Memory

Committing memory errors with high confidence: Older adults do but children don't

Yee Lee Shing [a]; Markus Werkle-Bergner [a]; Shu-Chen Li [a]; Ulman Lindenberger [a]
[a] Max Planck Institute for Human Development, Berlin, Germany

First Published:February2009

PLEASE SCROLL DOWN FOR ARTICLE

ψ **Psychology Press**
Taylor & Francis Group

Committing memory errors with high confidence: Older adults do but children don't

Yee Lee Shing, Markus Werkle-Bergner, Shu-Chen Li, and Ulman Lindenberger

Max Planck Institute for Human Development, Berlin, Germany

We investigated lifespan differences of confidence calibration in episodic memory, particularly the susceptibility to high-confidence errors within samples of children, teenagers, younger adults, and older adults. Using an associative recognition memory paradigm, we drew a direct link between older adults' associative deficit and high-confidence errors. We predicted that only older adults would show high-confidence error even though their memory performance was at a similar level to that of children. Participants of all ages showed higher confidence following correct responses compared to incorrect responses, demonstrating the ability to calibrate subjective confidence in relation to memory accuracy. However, older adults were disproportionately more likely to indicate high confidence following erroneously remembered word pairs than participants of the other three age groups. Results are discussed in relation to the misrecollection account of high-confidence errors and ageing-related decline in hippocampus-dependent episodic memory functions.

In everyday life we often find ourselves introspecting about how sure we are of our memories about certain issues or events. The ability to match (or "calibrate") subjective confidence with memory accuracy reflects the efficacy of one aspect of metacognitive memory monitoring (e.g., Brewer & Wells, 2006; Johnson, 2006; Loftus, 2003). Young adults typically provide higher confidence judgements after correct than after incorrect answers, indicating well-attuned calibration in matching accuracy and confidence (e.g., Roebers, 2002). In situations when young adults are "miscalibrated", overconfidence with respect to incorrect responses (henceforth termed *high-confidence error*) is a common phenomenon (e.g., Maki & Swett, 1987; Schneider & Laurion, 1993).

MEMORY ACCURACY AND CONFIDENCE CALIBRATION: TRENDS IN DEVELOPMENT AND AGEING

Researchers in the fields of child development (e.g., Pressley, Levin, Ghatala, & Ahmad, 1987; Roebers, 2002) and cognitive ageing (e.g., Jacoby & Rhodes, 2006) have noted that the ability to be well "calibrated" in judging the "goodness" of one's memory changes with age. An early study

(Pressley et al., 1987) showed that early-grade schoolchildren provided high-confidence ratings after both correct and incorrect responses, demonstrating a lack of differentiation according to accuracy. Children in later elementary school years can already differentially rate correct responses with higher confidence than incorrect ones (see also Roebers, von der Linden, Schneider, & Howie, 2007, with an event-recall paradigm). However, if prompted with misleading information, as is often done in false memory and eyewitness paradigms, even older children's confidence judgements are then less well calibrated than the judgements of younger adults, showing overconfidence with respect to incorrect responses (Roebers, 2002; Roebers & Howie, 2003). At the other end of the lifespan, older adults have been shown to experience greater difficulties in monitoring newly learned information than younger adults (e.g., Dodson & Krueger, 2006; Souchay, Isingrini, & Espagnet, 2000). Using an eyewitness misinformation paradigm, Dodson and Krueger (2006) found that older adults give high-confidence ratings to a greater proportion of falsely recognised items than younger adults.

Various potentially overlapping mechanisms have been suggested to account for the increase of false memory in ageing. Prominent accounts include compromised recollection (Jacoby, 1999), deficient inhibitory control (Jacoby & Rhodes, 2006), and impaired source monitoring (Henkel, Johnson, & DeLeonardis, 1998). Recently Dodson, Bawa, and Krueger (2007) found that high-confidence errors are particularly likely under conditions that require recollections of specific details of a memory episode. Specifically, this "misrecollection account" suggests that older adults' high-confidence errors may arise from their susceptibility to miscombine features from separate events, such that associations based on miscombined features become subjectively indistinguishable from associations based on correctly combined features (e.g., Chalfonte & Johnson, 1996; Naveh-Benjamin, Hussain, Guez, & Bar-On, 2003). This account is consistent with findings pointing towards pronounced age-associated declines in the hippocampus (e.g., Persson et al., 2006; Raz et al., 2005), which is important for episodic memory functioning. Age-associated impairments in this region have been linked to older adults' difficulties in forming new associations, and in separating new associations from memories stored in long-term memory (Dase-laar, Fleck, Dobbins, Madden, & Cabeza, 2006; Wilson, Gallagher, Eichenbaum, & Tanila, 2006).

AIMS AND OVERVIEW OF THE PRESENT STUDY

Thus far, child developmental and ageing research on memory accuracy and confidence calibration have been pursued separately and relied primarily on the paradigms of eyewitness misinformation (e.g., Ghetti, Qin, & Goodman, 2002; Roebers, 2002; Jacoby, Bishara, Hessels, & Toth, 2005) or source monitoring (e.g., Dodson et al., 2007). In these paradigms basic associative mechanisms of episodic memory (i.e., binding processes) are usually couched within various contextual details whose features could be manipulated to yield misinformation or shifts in memory sources. Under these conditions, younger children and older adults tend to overestimate their memory performance, particularly with respect to false memories. However, no study to date has directly compared children and older adults on memory-accuracy calibration and their susceptibility to high-confidence memory errors. To minimise age differences in the ability to process contextual details and to more directly relate memory-monitoring functions to basic associative mechanisms of memory, an associative recognition paradigm without addition of misinformation or source information was used.

According to the two-process account (Yonelinas, 2002), memory for past events can be based on retrieval accompanied by specific contextual details (recollection) or on the feeling that an event is old or new without recovering specific details (familiarity). Existing evidence suggests that recollection depends more on the hippocampus, whereas familiarity depends more on the rhinal cortex, and that healthy ageing has greater effects on recollection than on familiarity (e.g., Jacoby, 1999). Recent findings by Daselaar et al. (2006) suggest that older adults compensate for hippocampal (recollection) deficits by relying more on the rhinal cortex (familiarity), possibly guided by top-down frontal modulation. Based on these findings we predicted that older adults, due to their deficits in forming associations during encoding and their greater reliance on familiarity during retrieval, would be more likely to miscombine features originating from different familiar events into illusory memory with high confidence than younger adults. In contrast to

our prediction regarding older adults, we did not expect that 10- to 12-year-old children and adolescents would be particularly prone to commit memory errors with high confidence because the medial temporal brain regions (including the hippocampus and the rhinal cortex) operate relatively well by middle childhood (e.g., Giedd et al., 1999; Menon, Boyett-Anderson, & Reiss, 2005; Ofen et al., 2007; Sowell et al., 2003).

In addition to individual differences in the efficacy of associative mechanisms, strategic processes also affect episodic memory (see also Shing, Werkle-Bergner, Li, & Lindenberger, in press; Werkle-Bergner, Müller, Li, & Lindenberger, 2006). These processes may include the organisation and manipulation of the elements of a memory episode during encoding, storage, or retrieval (Craik & Lockhart, 1972; Levin, 1988; Paivio, 1971). For example, memory encoding can be aided by the use of mediators generated from verbal and imagery elaboration (Richardson, 1998). The child developmental literature suggests that it is not until the end of the elementary-school years that children make use of the full range of memory strategies (for a review, see Schneider & Pressley, 1997). On the other hand, the cognitive ageing literature suggests that older adults do not utilise memory strategies as efficiently as younger adults (for a review, see Kausler, 1994). Therefore in this study we also examined the extent to which high-confidence errors can be reduced by instructing an elaborative imagery strategy to the participants. Given recent evidence suggesting that older adults show more limited memory plasticity than children as a function of mnemonic training (Brehmer, Li, Müller, v. Oertzen, & Lindenberger, 2007), we expected that older adults would continue to show more high-confidence errors than individuals in other age groups even after strategy instruction.

METHOD

Participants

Our lifespan sample consisted of four age groups: 43 children (aged 10–12, $M = 11.2$, $SD = 0.6$), 43 teenagers (aged 13–15, $M = 14.4$, $SD = 0.4$), 42 younger adults (aged 20–25, $M = 23.3$, $SD = 1.6$), and 42 older adults (aged 70–75, $M = 73.2$, $SD = 1.7$). The age differences between the children and teenagers were chosen to reflect assumed

developmental differences in maturity of the prefrontal cortex (Giedd et al., 1999). All participants were residents of Berlin, Germany and travelled to our laboratory for testing on their own, or were accompanied by a parent in the case of younger children. The children and teenagers were attending the highest school track in Germany (the *Gymnasium*) that allows for university admission. The younger adults were mostly university students but none of them was a psychology student (to minimise pre-existing difference in knowledge of memory strategy). The older adults lived independently in the community. All participants reported having normal or corrected-to-normal visual and auditory acuity. Participants also filled out demographic questionnaires assessing subjective well-being and subject health condition, in which participants showed no significant age difference in these dimensions. Descriptive information of the participants is summarised in Table 1.

Design and procedure

We systematically varied demands on an associative recognition memory task (cf. Naveh-Benjamin, 2000) along the dimensions of associative demand and strategic elaboration in a fully crossed within-person repeated measures design. The task entailed presenting lists of word pairs for study and then testing either for the single words (item recognition) or for the associations between the words (pair recognition). Given the emphasis on binding as a candidate mechanism for high-confidence error, here we focused only on pair recognition.[1]

Associative-demand manipulation. Word pairs of either two unrelated German–German (GG) words or of one German word and a Malay word[2] (GM) were used. The Malay word was the direct translation of the German word. For our German participants the GM pairs demanded more associative processing than the GM pairs, as there is a lack of pre-existing knowledge of the unfamiliar Malay language.

---

[1] Other details of the study can be obtained by contacting the authors.
[2] Malay is written in the Latin alphabet and the phonemes are pronounceable for German speakers. None of the participants in our sample knew this language, thus it was equally unfamiliar for all age groups.

TABLE 1
Descriptive characteristics of sample

Measures	Children M (SD)	Teenager M (SD)	Younger adults M (SD)	Older adults M (SD)
Age	11.2 (0.6)	14.4 (0.4)	23.3 (1.6)	73.2 (1.7)
Male: Female percentage (%)	47: 53	51: 49	48: 52	50:50
Well-being	4.9 (2.3)	4.6 (2.4)	4.9 (2.7)	5.4 (2.9)
Subjective health	3.4 (0.7)	3.1 (0.6)	3.1 (0.8)	3.1 (0.7)
Digit symbol (cognitive mechanics)	45.7 (7.3)	57.3 (9.3)	62.8 (10.9)	43.4 (9.5)
Vocabulary (cognitive pragmatics)	9.4 (4.1)	13.5 (4.4)	20.6 (4.5)	27.9 (4.1)

Standard deviation in brackets. See reference for calculation of well-being score (total score 12; the lower the better) in Radloff (1977). Questionnaire of subjective health was constructed by the authors (scale from 1 to 5; the higher the better).

Strategic elaboration manipulation. The levels of strategic elaboration were manipulated using the encoding instruction. At pre-strategy memory assessment, participants were simply instructed to study the word pairs for an upcoming pair recognition test. This was followed by a strategy training session, in which participants were instructed and practised on an elaborative imagery strategy (Richardson, 1998). The essence of the strategy was to elaborate on the word pairs with visual imagery that dynamically integrated the two words (Paivio, 1969). For the German–Malay condition, a variant of the imagery strategy known as the keyword strategy was instructed. Participants were told to first find a meaningful connection (i.e., the keyword) for the Malay word through either its phonological or orthographic characteristics. Then participants integrated the keyword with the German word through imagery. In the strategy instruction session participants were introduced to the main principles of the strategy and intensively practised applying it. Detailed feedback was provided to improve the quality of the imageries. After the strategy-training session, participants' memory performance was assessed in two sessions of post-strategy memory assessment. On average, the time interval between testing sessions was between 2 and 4 days.

Pre- and post-strategy memory assessments. In a counterbalanced block design, 45 GG or GM word pairs were presented sequentially on the computer screen during the encoding phase. Presentation time was 3 seconds for younger adults and 6 seconds for the other age groups to avoid ceiling or floor performance in the different age groups. At the recognition phase, 60 memory probes were presented consecutively on the computer screen. Half of the pair probes were exact replication pairs from the encoding phase

(target pairs); 15 probes were rearranged pairs, composed of recombinations of words taken from different study pairs at encoding; 15 probes were new–new pairs, composed of words that never appeared at encoding. Participants were instructed to press the ''old'' response button for intact pairs and the ''new'' response button for both the rearranged and new–new pairs. The distinction between the two types of lure probes was based on the rationale that the rearranged pairs elicited a higher familiarity signal than the new–new pairs, which also required higher fidelity of the recollection process to avoid committing false alarm (FA) responses (Yonelinas, 2002). In the analysis we focused on confidence judgements following FAs on rearranged pairs.

Confidence rating. Following each old/new judgement during the recognition phase, participants made a confidence judgement of their decision on a 3-point scale ranging from 1 (unsure) to 3 (sure), without time restriction. To aid the children in understanding the scale, each of the response buttons corresponding to the three increasing levels of confidence was labelled by stickers with one, two, or three stars.

RESULTS

Overview of analyses

Confidence judgements of trials in which an old/new judgement was given in less than 400 ms were likely to be invalid or anticipatory responses and were discarded from analyses (less than 1% of overall data). Rates of hits (''old'' responses to target) and FAs (''old'' responses to rearranged lures) were computed respectively for each block types. On average, teenagers and younger adults produced much fewer FAs than children and older

adults. Whereas children showed comparable levels of FAs to older adults at pre-strategy assessment, they reduced the FA rate more than older adults after strategy instruction (see Table 2). For the two response types (hit or FA), percentages of "sure" judgements were calculated. For analyses of confidence rating with respect to these two response types, individuals who did not produce FAs logically yielded missing data and were excluded from the statistical analyses. Overall, this resulted in 7% and 9% of missing data in children and older adults for the GG condition, respectively, before strategy instruction, and 12% missing data in both age groups after strategy instruction. In teenagers and younger adults, the amount of missing data is larger (i.e., 10–43%). There were no missing data in the GM condition. Given that we did not have specific predictions regarding the two post-strategy sessions, we collapsed across the participants' performance on the two post-strategy sessions to reduce data complexity.

The percentages of "sure" judgements following hits and FAs (see Figure 1) were analysed with a $2 \times 2 \times 4$ (response type: hits vs FAs x assessment: pre- vs post-strategy $\times$ 4 age groups) mixed MANOVA, separately for the GG and GM conditions. We examined three main hypotheses: (1) all participants would show more "sure" judgements on hit than FA responses; (2) older adults would show more "sure" responses on FAs (i.e., high-confidence error) than other age groups, including children; (3) high-confidence errors in older adults would remain even after strategy instruction. For all effects we also calculated p_{rep} using the approximation of Killeen (2005) and partial eta squared (η^2). Significant age effects were followed up by post hoc all possible pairwise comparisons with Bonferroni adjustment.

GG condition

The results are shown in the upper panels of Figure 1. The $2 \times 2 \times 4$ omnibus test yielded a significant main effect of response type, $F(1, 109) = 275.35$, $p_{\text{rep}} > .99$, $\eta^2 = .72$, and significant interactions between response type and assessment, $F(1, 109) = 14.81$, $p_{\text{rep}} > .99$, $\eta^2 = .12$, as well as between response type and age, $F(3, 109) = 10.56$, $p_{\text{rep}} > .99$, $\eta^2 = .23$. The three-way interaction was not significant, $F < 1$. The main effect of response type was driven by a higher percentage of "sure" judgements following hit responses than FA responses ($M_{\text{hit}} = 80.40$ vs $M_{\text{FA}} = 35.40$). Therefore, as expected, participants in all age groups differentiated between correct and incorrect responses, demonstrating appropriate memory accuracy and confidence calibration in general. In addition, the interaction between response type and assessment showed that the difference in confidence following hit and FA responses was enlarged after strategy instruction ($\triangle_{\text{pre-strategy}} = 37.40$ vs $\triangle_{\text{post-strategy}} = 52.60$).

TABLE 2

Mean and standard error (in parentheses) of hit rates for target pairs and FA rates for rearranged lure pairs at pre- and post-strategy assessments

Age group	Hit rate		False alarm rate	
	Pre-strategy	Post-strategy	Pre-strategy	Post-strategy
German–German				
Children	.69 (.02)	.81 (.02)	.31 (.03)	.12 (.01)
Teenager	.80 (.02)	.86 (.02)	.22 (.02)	.08 (.01)
Younger adults	.82 (.02)	.88 (.01)	.16 (.03)	.06 (.01)
Older adults	.81 (.02)	.84 (.02)	.34 (.04)	.21 (.03)
German–Malay				
Children	.60 (.02)	.68 (.02)	.40 (.02)	.32 (.02)
Teenager	.68 (.03)	.73 (.02)	.38 (.02)	.29 (.02)
Younger adults	.77 (.02)	.79 (.01)	.45 (.03)	.27 (.02)
Older adults	.76 (.02)	.80 (.02)	.55 (.03)	.48 (.03)

In the GG condition children showed comparable FA rates to older adults at pre-strategy assessment, $t(76) = -.76$, $p = .45$, and lower FA rates compared to older adults after strategy instruction, $t(61) = -3.27$, $p = .002$. In the GM condition children and older adults differed in the FA rate both before, $t(74) = -3.48$, $p = .001$, and after strategy instruction, $t(77) = -4.66$, $p = .00$.

German - German Condition

German - Malay Condition
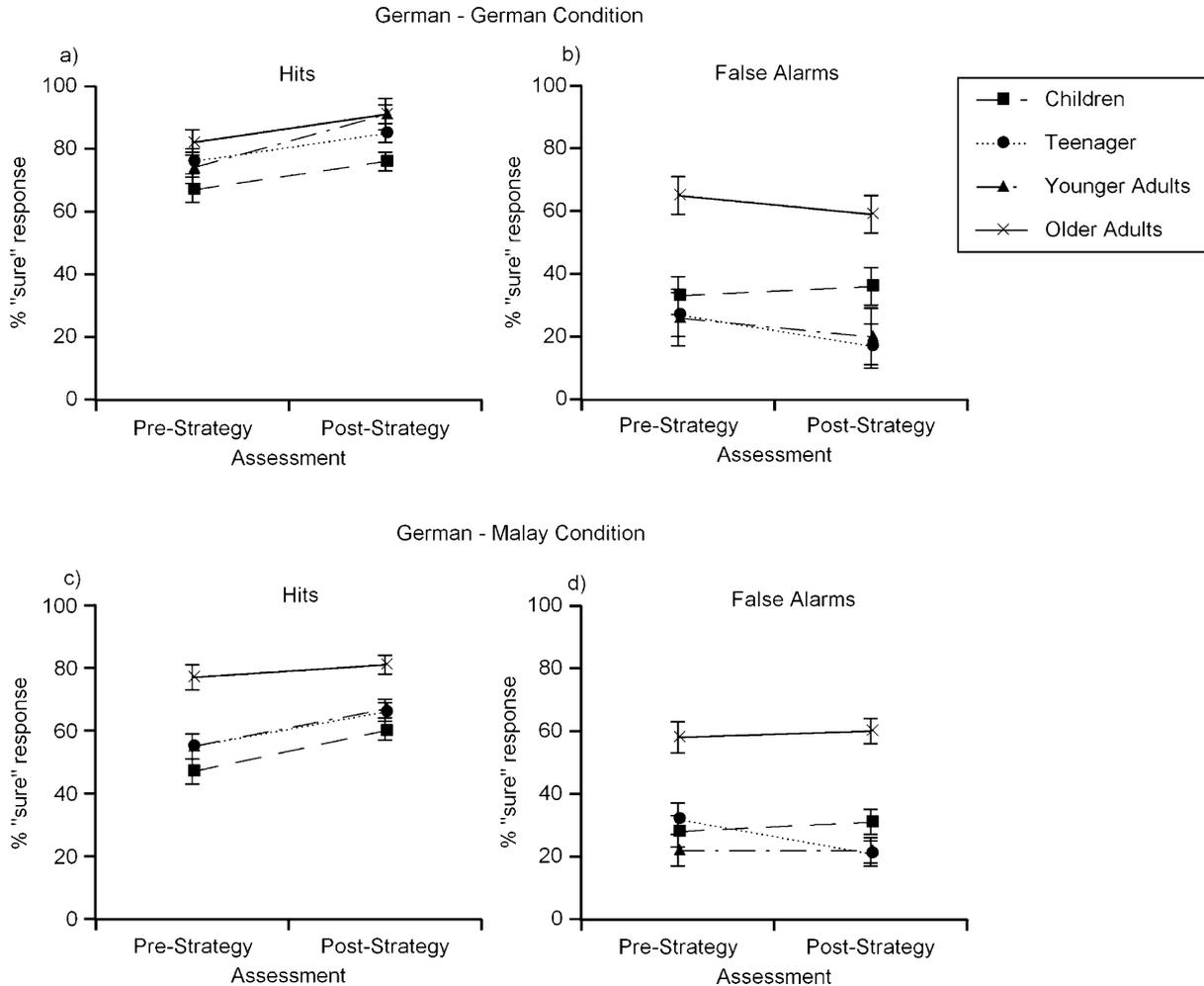
Figure 1.   Mean percentages of "sure" responses from pre- to post-strategy instruction as a function of age group and response type (i.e., hit or FA) in the GG and GM conditions.

To interpret the interaction between response type and age, we examined mean percentages of "sure" judgements (separately for hits or FAs) derived from the pre- and post-strategy assessments with one-way ANOVAs (age as factor). For both response types, there was a significant age effect—hit: $F(3, 166) = 6.99$, $p_{rep} > .99$, $\eta^2 = .11$; FA: $F(3, 154) = 8.19$, $p_{rep} > .99$, $\eta^2 = .14$. For confidence judgements following hits, post hoc comparisons showed that children were significantly lower in the percentage of "sure" judgements ($M_{children} = 73.77$) compared to the other age groups ($.93 < p_{rep} < .99$; $82.22 < M < 87.26$). There was no difference among the remaining three age groups. For confidence judgements associated with FAs, post hoc comparisons showed that older adults were significantly higher in the percentage of "sure" judgements ($M_{olderadults} = 56.83$) compared to the other age groups ($p_{rep} > .92$, $22.00 <$

$M < 35.17$), including children ($M_{children} = 35.17$, $p_{rep} = .97$). There was no difference among the remaining three age groups. Therefore our prediction that older adults should show more high-confidence errors, including the comparison to children, was supported. Furthermore, this effect was observable both at pre- and post-strategy instruction sessions as indicated by the non-significant three-way interaction.

GM condition

The results are shown in the bottom panel of Figure 1. The $2 \times 2 \times 4$ omnibus test yielded significant main effect of assessment, $F(1, 164) = 6.80$, $p_{rep} = .95$, $\eta^2 = .04$, and response type, $F(1, 164) = 418.17$, $p_{rep} > .99$, $\eta^2 = .72$; significant interactions between response type and

age, $F(3, 164) = 9.54$, $p_{rep} > .99$, $\eta^2 = .15$, as well as between assessment and response type, $F(1, 164) = 27.42$, $p_{rep} > .99$, $\eta^2 = .14$; and a significant three-way interaction, $F(3, 164) = 3.61$, $p_{rep} > .94$, $\eta^2 = .06$. To interpret the highest-order interaction, we examined the confidence judgements following hits or FAs separately using two-way ANOVAs (assessment by age group).

For confidence judgements following hit responses there were significant main effects of assessment, $F(1, 166) = 31.13$, $p_{rep} > .99$, $\eta^2 = .16$, and age group, $F(3, 166) = 12.47$, $p_{rep} > .99$, $\eta^2 = .18$. Specifically, participants showed higher percentages of "sure" responses following hit responses after strategy instruction ($M_{pre\text{-}strategy} = 58.65$, $M_{post\text{-}strategy} = 68.09$). Older adults, on average, showed significantly more "sure" responses ($M_{older\ adults} = 78.70$) compared to the other age groups ($p_{rep} > .99$, $53.10 < M < 61.10$). The remaining three age groups did not differ from each other. The interaction between assessment and age group was not significant, $F(3, 166) = 1.31$, $p_{rep} = .66$, indicating that the change from pre- to post-strategy assessment was equal across age group.

For confidence judgements following FA responses, the only significant effect was age, $F(3, 164) = 18.91$, $p_{rep} > .99$, $\eta^2 = .26$. Post hoc comparisons revealed that older adults, similar to the GG condition, showed significantly more high-confidence errors ($M_{older\ adults} = 59.00$) compared to the other age groups ($p_{rep} > .99$, $22.10 < M < 29.10$), including children ($M_{children} = 29.10$). The three remaining age groups did not differ from each other. Therefore our prediction that older adults should show more high-confidence error, including the comparison to children, was supported. Furthermore, this pattern did not diminish even after strategy instruction. We also examined whether older adults showed even more "sure" responses on FAs than other age groups in the GM condition. In a 2×4 (associative demand; GG vs GM $\times$ four age groups) ANOVA, we did not find main or interaction effect involving the associative condition factor. Therefore our prediction was not supported.

DISCUSSION AND CONCLUSION

Results from this study can be summarised in three main points. First, participants of all age groups showed more "sure" judgements following hit than FA responses, demonstrating their general ability to calibrate confidence level according to memory accuracy. In both the GG and GM conditions the differences between the confidence ratings of hit and FA responses were enlarged after strategy instruction, indicating the effects of strategy use on the calibration of confidence judgement. Inspecting Figure 1 indicates that this effect was driven primarily by an increase of "sure" response following hit responses. In other words, changes in strategy use as induced by instruction increased participants' performance as well as the confidence in their performance.

Second, as predicted by the misrecollection account, older adults showed more "sure" responses following FA errors than all other age groups. Children, while performing the memory task at a similar level as older adults (before strategy instruction), did not show such a pattern of high-confidence errors when they produced FAs. This finding is consistent with evidence gathered from developmental and lifespan studies showing that children's ability in associative binding is mature relatively early (Cowan, Naveh-Benjamin, Kilb, & Saults, 2006; Shing et al., in press; Sluzenski, Newcombe, & Kovacs, 2006; Werkle-Bergner et al., 2006). Our results also lend support to the view that the efficacy of associative binding mechanisms is compromised in old age and, furthermore, contributes to lifespan differences in high-confidence errors for rearranged lure pairs.

Third, our study extended the literature by showing that instructing an elaborative imagery strategy does not eliminate or reduce negative adult age differences in high-confidence errors. It is important to point out that, overall, the participants' performance was significantly improved after strategy instruction, as reflected in increased hit and reduced FA rates. However, as long as FAs were committed, older adults persistently experienced the phenomenological feeling of high confidence and more so than the other age groups. Similar observations can be made on the pattern of high-confidence errors in the GG and GM conditions. The two conditions were designed to vary the demands on associative binding. As can be seen in Table 2, participants' performance in the GM condition was very much lowered (i.e., lower hit and higher FA rates), demonstrating the difficulty of the high associative-demand condition. However, given the occurrence of FAs, participants (including older adults) were equally likely to associate the

FAs with high confidence (as reflected in the percentage of "sure" responses) in the GM and GG conditions. This seems to suggest that, while older adults are vulnerable in making high-confidence errors when memory retrieval require specific details, the extent of the difficulty in integrating the memory feature at encoding does not affect the relative occurrence of high-confidence errors. At the same time, the greater associative difficulty in the GM condition resulted in poorer performance and may have led children, teenagers, and younger adults to be less likely to give "sure" judgements following hit responses in the GM condition. However, this was not the case for older adults, as reflected in the significant difference between older adults and the other age groups in percentages of "sure" judgements following hit responses. Apparently, older adults lacked sensitivity to the difference in difficulty between the two conditions. These findings need to be corroborated by further investigations to determine whether they are substantially related to more general ageing-induced deficits in meta-cognitive monitoring (cf. Hertzog & Dunlosky, 2005; Souchay & Isingrini, 2004).

The findings above also suggest that the locus of the mechanism causing the high-confidence errors in older adults is not yet clearly identified. Developmentally, differences in children and older adults' memory functioning may manifest themselves at encoding and/or retrieval processes. The misrecollection account suggests that older adults have a propensity to miscombine features of different events, which in turn causes convincing high-confidence errors in false memories (Dodson et al., 2007). Assuming the manipulations on strategy instruction and associative demand mainly influenced the encoding operation, the similar levels of high-confidence errors in older adults across the conditions indicate that the errors were driven not by encoding but by retrieval difficulty. That is, older adults may have difficulty in specifying the sources of the familiarity signals at retrieval (e.g., Henkel et al., 1998; Jacoby, 1999). An alternative memory strategy that is worth investigating is the distinctiveness heuristic, which is a mode of responding emphasising participants' metacognitive awareness at retrieval, such that accurate recognition of studied items should require recollection of distinctive details (e.g., Schacter, Israel, & Raccine, 1999). The heuristic has been shown to reduce false memory and improve confidence-accuracy calibration in children (Ghetti et al., 2002) and older adults (Dodson & Schacter, 2002; Schacter et al., 1999). In as much as the pattern of high-confidence errors is reduced in older adults after the distinctiveness heuristic is instructed, the results would point towards retrieval problems as the cause of high-confidence errors.

It is important to point out that, while our study directly tested the link between the misrecollection account and associative mechanisms, our data do not disprove other alternative explanatory frameworks of false memories in ageing (e.g., inhibitory deficit, decline in recollection process). On the contrary, it is most likely that the various mechanisms at memory encoding, representation, and retrieval interact and underlie different aspects of false memory and high-confidence error phenomena. One line of research that may be informative is the neurocomputational framework proposed by Li and colleagues (e.g., Li, Lindenberger, & Sikström, 2001), in which senescent changes in aspects of cognition are simulated in relation to age-related reductions in the efficacy of dopaminergic neuromodulation (see Bäckman & Farde, 2005, for a review). Using the model it was shown that neuromodulational deficiency results in less distinct internal representation, especially in task conditions that require associative binding (Li, Naveh-Benjamin, & Lindenberger, 2005). In accordance with the misrecollection account, although at a different level, the neurocomputational model suggests that older adults' high-confidence errors may have been the behavioural manifestation of the highly activated but less distinctive internal hidden-layer representations of memory items in the simulated old networks. From the perspective of the ageing hippocampal model (Wilson et al., 2006), this would correspond to the lack of distinction between newly learned information and existing memory traces, possibly due to deteriorating functional connectivity between the entorhinal cortex and hippocampus. Extending the neurocomputational theory on deficient neuromodulation leading to less distinctive representations (e.g., Li et al., 2001, 2005), a conjecture can also be derived to interpret why the ageing-related deficit in memory calibration is specific to false-alarm responses but not to hits. This is because, by default, hit responses are derived from having "afferent sensory copies" of the associations formed in memory. There is no evidence indicating that older adults are less able to calibrate when they detect a match between stimuli and memory representation. On the other hand, for the FAs for rearranged pairs, although

there were separate afferent sensory copies of each of the items encoded in memory, there were no actual "afferent sensory copies" of these associations. Any associations of these come from illusory conjunctions. Thus, one would expect the negative consequences of less distinctive representations to be more prominent when needing to verify memory representations that lack the actual sensory encoding aspects; that is, the rearranged lure pairs (cf. Craik, 1983).

At the same time, ageing affects the functioning of strategic processes possibly via senescent changes in structure and connectivity of the prefrontal cortex (Buckner, 2004; Raz et al., 2005; West, 1996). Prefrontally mediated strategic processes during memory encoding, storage, and retrieval are mainly understood to increase the signal-to-noise ratio in the face of competing representations and to bias memory search operations in task-appropriate ways (Miller & Cohen, 2001; Rugg & Wilding, 2000; Simons & Spears, 2003). Given that rejecting the rearranged pairs puts particularly high demands on controlled processing during retrieval to overcome response tendencies triggered by familiarity signals, our findings suggest that older adults' high-confidence errors may have resulted from a reduced distinctiveness of memory traces in combination with less efficient strategic support at retrieval. This chain of effects may provide a viable explanation for the high-confidence error phenomenon in old age and calls for future studies in this direction for corroboration. Furthermore, it is important to recognise that older adults differ considerably in their mean levels and rates of decline of the cognitive and neural functioning. The multiple factor framework (Buckner, 2004) postulates that distinct age-related cascades targeting different brain systems may vary in their levels of progression across individuals (see also Hedden & Gabrieli, 2004). Specifically, a dissociation between nondemented ageing and Alzheimer's disease has been documented, such that individuals with dementia show substantial reduction in hippocampal volume while only mild effects are present in nondemented ageing (e.g., Head, Snyder, Girton, Morris, & Buckner, 2005). In this context it is worth noting that older adults varied considerably in the proportion of high confidence. Some older adults always expressed high confidence following FA responses, whereas others usually expressed lower confidence following FA responses (similar to younger adults and children). Given our conjecture that high-confidence error

may be related to hippocampal decline, future research should combine behavioural experiments with neuroimaging measures and longitudinal assessments to find out the extent to which high proportions of high-confidence false alarms form part of normal ageing or signal later risk for dementia (cf. Bäckman & Small, 2007).

REFERENCES

Bäckman, L., & Farde, L. (2005). The role of dopamine systems in cognitive aging. In R. Cabeza, L. Nyberg, & D. Park (Eds.), *Cognitive neuroscience of aging: Linking cognitive and cerebral aging* (pp. 58–84). New York: Oxford University Press.

Bäckman, L., & Small, B. J. (2007). Cognitive deficits in preclinical Alzheimer's disease and vascular dementia: Patterns of findings from the Kungsholmen project. *Physiology & Behavior*, *92*, 80–86.

Brehmer, Y., Li, S-C., Müller, V., v. Oertzen, T., & Lindenberger, U. (2007). Memory plasticity across the lifespan: Uncovering children's latent potential. *Developmental Psychology*, *43*(2), 465–478.

Brewer, N., & Wells, G. L. (2006). The confidence–accuracy relationship in eyewitness identification: Effects of lineup instructions, foil similarity, and target-absent base rates. *Journal of Experimental Psychology: Applied*, *12*, 11–30.

Buckner, R. L. (2004). Memory and executive function in aging and AD: Multiple factors that cause decline and reserve factors that compensate. *Neuron*, *44*, 195–208.

Chalfonte, B. L., & Johnson, M. K. (1996). Feature memory and binding in young and older adults. *Memory & Cognition*, *24*(4), 403–416.

Cowan, N., Naveh-Benjamin, M., Kilb, A., & Saults, J. S. (2006). Life-span development of visual working memory: When is feature binding difficult. *Developmental Psychology*, *42*(6), 1089–1102.

Craik, F. I. M. (1983). On the transfer of information from temporary to permanent memory. *Philosophical Transactions of the Royal Society of London*, *B302*, 341–359.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*, 671–684.

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

Dodson, C. S., Bawa, S., & Krueger, L. E. (2007). Ageing, metamemory, and high-confidence errors: A misrecollection account. *Psychology and Ageing*, *22*(1), 122–133.

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). Ageing and strategic retrieval processes: Reducing false memories with a distinctiveness heuristic. *Psychology and Aging*, 17(3), 405–415.

Ghetti, S., Qin, J., & Goodman, G. S. (2002). False memories in children and adults: Age, distinctiveness, and subjective experience. *Developmental Psychology*, 38(5), 705–718.

Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., Zijdenbos, A., et al. (1999). Brain development during childhood and adolescence: A longitudinal MRI study. *Nature Neuroscience*, 2(10), 861–863.

Head, D., Snyder, A. Z., Girton, L. E., Morris, J. C., & Buckner, R. L. (2005). Frontal–hippocampal double dissociation between normal ageing and Alzheimer's disease. *Cerebral Cortex*, 15, 732–739.

Hedden, T., & Gabrieli, J. D. E. (2004). Insights into the aging mind: A view from cognitive neuroscience. *Nature Reviews Neuroscience*, 5, 87–97.

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

Hertzog, C., & Dunlosky, J. (2005). Ageing, metacognition, and cognitive control. In B. H. Ross (Ed.), *The psychology of learning and motivation* (pp. 215–251). San Diego, CA: Academic Press.

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

Jacoby, L. L., Bishara, A. J., Hessels, S., & Toth, J. P. (2005). Aging, subjective experience, and cognitive control: Dramatic false remembering by older adults. *Journal of Experimental Psychology: General*, 134(2), 131–148.

Jacoby, L. L., & Rhodes, M. G. (2006). False remembering in the aged. *Current Directions in Psychological Science*, 15(2), 49–53.

Johnson, M. K. (2006). Memory and reality. *American Psychologist*, 61, 760–771.

Kausler, D. H. 1994. *Learning and memory in normal aging*. New York: Academic Press.

Killeen, P. R. (2005). An alternative to null-hypothesis significance tests. *Psychological Science*, 16(5), 345–353.

Levin, J. R. (1988). Elaboration-based learning strategies: Powerful theory = powerful application. *Contemporary Educational Psychology*, 13, 191–295.

Li, S-C., Lindenberger, U., & Sikström, S. (2001). Aging cognition: From neuromodulation to representation. *Trends in Cognitive Sciences*, 5(11), 479–486.

Li, S-C., Naveh-Benjamin, M., & Lindenberger, U. (2005). Aging neuromodulation impairs associative binding: A neurocomputational account. *Psychological Science*, 16(6), 445–450.

Loftus, E. F. (2003). Make-believe memories. *American Psychologist*, 58(11), 867–873.

Maki, R. H., & Swett, S. (1987). Metamemory for narrative text. *Memory & Cognition*, 15(1), 72–83.

Menon, V., Boyett-Anderson, J. M., & Reiss, A. L. (2005). Maturation of medial temporal lobe response and connectivity during memory encoding. *Cognitive Brain Research*, 25, 379–385.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, 24, 167–202.

Naveh-Benjamin, M. (2000). Adult age differences in memory performance: Tests of an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 26, 1170–1187.

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.

Ofen, N., Kao, Y-C., Sokol-Hessner, P., Kim, H., Whitfield-Gabrieli, S., & Gabrieli, J. D. E. (2007). Development of the declarative memory system in the human brain. *Nature Neuroscience*, 10(9), 1198–1205.

Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review*, 76, 241–263.

Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart & Winston.

Persson, J., Nyberg, L., Lind, J., Larsson, A., Nilsson, L-G., Ingvar, M., et al. (2006). Structure–function correlates of cognitive decline in aging. *Cerebral Cortex*, 16, 907–915.

Pressley, M., Levin, J. R., Ghatala, E. S., & Ahmad, M. (1987). Test monitoring in young grade school children. *Journal of Experimental Child Psychology*, 43, 96–111.

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Journal of Applied Psychology Measurement*, 1, 385–401.

Raz, N., Lindenberger, U., Rodrigue, K. M., Kennedy, K. M., Head, D., Williamson, A., et al. (2005). Regional brain changes in ageing healthy adults: General trends, individual differences, and modifiers. *Cerebral Cortex*, 15(11), 1676–1689.

Richardson, J. T. E. (1998). The availability and effectiveness of reported mediators in associative learning: A historical review and an experimental investigation. *Psychonomic Bulletin & Review*, 5(4), 597–614.

Roebers, C. M. (2002). Confidence judgements in children's and adults' event recall and suggestibility. *Developmental Psychology*, 38(6), 1052–1067.

Roebers, C. M., & Howie, P. (2003). Confidence judgements in event recall: Developmental progression in the impact of question format. *Journal of Experimental Child Psychology*, 85, 352–371.

Roebers, C. M., von der Linden, N., Schneider, W., & Howie, P. (2007). Children's metamemorial judgements in an event recall task. *Journal of Experimental Child Psychology*, 97, 117–137.

Rugg, M. D., & Wilding, E. L. (2000). Retrieval processing and episodic memory. *Trends in Cognitive Sciences*, 4(3), 108–115.

Schacter, D. L., Israel, L., & Raccine, C. (1999). Suppressing false recognition in younger and older adults: The distinctiveness heuristics. *Journal of Memory and Language*, 40, 1–24.

Schneider, S. L., & Laurion, S. K. (1993). Do we know what we've learned from listening to the news? *Memory and Cognition*, *21*(2), 198–209.

Schneider, W., & Pressley, M. (1997). *Memory development between two and twenty* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates Inc

Shing, Y. L., Werkle-Bergner, M., Li, S-C., & Lindenberger, U. (in press). Associative and strategic components of episodic memory: A lifespan dissociation. *Journal of Experimental Psychology: General*.

Simons, J. S., & Spiers, H. J. (2003). Prefrontal and medial temporal lobe interactions in long-term memory. *Nature Reviews Neuroscience*, *4*, 637–648.

Sluzenski, J., Newcombe, N., & Kovacs, S. L. (2006). Binding, relational memory, and recall of naturalistic events: A developmental perspective. *Journal of Experimental Psychology: Learning, Memoryand Cognition*, *32*(1), 89–100.

Souchay, C., & Isingrini, M. (2004). Age related differences in metacognitive control: Role of executive functioning. *Brain And Cognition*, *56*(1), 89–99.

Souchay, C., Isingrini, M., & Espagnet, L. (2000). Aging, episodic memory feeling-of-knowing, and frontal functioning. *Neuropsychology*, *14*, 299–309.

Sowell, E. R., Peterson, B. S., Thompson, P. M., Welcome, S. E., Henkenius, A. L., & Toga, A. W. (2003). Mapping cortical change across the human lifespan. *Nature Neuroscience*, *6*(3), 309–314.

Werkle-Bergner, M., Müller, V., Li, S-C., & Lindenberger, U. (2006). Cortical EEG correlates of successful memory encoding: Implications for lifespan comparisons. *Neuroscience and Biobehavioral Reviews*, *30*, 839–854.

West, R. L. (1996). An application of prefrontal cortex function theory to cognitive aging. *Psychological Bulletin*, *120*(2), 272–292.

Wilson, I. A., Gallagher, M., Eichenbaum, H., & Tanila, H. (2006). Neurocognitive aging: Prior memories hinder new hippocampal encoding. *Trends in Neurosciences*, *29*(12), 662–670.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*(3), 441–517.