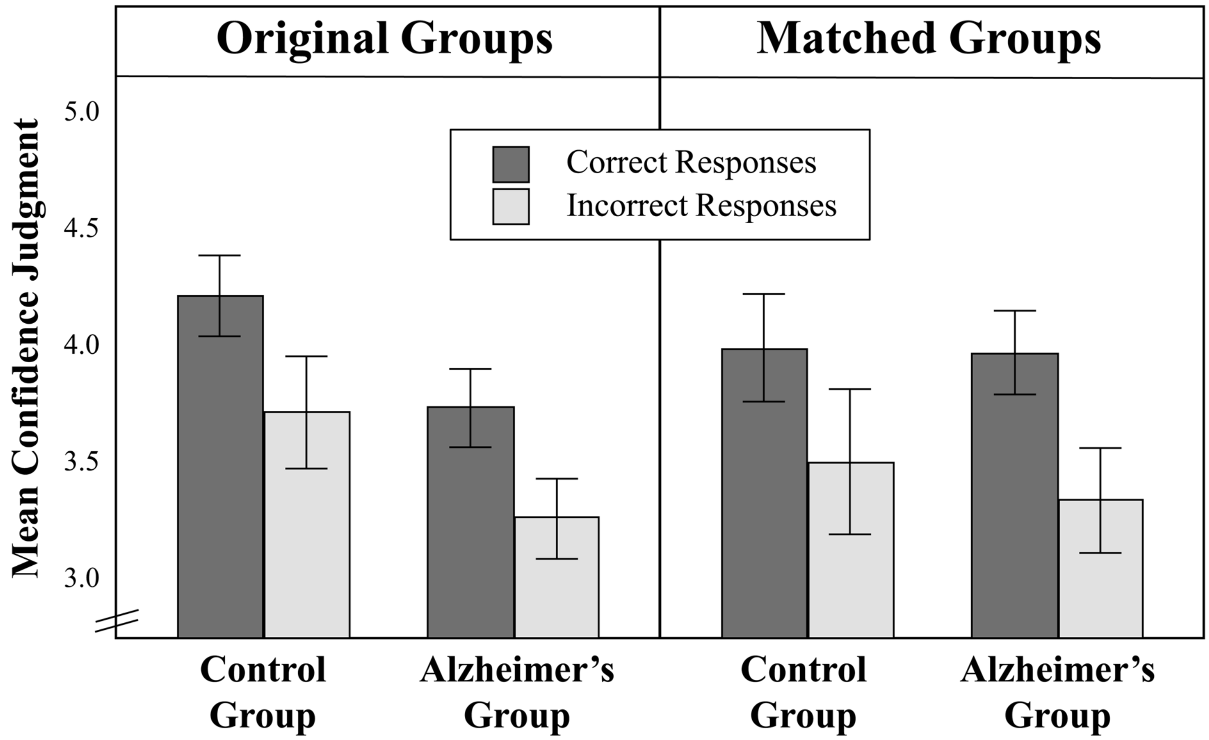

Figure 1.
Mean confidence judgments assigned to correct and incorrect responses in each participant group.

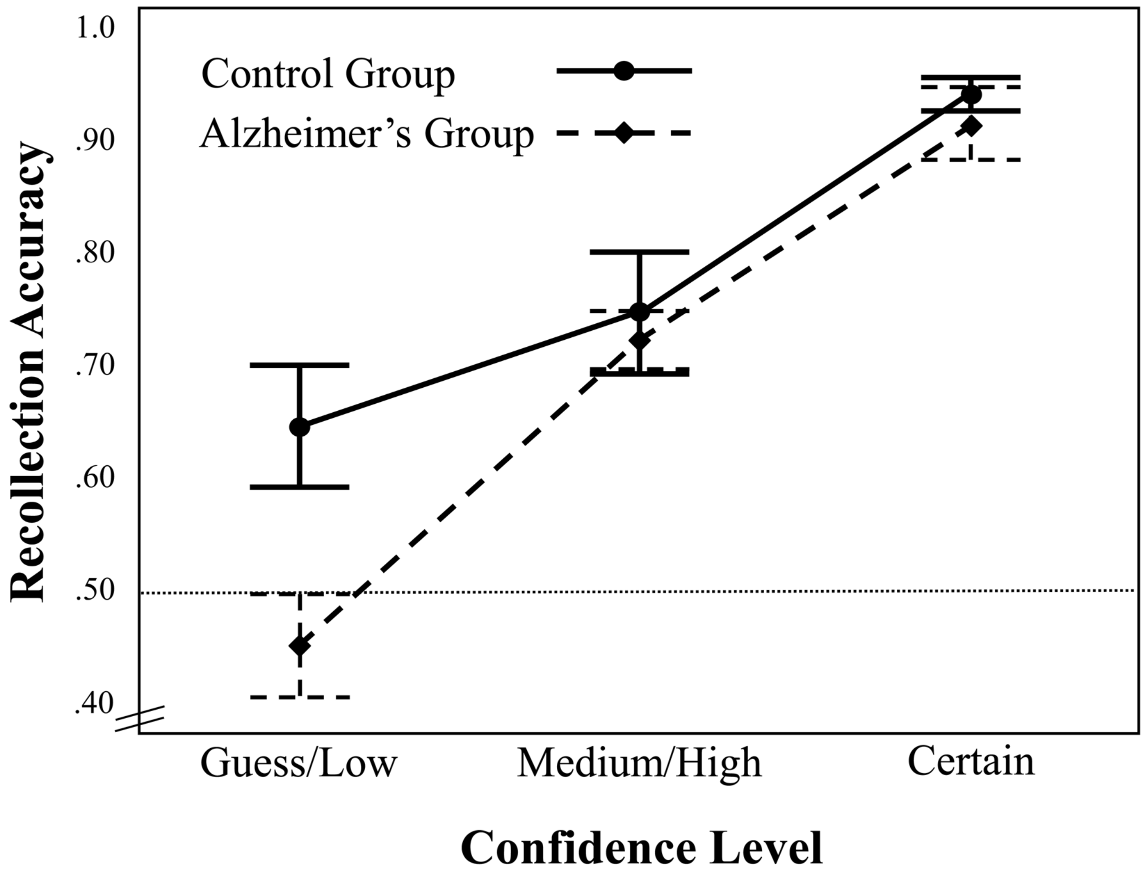

Figure 2.
Mean 2AFC recollection accuracy (and standard errors) for three confidence level bins in the original participant groups.

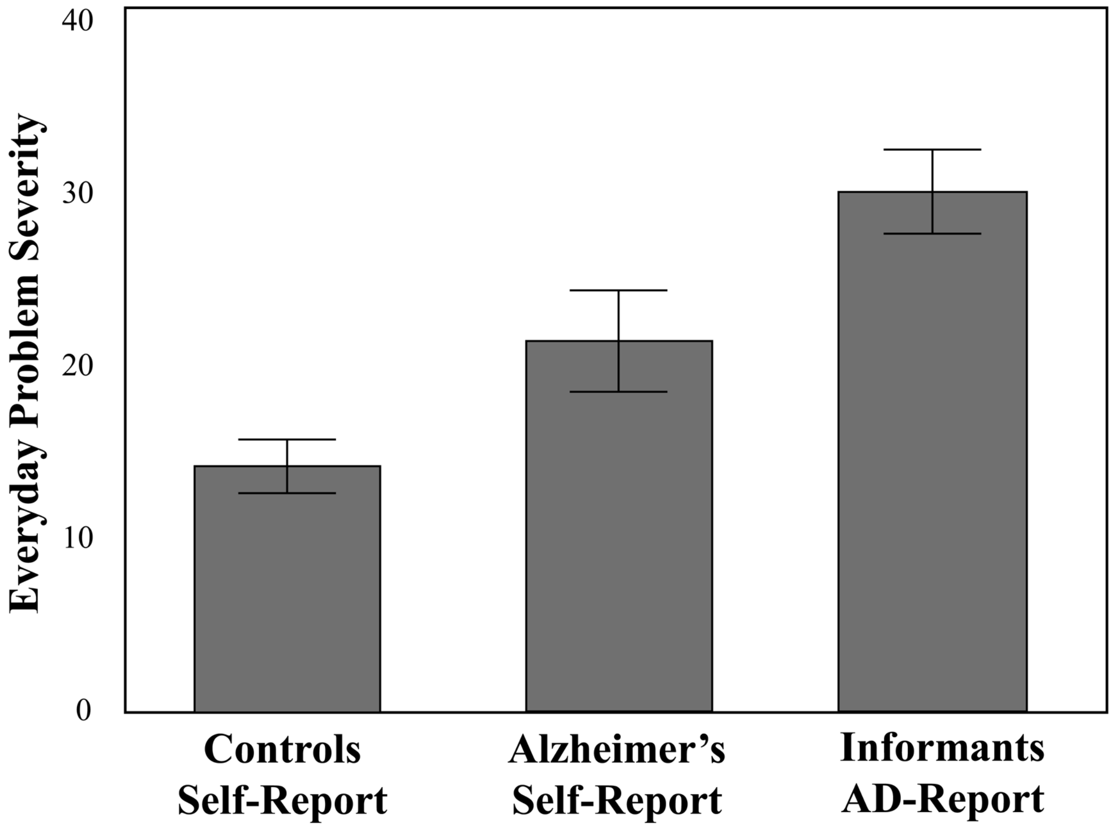

Figure 3.
Mean behavioral problems reported by the original groups of control participants (self-report), AD participants (self-report), and informants (reporting on the AD participants).

Table 1

Mean proportion of total test trials assigned to each of the confidence levels, independent from response accuracy, in each of the participant groups.

	Confidence Scale (1–5)				
	1 Guess	**2 Low**	**3 Medium**	**4 High**	**5 Certain**
Original Groups (n=18)					
Control Group	.03 (.01)	.10 (.04)	.14 (.03)	.18 (.04)	.55 (.07)
AD Group	.03 (.01)	.10 (.04)	.31 (.04)	.30 (.05)	.25 (.06)
Matched Groups (n=12)					
Control Group	.04 (.02)	.14 (.05)	.18 (.04)	.20 (.05)	.44 (.09)
AD Group	.01 (.01)	.08 (.05)	.27 (.05)	.34 (.05)	.30 (.07)

Note. Standard errors for each mean are in parentheses.

Table 2

Measures of confidence judgment accuracy for each of the participant groups.

	Somers' d Correlation	Calibration Error	Confidence Discrimination
Original Groups (n=18)			
Control Group	.18 (.04)	.09 (.01)	0.49 (.16)
AD Group	.19 (.04)	.15 (.02)	0.48 (.12)
Matched Groups (n=12)			
Control Group	.21 (.04)	.10 (.02)	0.48 (.18)
AD Group	.26 (.04)	.11 (.01)	0.62 (.16)

Note. Standard errors for each mean are in parentheses. Somer's d potentially ranged from −1 to +1, calibration error from 0 to .50, confidence discrimination from 0 to 4.

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 3

Mean proportion of test trials within each of the confidence levels, calculated separately for correct and incorrect responses, in each of the participant groups.

	Confidence Scale (1–5)				
	1 Guess	**2 Low**	**3 Medium**	**4 High**	**5 Certain**
Correct Responses					
Original Controls	.02 (.01)	.09 (.04)	.13 (.02)	.17 (.04)	.58 (.07)
Original AD	.02 (.01)	.09 (.03)	.30 (.05)	.31 (.05)	.28 (.07)
Matched Controls	.03 (.02)	.12 (.05)	.17 (.03)	.20 (.06)	.48 (.09)
Matched AD	.01 (.01)	.06 (.04)	.24 (.05)	.35 (.05)	.34 (.08)
Incorrect Responses					
Original Controls	.06 (.03)	.17 (.06)	.22 (.05)	.17 (.04)	.38 (.10)
Original AD	.06 (.02)	.14 (.05)	.41 (.06)	.25 (.05)	.14 (.06)
Matched Controls	.08 (.03)	.21 (.08)	.23 (.06)	.23 (.05)	.26 (.11)
Matched AD	.05 (.03)	.12 (.06)	.42 (.08)	.26 (.06)	.15 (.06)

Note. Standard errors for each mean are in parentheses.