Abstract:

The noun-pair lookup (NP) task was used to evaluate strategic shift from visual scanning to retrieval. We investigated whether age differences in feeling-of-knowing (FOK) account for older adults' delayed retrieval shift. Participants were randomly assigned to one of three conditions: (1) standard NP learning, (2) fast binary FOK judgments, or (3) Choice, where participants had to choose in advance whether to see the look-up table or respond from memory. We found small age differences in FOK magnitudes but major age differences in memory retrieval choices that mirrored retrieval use in the standard NP task. Older adults showed lower resolution in their confidence judgments (CJs) for recognition memory tests on the NP items, and this difference appeared to influence rates of retrieval shift, given that retrieval use was correlated with CJ magnitudes in both age groups. Older adults had particular difficulty with accuracy and confidence for rearranged pairs, relative to intact pairs. Older adults' slowed retrieval shift appears to be attributable to (1) impaired associative learning early in practice, not just a lower FOK; but also (2) retrieval reluctance later in practice after the degree of associative learning would afford memory-based responding.

Keywords: age differences | aging | human information storage | memory | metacognition | associative processes | skill learning | strategy | psychology

Article:

Acknowledgement: We thank Elizabeth Kirk, Jared Pryor, Shannon Langston, Miranda Pryor, and Cassandra Pryor for assistance with program implementation and data collection, and Jarrod Hines for assistance with data preparation and management. This research was supported by a grant from the National Institute on Aging, one of the National Institutes of Health (R01 AG024485).

Correspondence concerning this article should be addressed to: Christopher Hertzog, School of Psychology, Georgia Institute of Technology, Atlanta, GA 30332-0170 Electronic Mail may be sent to: christopher.hertzog@psych.gatech.edu.

The present study concerns age differences in feelings-of-knowing and strategy choice during a relatively simple skill acquisition paradigm, the noun-pair (NP) lookup task. The NP task presents a pair of nouns at the center of the screen (henceforth termed the probe). The task requires a visual search of table of NPs at the top of the screen to determine whether any of the tabled pairs matches the probe (Ackerman & Woltz, 1994). At the start of the task, individuals can only make correct decisions via visual search, scanning the table until they find a pair that confirms or disconfirms a match to the table. However, if the tabled pairings remain unchanged on every trial [consistent mapping in Shiffrin & Schneider's (1977) terms], participants can incidentally learn them and respond to the probe on the basis of memory retrieval rather than visual scanning. The strategic shift from slow rates of visual scanning to faster rates of memory retrieval is the primary source of response time (RT) improvement with practice in the NP task.

The NP task is an example of skill acquisition tasks involving a shift from algorithmic processing to memory-based responding. Theoretical accounts of the retrieval shift in such tasks include mechanistic accounts based on associative strengthening, including Logan's (1988) instance theory and its successors (e.g., Nosofsky and Palmieri, 1997). By these accounts, the algorithm and retrieval are executed in parallel, and the fastest process wins. Only with extended repetitions does the associative strength achieve a level that supports fast retrieval as the winning process. An alternative view is that there is an early strategy selection process for each trial (e.g., Rickard, 1997), such that individuals first decide either to execute the algorithm or retrieve the answer (Bajic & Rickard, 2009).

Strategic Shift in the NP Task

Individual differences in NP retrieval shift are influenced by cognitive abilities (Ackerman & Woltz, 1994; Rogers, Hertzog, & Fisk, 2000). Individuals high in fluid intelligence, processing speed, and episodic memory shift to the retrieval strategy more quickly. Retrieval shift is also influenced by individuals' processing goals and response criteria (e.g., Bourne, Raymond, & Healy, 2010). For example, monetary incentives or instructions to use the retrieval strategy also speed the shift to retrieval in the NP task (Touron, Swaim, & Hertzog, 2007), as they do in other skill acquisition tasks like the pound arithmetic task (Touron & Hertzog, 2009). Such outcomes suggest that the decision to scan or retrieve in the NP task is influenced by top-down mechanisms, including confidence in one's ability to rely on memory retrieval (Lamson & Rogers, 2008; Touron & Hertzog, 2004a).

Older adults' performance patterns suggest they are more likely to avoid retrieval – even when they could do so without making errors. Monetary incentives to respond more quickly substantially improve older adults' retrieval shift rates (Touron, Swaim, & Hertzog, 2007; Touron & Hertzog, 2009). Touron, Swaim, & Hertzog (2007) showed that the faster retrieval shift was not accompanied by higher NP error rates, as would be expected if monetary incentives had simply induced a liberal speed-accuracy tradeoff response criterion.

Older adults' retrieval avoidance is also indicated by their lower conditional probability of retrieval strategy use on a standard NP task trial, given that they had correctly recognized a probe for that same item in a preceding recognition memory trial (Touron & Hertzog, 2004a). Even so, retrieval shift is actually facilitated more for older adults by inserting such recognition memory test trials into the NP task. These recognition test trials the word pairs without the lookup table, so that the response must be based on memory (Rogers & Gilbert, 1997; Touron & Hertzog, 2004a). The speeding of retrieval shift when recognition probes are inserted could be caused by overcoming a metacognitive deficit – alerting individuals to the fact that they know the pairings, thereby hastening their retrieval shift. Retrieval avoidance could represent a dampening influence of a broader strategic task set that promotes scanning despite older adults' feelings of knowing the item pairings.

In this sense, older adults appear to be more cautious about using the retrieval strategy (e.g., Botwinick, 1984), although this effect may be attributable to a subset of older adults being highly resistant to using the retrieval strategy (Touron & Hertzog, 2004b). Older adults may be more risk-averse because reliance on memory retrieval at intermediate levels of learning risks response errors. In addition, older adults may simply be resistant to change a strategy that is effective in minimizing error rates (Spieler, Mayr, & LaGrone, 2006) unless provided with compelling motivation to change their strategic approach to the task.

Feeling of Knowing and Choice Mechanisms

What is the basis for retrieval strategy choice in the NP task? One important perspective on this question derives from the Source-Activation Confusion model of Reder and colleagues (Reder & Ritter, 1992; Schunn, Reder, Nhouyvanisvong, Richards, & Stroffolino, 1997). This model states that an early familiarity process drives strategy selection. Presentation of a candidate stimulus rapidly activates a familiarity-based feeling-of-knowing (FOK) state. The FOK is a feeling or belief that one can successfully retrieve information from memory, either before retrieval is attempted or after a retrieval attempt is initially unsuccessful. A high FOK leads to a decision to search memory rather than to compute the answer (in the case of the NP task, to visually search the look-up table). Strategic choice involves an FOK state rather than direct access to the contents of memory, because individuals can be lured into an incorrect retrieval choice when given new foils that closely resemble target stimuli (Reder & Ritter, 1992).

In typical FOK studies, individuals are asked to search memory for a fact likely to be in semantic memory (e.g., “what is the capital city of France?”) or for information in episodic memory (e.g., “what word was previously paired with QUAIL?”). FOKs for unrecalled semantic information (e.g., Connor, Balota, & Neely, 1992) or episodic items (e.g., Hertzog, Dunlosky, & Sinclair, 2010; Metcalfe, Schwartz, & Joaquim, 1993) are affected by experimental manipulations of FOK cue familiarity. However, FOKs are also known to be influenced by deliberate searches of memory for information about an unrecalled target, and products of such searches, irrespective of whether these products derive from the correct target, also influence FOKs (e.g., Koriat &

Levy-Sadot, 2001). Cue familiarity may be accessed rapidly and hence may be the major influence on an initial FOK. In recognition memory paradigms, an accurate judgment of familiarity can be completed faster than an accurate recollective experience, consistent with this account (e.g., Yonelinas, 2002). In the NP task environment, an initial FOK that would influence an early strategy choice (Bajic & Rickard, 2009) could well be based on cue familiarity.

Studies of FOK generally require individuals to make ordinal FOK judgments rating the likelihood of subsequent recognition. Reder's studies used a different method. Individuals were asked to make an initial strategy choice (e.g., retrieve vs. compute), which was taken as a direct outcome of the FOK state that generates the strategic choice. The present study compares both methods within the NP task. We do so because we assume that the FOK state and the strategy choice process are separable and distinct. Namely, individuals can behave suboptimally, selecting to scan even when they have a high FOK that would lead to a correct retrieval-based response with faster RT. By this view, the FOK state seeds the choice process, but the strategic choice is influenced by other variables as well.

FOK and Older Adults' NP Retrieval Shift

Using a strategy choice procedure modeled after Reder and Ritter (1992), Lamson and Rogers (2008) showed that older adults are less likely to choose to retrieve in an arithmetic computation task. They used bonus points as incentives to shift to retrieval, with a fast deadline for choosing to compute or retrieve. A subset of their older adults successfully shifted to retrieval, whereas a subset did not. These groups differed in perceived change in memory and in cognitive performance (see also Rogers, Hertzog, & Fisk, 2000; Touron & Hertzog, 2004a). Several participants did not increase retrieval rates with practice. Individuals who did not shift showed similar sensitivity of retrieval shift to correct responses and RT, suggesting that their retrieval reluctance was associated with a failure to learn the correct answers to the problems with extensive task practice. Such effects could be consistent with an associative memory deficit (Cerella, Onyper, & Hoyer, 2006; Naveh-Benjamin, 2000). Lamson and Rogers (2008) also found that older adults not meeting performance criteria actually had high associations of strategy choice with subsequent performance – a conjunction of correct responses with fast RT after retrieval (versus slow RT after choosing to calculate). They argued that older adults were making valid choices to retrieve or calculate based on FOK confidence.

On the surface, some of Lamson and Rogers' (2008) findings seem inconsistent with results from our laboratories regarding recognition memory test confidence judgments (CJs) in the NP task. CJs rate the likelihood that the test decision is correct (e.g., that the candidate word pair actually matches an entry in the table). Touron and Hertzog (2004a) found older adults to have high confidence in the accuracy of their recognition test responses, even when they subsequently used the scanning strategy for the same item on a later standard NP trial. Furthermore, within-person correlations of CJs with recognition memory test accuracy were high for older adults, suggesting accurate retrieval monitoring. The apparent discrepancy could be attributable to different kinds of metacognitive judgments. CJs after a recognition memory response do not have the same temporal dynamics as initial FOKs. Initial FOKs governing early strategy choice at the onset of a standard NP trial should be relatively fast, because participants are instructed to respond as quickly as possible. It could be the case that older adults have low FOKs at NP trial onset, when a strategic choice must be made, but have high CJs for recognition memory test trials after a slow retrieval success (Cerella, Onyper, & Hoyer, 2006). Low initial FOKs would lead to a choice to scan rather than retrieve, as in Lamson and Rogers (2008).

Alternatively, the difference in results could also have been influenced by how frequently Touron and Hertzog (2004a) assessed recognition memory for the pairings. In that experiment, recognition memory test trials were randomly intermixed with standard NP trials, with an equal number of both types of trials. As noted earlier, testing recognition memory during NP practice greatly speeded older adults' retrieval shift, and hence their high CJs could have resulted from the frequent assessments of recognition memory (with accuracy feedback) that informed them about their knowledge of the word pairs. By using the choice method of Reder and colleagues, Lamson and Rogers' (2008) experiment probably did not guarantee that older adults would glean the same quality of diagnostic information about memory improvements with practice from task experience.

The present study was designed to assess the hypothesis that initial FOKs would show age differences and explain older adults' delayed strategy shift. In doing so, we contrasted the initial FOK as measured by a metacognitive judgment from an initial strategy choice. We used two different conditions: a binary high-low FOK rating (i.e., how confident am I that I can correctly recognize this pairing), and a fast retrieve/scan strategy choice, as in Reder & Ritter (1992). Individuals in the Choice condition were first presented with the NP probe item without the look-up table and asked to make a fast decision to scan or to retrieve the answer from memory. If scanning was chosen, the look-up table appeared. Otherwise the response had to be made on the basis of memory. In the FOK condition, individuals were also first presented with the NP probe item without the look-up table but then asked to make a binary (high-low) FOK, after which the look-up table appeared, regardless of the FOK response.

The two conditions aided us in discriminating between an associative deficit hypothesis and a strategic retrieval reluctance hypothesis as competing accounts of the delayed retrieval shift for older adults. A pure associative learning deficit stipulates that the delay is solely a function of degree of item learning. Hence it predicts reliably lower FOKs for older adults. Likewise, it predicts no difference in the likelihood of a high FOK response and a selection of the retrieval strategy in the Choice condition. Both measures should reflect little more than underlying associative strength of that item. Assuming no age differences in the accuracy of initial FOKs, these judgments should also show equivalent relationships in younger and older adults to retrieval strategy use, as reported at the end of the trial, and to subsequent recognition memory tests on the same items.

In contrast, the retrieval reluctance hypothesis predicts that older adults would be less likely to choose to retrieve, even when in a high FOK state. Hence, it predicts a significant difference in the likelihood of retrieval choices, relative to the likelihood of high FOKs. Furthermore, it predicts a lower correlation of older adults' FOKs to their retrieval choices, relative to younger adults.

Recognition Memory Test Trials

Although we still wished to obtain recognition memory test trials and CJs, we also wished to avoid providing experimental feedback based on recognition memory for the NPs early in practice. We also did not want to further complicate the design by having recognition memory tests as a third type of mixed trial in the FOK and Choice conditions. Hence we gave recognition memory tests in separate, discrete blocks during the NP task, beginning with the fourth block of trials (see Methods for details).

Intact Versus Rearranged Pairs

One nuance regarding CJs is that they differ substantially by whether the probe is an intact pair (matching the corresponding pair in the table) or a rearranged pair. Rearranged pairs are subject to recollection-based strategies, such as recall-to-reject (Cohn & Moscovitch, 2007). Recollection experiences during recognition occur less frequently for older adults (Light, Prull, LaVoie, & Healy, 2000), which in turn affects their accuracy for recognizing rearranged pairs (Cohn, Moscovitch, and Emrich, 2008). Hines, Touron, & Hertzog (2009) found a differential association of CJs for intact versus rearranged pairs on subsequent item study times in a multi-trial recognition memory task for both younger and older adults. Hines, Touron, & Hertzog (2009) also found an age difference in CJ resolution (accuracy) favoring young adults. By analogy, individuals in the NP task may receive less reinforcement for incidental learning when confronted with a rearranged pair, especially if they cannot recollect the answer. Unlike the intentional learning task of Hines, Touron, & Hertzog (2009), older adults' retrieval choices in the NP task may be differentially affected by less accurate monitoring of incidentally produced recognition accuracy for rearranged pairs (as measured by CJs). We therefore evaluated whether FOK and Choice behaviors differed as a function of whether the probe matched or did not match the item in the look-up table.

In summary, our study was designed to evaluate the following hypotheses: (1) slowed retrieval shift would reflect low initial FOKs by older adults, versus (2) older adults would show conservative choices to retrieve, relative to their level of FOK; (3) older adults would manifest lower recognition accuracy and lower CJs on rearranged pairs, relative to intact pairs; and (4) these effects would be associated with lower levels of retrieval strategy use on standard NP trials.

Method

Participants and Design

We tested 73 young adults (ages 18-24, M_{age} = 19.6, 47% female) and 79 older adults (ages 61-81, M_{age} = 69.5 years, 67% female). Young adults were University students who participated for extra credit. Older adults were recruited from the community and received a modest honorarium for their participation. Table 1 reports participant characteristics for the older and younger adults. We administered the ETS Advanced Vocabulary and Letter Sets tests to measure crystallized intelligence (vocabulary) and fluid intelligence (inductive reasoning; see Ekstrom, French, Harman, & Dermen, 1976). Typical patterns of age differences were found, with older adults having reliably lower inductive reasoning but higher vocabulary scores. Older adults also reported lower self-reported health. On average the older adults were relatively well-educated, with about 15 years of formal schooling.

Table 1 is omitted from this formatted document.

Participants from both age groups were randomly assigned to the following conditions: (1) standard NP learning, (2) FOK judgments, or (3) Choice, where participants chose to see the look-up table or respond from memory. Participants in the standard NP condition completed only standard NP trials, whereas those in the FOK and Choice conditions completed standard trials as well as FOK or choice trials. There were 32 young and 23 older adults in the standard NP condition, 26 young and 25 older adults in the FOK condition, and 25 young and 27 older adults in the Choice condition.

Materials and Procedures

In session 1, participants completed a synonym matching task and a lexical decision task. These measures are not relevant to the present project other than the speed of synonym matching, which was used as an individually-tailored response prompt for fast FOK and Choice responses (see below).

Participants completed the NP task in session 2. A Visual Basic 6.0 program (Microsoft Visual Studio, 2007) controlled stimulus presentations and response recordings. Stimuli were presented in 15-point Arial font on a15-inch LCD monitor with a resolution of 1024 × 768. Seating and monitors were adjusted to a height and distance that optimized each participant's viewing and comfort. Throughout the task, each block was followed by a short break, during which participants received feedback on their mean RT and accuracy. However, participants were not given feedback on RT and accuracy after the first block of trials or for any recognition memory test trial blocks. Participants were given the option to extend these breaks as needed.

The stimulus set contained 24 semantically unrelated concrete nouns which were randomly paired for each participant to form 12 paired-associate items (e.g., TABLE – APPLE) that constituted the elements of the lookup table. On a given NP task trial, a target word pair was presented in the center of the screen. The target pair was intact if it matched (i.e., was identical to) one of the pairs in the lookup table. Mismatched or rearranged trials paired a left-hand word from one pair with a randomly selected right-hand word from a different pair. The physical location of the intact pairs in the table was randomly rearranged for each trial. On a standard NP trial, only a target pair was presented. Participants pressed a key labeled “Y” if the target pair was intact or a key labeled “N” if the target pair was rearranged from pairs in the lookup table. If participants responded to the NP trial incorrectly, they received a centrally presented “ERROR” message after their response.

Twelve blocks of trials were completed, each containing 48 trials. For all participants, the first trial block consisted of 48 standard NP trials, half of which presented intact and half of which presented rearranged word pairs. Starting with the second trial block, the nature of trials varied for the three between-subjects groups. In the standard NP condition, all trials in a block consisted of the standard NP trial just described. In the FOK and Choice conditions, half the trials were standard NP trials. In the FOK condition, the other 24 trials presented the target pair without the table, requesting a binary (high, low) FOK response. FOKs were prompted by the question “How confident are you that you know whether this work pair is matched or unmatched?” Individuals were instructed to press 1 for low confidence, 2 for high confidence. After the FOK response, the lookup table was displayed and remained on the screen until the participant responded.

In the Choice condition, the other 24 trials also began with presentation of the target pair. Participants were asked to indicate whether they chose to scan (which would cause presentation of the lookup table, as in the FOK condition) or retrieve (in which case the lookup table would not be displayed. The choice was prompted by the instruction, “Please make a choice as to which strategy you will use for this word pair on the following trial.” Participants pressed a key labeled “S” for scan or one labeled “M” for memory. If the participant chose scanning, the lookup table was displayed immediately, as in the FOK condition. If the participant chose to retrieve, the lookup table was not displayed, forcing the decision to be based on memory.

Individuals were instructed to make an FOK or select a strategy choice quickly, but no firm deadline was imposed. Instead, the probe pair changed color from black to red, signaling that individuals should make their strategy choice or FOK immediately. The timing of this response prompt was based on each individual's RT in the synonym matching task collected earlier (e.g., Hertzog, Raskind, & Cannon, 1986). The task required rapid judgments of whether two common nouns had the same meaning (e.g., THIEF – BURGLAR) as opposed to different meanings. Participants' 90th percentile of their cumulative RT distribution for correct trials was used as the response prompt in the FOK and Choice conditions. Participants were instructed to give their FOK or Choice as rapidly as possible, but to respond no later than when the font color of the words changed from black to red. Given age-related slowing of RTs, the response prompts differed for the two age groups; the mean 90th percentile was 1.6 s for young adults ($SD = 0.6$) and 2.3 s for old adults ($SD = 0.8$).

After each standard NP trial, participants in all three conditions were next asked to report the strategy they had used to make the discrimination of intact or rearranged pairing. They did so by pressing labeled keys: “S” if they used the scanning strategy, “M” if they used the memory retrieval strategy, “B” if they used both strategies, or “O” (signifying “other”) if they used a strategy not listed above. The same strategy report procedure was used for FOK trials in the FOK condition. That report procedure was also used in the Choice condition when participants chose the scan strategy, as reversion to the memory strategy remained possible. When participants in the Choice condition chose the memory strategy, the lookup table was not presented, so scanning was not a viable strategy, For these trials, the strategy reports were therefore not collected.

In Blocks 4, 8, 12, and 16, participants completed recognition memory tests, which were the same as standard NP trials except that the lookup table was absent, requiring a memory-based response. After each memory test, participants were asked to provide a CJ, reporting their level of confidence that their preceding answer was correct by pressing a key labeled “0%” through “100%” in increments of 10%. If participants answered the recognition memory test incorrectly, the CJ was followed by the word “ERROR” presented centrally on the following screen for 1 second, followed by the next trial.

After the NP task, participants completed a cued-recall test of their memory for the NPs. The first word in the pair was presented, and participants were asked to report its associated word. The cued recall test also included judgments of learning (JOLs) to evaluate confidence in how well the associations had been learned. JOL-recall correlations were used to determine whether individuals could discriminate items they had learned from items they had not yet learned (see Touron & Hertzog, 2004b, for a more complete description of this task).

Results and Discussion

Standard NP Trials

Results from the standard NP trials in the FOK and Choice condition both (1) replicated previous findings (e.g., Touron & Hertzog, 2004a) on age differences in NP RT, accuracy, and retrieval reports, and (2) were similar to the standard NP trials that comprised the Control condition. To save space, we do not report on standard NP trials in detail. To briefly summarize, RT decreased reliably from the beginning to the end of practice for both intact and rearranged items (see Figure 1). RT improved reliably faster for younger compared to older adults. Error rates were low and consistent across blocks of practice.

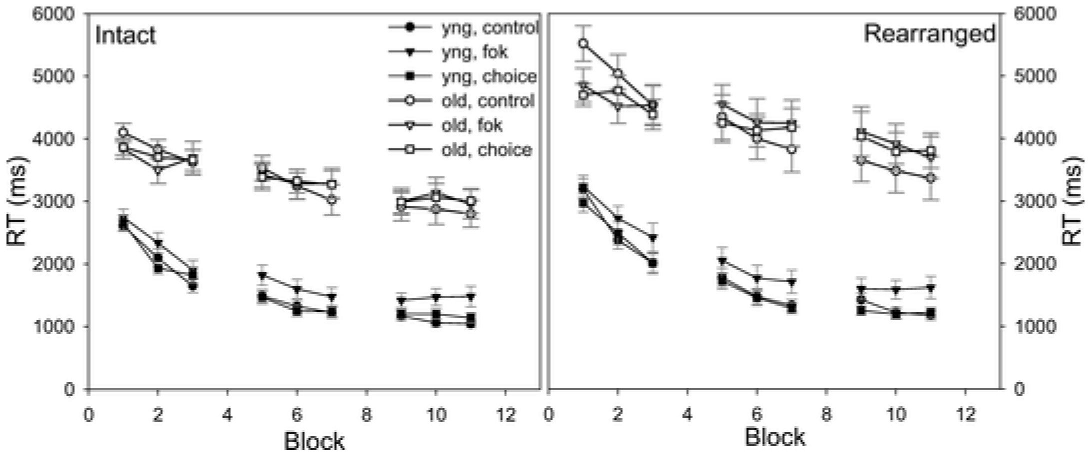


Figure 1. Standard trial RTs as a function of block by age and condition.

As in previous studies, the RT improvement corresponded to a major increase in self-reported retrieval strategy use. Figure 2 (top) shows the percentage of retrieval strategy reports for the different conditions. Reported retrieval use increased monotonically over blocks for both groups, but the rate of increase was lower for older adults. This inference was supported by a reliable Age × Block interaction, $F(7, 1022) = 5.18$, $p < 0.001$ attributable to the quadratic trend on blocks, $F(1, 146) = 21.72$, $p < 0001$. Retrieval rates were lower for rearranged pairs, $F(1, 146) = 26.69$, $p < 0.001$, $d = 0.11$, with the effect being mainly attributable to older adults' lower likelihood of reporting retrieval on rearranged trials, $F(1, 146) = 16.71$, $p < 0.001$, $d = 0.19$. Whether probes were intact or rearranged pairs had little effect on younger adults' retrieval strategy use, $d = 0.03$.

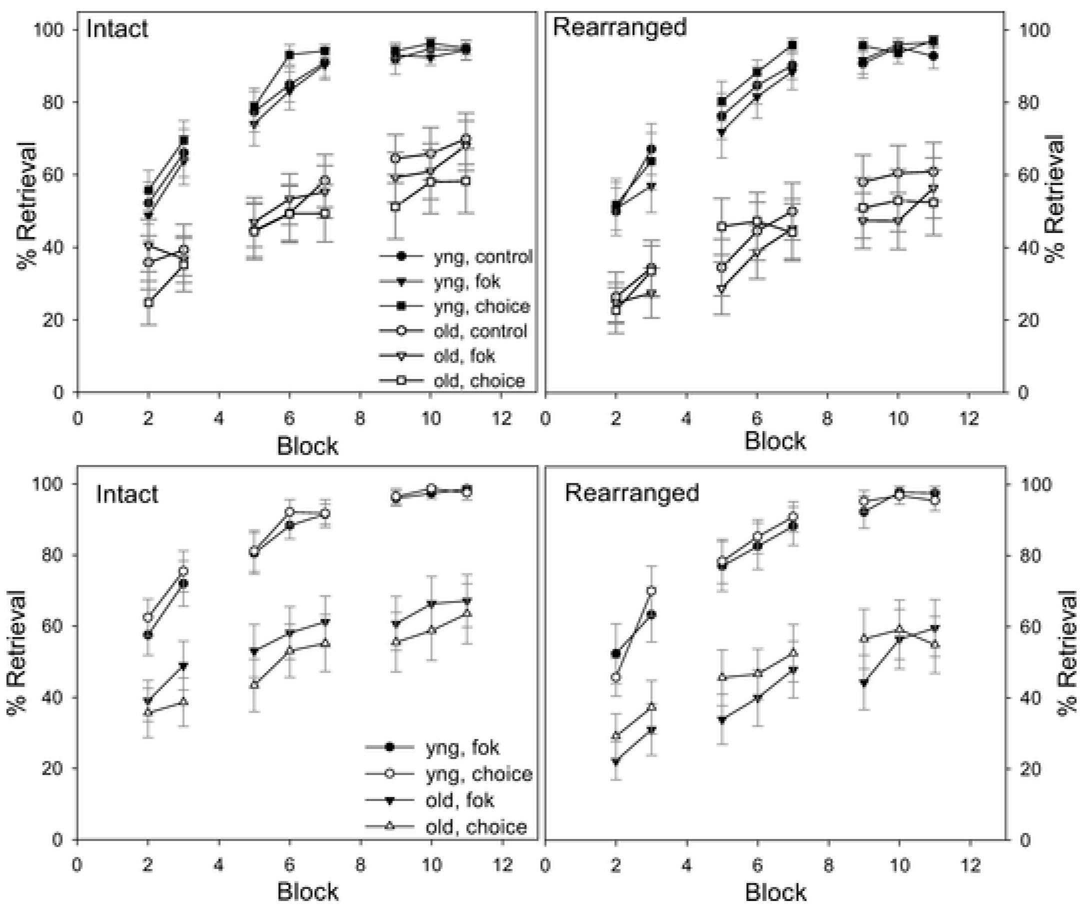


Figure 2. Retrieval reports for standard trials (top) and after FOK and Choice decisions (bottom) as a function of block by age and condition.

These findings are consistent with other studies indicating that older adults have special difficulty with rejecting rearranged pairs during associative recognition tests (Cohn, Emrich, & Moscovitch, 2008; Hines, Touron, & Hertzog, 2009). We return to this point when evaluating the recognition memory test data. More generally, the pattern of age differences in retrieval reports for NP trials are consistent with an age-related associative learning deficit (Cerella, Onyper, & Hoyer, 2006).

Choice and FOK Trials

The critical questions for the study involved behavior on choice and FOK trials; we examine hypotheses regarding FOK as a factor in producing older adults' delayed retrieval shift.

Retrieval reports

Figure 2 (bottom) plots the aggregate retrieval choices and reports in the Choice condition and the retrieval reports after FOK trials. These results paralleled the retrieval reports for standard NP trials. There were large age differences in the likelihood of either a retrieve choice or a reported retrieval after scanning choice or FOK, $F(1, 98) = 49.34$, $p < 0.001$, $d = 1.12$, $M_{young} = 88$, $M_{old} = 51$). This difference was qualified by a three-way interaction of Age × Match × Condition (FOK or choice), $F(1, 98) = 5.85$, $p < 0.05$. This interaction reflected the fact that older adults were far less likely to report retrieval for rearranged items relative to intact pairs in the FOK condition, $M_{Rearranged} = 42$ and $M_{Intact} = 57$, respectively. The comparable marginal means in older adults' Choice condition decisions were $M_{Rearranged} = 52$ and $M_{\text{Intact}} = 55$. As with retrieval reports on standard trials, young adults' retrieval selections were less affected by whether probes were rearranged ($M_{Rearranged} = 81$ in the FOK condition, $M_{Rearranged} = 92$ in the Choice condition) or intact ($M_{Intanct} = 85$ in the FOK condition, $M_{Intanct} = 94$ in the Choice condition). Instead, younger adults were more likely to choose to retrieve in the Choice condition, relative to reported retrieval the FOK condition. Retrieval increased as a function of blocks, primarily because of the linear trend on blocks, $F(1, 98) = 70.69$, $p < 0.001$. There was a reliable age × block interaction, indicating less of an increase in retrieval for older adults, $F(7, 686) = 2.57$, $p < 0.05$.

Choice and FOK responses

The main data of interest concern the similarity of choice condition selection behavior and FOKs for young and old adults. For the Choice condition, we evaluated the mean proportion of trials for which each age group selected retrieval (i.e., chose not to view the lookup table) in comparison to the probability of a high FOK in the FOK condition. Figure 3 plots the percentage of trials where participants either chose to retrieve (in the Choice condition) or gave a high FOK (in the FOK condition), separating the data by intact and rearranged pairs. We conducted a mixed model analysis treating Condition (FOK versus Choice) as a between-subjects factor. High FOKs and retrieval choices were more likely for intact NP probes than for rearranged probes, $F(1, 98) = 142.49$, $p < 0.001$, $d = 0.31$. Furthermore, Condition interacted with Match, $F(1, 98) = 67.97$, $p < 0.001$, and the three-way interaction of Age × Match × Condition was also reliable, $F(1, 98) = 27.52$, $p < 0.001$. Hence we evaluated responses separately for intact and rearranged pairs.

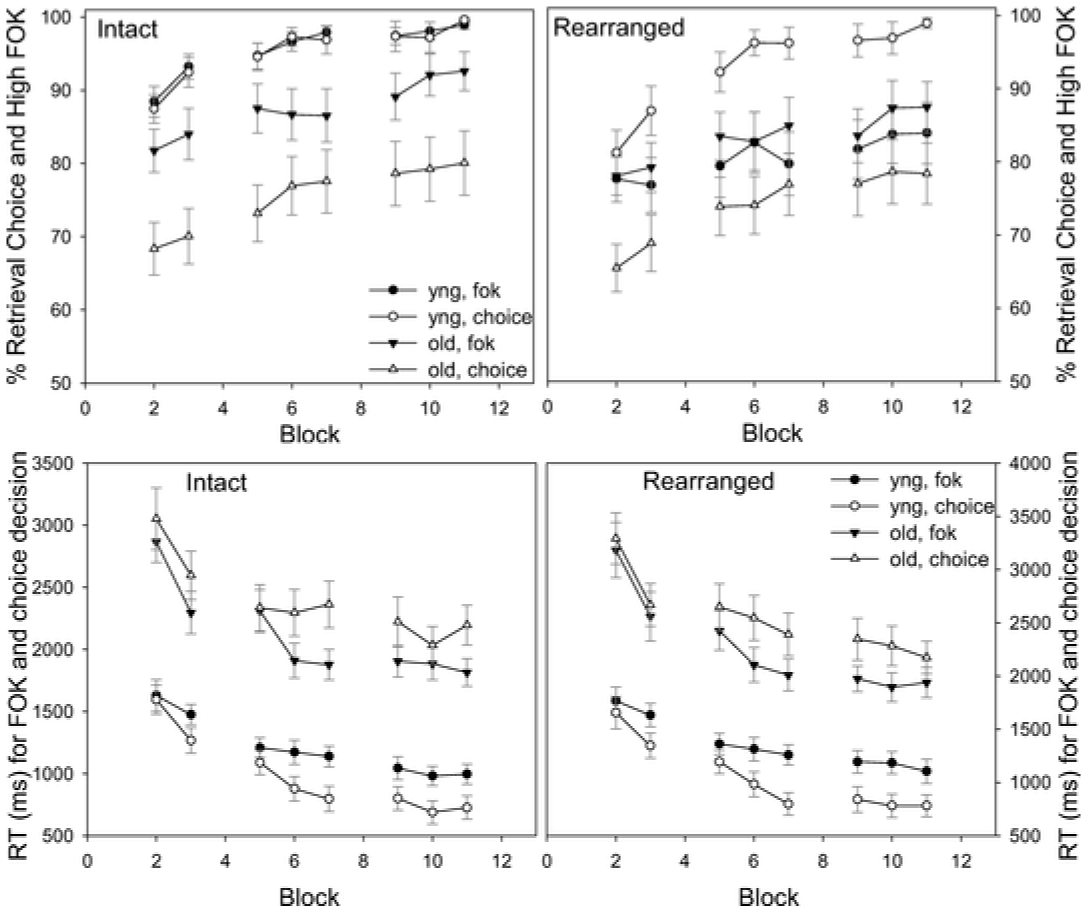


Figure 3. Choice behavior and FOK judgments (top) and decision RTs (bottom) for Choice and FOK judgments for young and old adults.

For intact pairs, there were reliable age differences in the likelihood of high FOK or retrieval choice, $F(1, 98) = 30.98$, $p < 0.001$, $d = 0.91$, and this age difference was moderated by Condition, $F(1, 98) = 5.42$, $p < 0.05$. On average, older adults had lower FOKs and were more likely to choose seeing the look-up table (opting for scanning over retrieval) compared with young adults. There was little difference between the two conditions for the young ($M_{FOK} = 96$, $M_{Choice} = 95$, $d = 0.05$), but older adults had a reliably lower probability of choosing to retrieve in the Choice condition, relative to the likelihood of giving a high FOK ($M_{FOK} = 88$, $M_{Choice} = 75$, $d = 0.63$). This effect suggests retrieval reluctance by the older adults; their choice to retrieve was lower than their level of FOK. Both retrieval choices and high FOK increased over blocks, $F(7, 686) = 21.00$, $p < 0.001$, because of the linear [$F(1, 98) = 40.52$, $p < 0.001$] and quadratic [$F(1, 98) = 7.52$, $p < 0.01$] trends.

Rearranged pairs produced a different and intriguing pattern of outcomes. Although there was a reliable age main effect, $F(1, 98) = 6.65$, $p < 0.05$, $d = 0.42$, there was also a robust Age × Condition interaction, $F(1, 98) = 11.65$, $p < 0.001$. As can be seen in Figure 3, the condition effects went in opposite directions for the two age groups. Younger adults had much higher likelihood of a retrieval choice than a high FOK (M_{choice} = 93.2, $MFOK$ = 80.1); the opposite was true for older adults (M_{choice} = 74.2, $MFOK$ = 83.4). This outcome suggested that younger adults had a general disposition to choose retrieval in the Choice condition, whereas older adults had the opposite disposition to choose scanning, relative to FOK confidence levels. On average, types of FOK increased over blocks, $F(7, 686) = 19.72$, $p < 0.001$, because of the reliable linear [$F(1, 98) = 34.54$, $p < 0.001$] and quadratic [$F(1, 98) = 8.91$, $p < 0.01$ trends].

These results refute the hypothesis that older adults might have low FOKs at the start of an NP trial, leading to retrieval reluctance. Older and younger adults actually had similar rates of experiencing high FOKs. Instead, there was a substantial difference between FOK magnitudes and the likelihood of choosing retrieval for older adults, but not for younger adults. Older adults in the Choice condition – which required an early scan choice to see the lookup table – were far less likely to commit early to the retrieval strategy. Instead, they were more likely to choose to see the lookup table. The difference between the FOK and Choice conditions for older adults is more consonant with the retrieval avoidance hypothesis than with a pure associative learning deficit.

This outcome also reinforces the argument that the strategy choice procedure is not merely an alternative method for collecting FOKs, as assumed by Reder and colleagues (e.g., Reder & Ritter, 1992). Instead, it involves both (1) metacognitive monitoring (as reflected in an FOK) and (2) a rapid strategic choice based on the FOK. Other studies have demonstrated that FOKs are linked, on a probabilistic basis, with control over memory, such as the decision to terminate searching for an answer to a question (e.g., Singer & Tiede, 2008; Nelson & Narens, 1990). However, as with any metacognitive state, exerting control based on FOK magnitude is possible but not inevitable. In the NP task, older adults often experience a relatively high FOK but still choose to scan, probably so that they can verify their answer by searching for it in the lookup table.

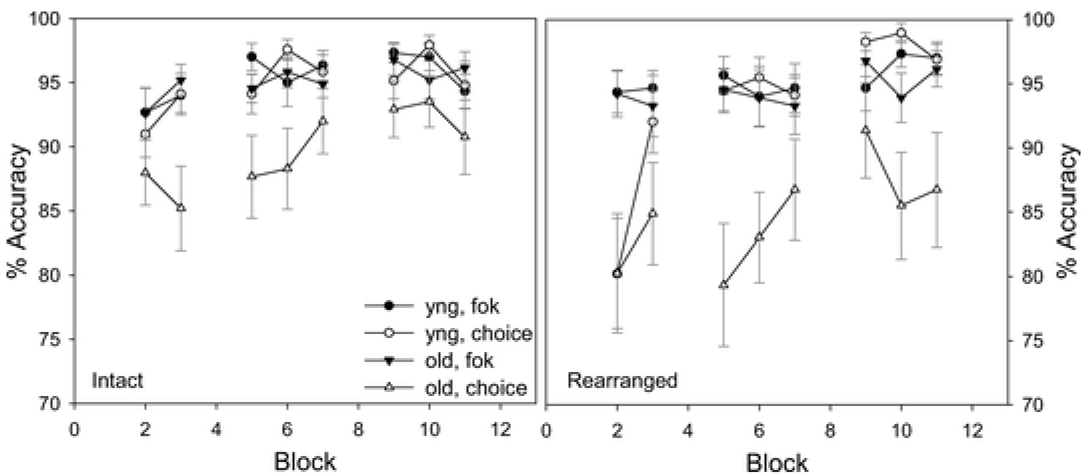
Decision Times in Choice and FOK conditions

Figure 3 also shows the RTs for FOKs and strategy choices. The pattern of results provides additional evidence that FOKs and strategy choices behave differently, especially between the two age groups. There were large age differences in RT, $F(1, 98) = 85.25$, $p < 0.001$, $d = 1.44$, and RT improved over blocks, $F(7, 686) = 69.01$, $p < 0.001$. The most interesting effect was a disordinal Age × Condition interaction, $F(1, 98) = 4.56$, $p < 0.05$. For younger adults, strategy choices were made more quickly than FOKs. Older adults showed a greater separation of FOK and choice RTs than younger adults, predominantly because of much slower strategy choices. This pattern persisted despite the speed-up of RT with practice. Decisions were also slower for rearranged than for intact pairs, $F(1, 98) = 41.44$, $p < 0.001$, $d = 0.13$. These outcomes suggest that older adults were more reluctant to make a rapid choice to retrieve.

Accuracy in Choice and FOK conditions

An important question was whether the accuracy of match-mismatch decisions in the Choice condition differed for the two age groups. Older adults were less likely to choose to retrieve. Were they able to respond accurately when they did so?

Accuracy was high in the FOK condition for both age groups (see Figure 4), but there was an apparent drop in accuracy in the choice condition relative to the FOK condition, where early decisions were required, $F(1, 98) = 8.58$, $p < 0.01$, $d = 0.37$, and this effect interacted with age, $F(1, 98) = 5.18$, $p < 0.05$, $d_{yng} = 0.12$, $d_{old} = 0.53$. For both intact and rearranged pairs, older adults were less accurate in the Choice condition, despite taking longer to make such choices.



We further analyzed the Choice condition accuracy data by whether individuals had chosen scanning or retrieval. Accuracy was relatively high for younger adults in all cases (> 90%), but older adults were less accurate when choosing retrieval for either intact pairs ($M = 84$) or rearranged pairs ($M = 72$), manifested in an age main effect, $F(1, 49) = 15.08$, $p < 0.001$, an Age × Choice interaction, $F(1, 44) = 17.16$, $p < 0.001$, and an Age × Match × Choice interaction, $F(1, 41) = 5.18$, $p < 0.05$. Accuracy increased over blocks, more so for rearranged pairs and after retrieval choices.

The accuracy data provide an important frame for interpreting older adults' lower retrieval choice behavior. When older adults ventured a retrieval choice, they were somewhat less likely to experience success when choosing to retrieve. This pattern is inconsistent with a pure retrieval avoidance. Instead it suggests that older adults more fragile associative learning may have inhibited their choice of the retrieval strategy. On the other hand, younger adults' apparent disposition to choose retrieval, noted above with respect to higher retrieval choices than FOK confidence for rearranged pairs, was not accompanied by a cost in decision accuracy.

Recognition Memory Trials

Recognition memory trials (with no lookup table provided) were given in Blocks 4, 8, and 12. We used these data to evaluate several aspects of our participants' learning of the new associations, including their metacognitive accuracy in judging their recognition responses. We analyzed all dependent variables as a function of Condition (Control, Choice, FOK), Match (Intact, Rearranged), and Block.

Recognition response times

The RTs for recognition memory trials can be conceptualized as a boundary condition of how fast RT in the other conditions could be if scanning were avoided. Table 2 reports the RTs for the recognition trials of Blocks 4, 8, and 12. Besides the typical age-related slowing in retrieval RT, $F(1, 144) = 170.01$, $p < 0.001$, $d = 1.57$ (e.g., Hertzog, Touron, & Hines, 2007), RT improved over blocks, $F(1, 288) = 122.12$, $p < 0.001$), more so for older adults, $F(1, 288) = 22.76$, $p < 0.001$, and more so for rearranged trials, $F(1, 288) = 67.68$, $p < 0.001$. In the first block of recognition trials, rearranged pairs took older adults about 800 ms longer to correctly judge than intact pairs, $d = 0.34$. There was a reliable Condition × Block effect, $F(1, 288) = 3.06$, $p < 0.05$, and a reliable Age × Condition by block effect, $F(1, 288) = 3.32$, $p < 0.01$. These interactions appeared to be mostly attributable to older controls responding more slowly in Block 4 ($M_{control} = 4040$, $M_{fok} = 3203$, $M_{choice} = 3455$), with this difference dissipating thereafter (block 8: $M_{control} = 2816$, $M_{fok} = 2591$,$M_{choice} = 2886$). The effect could have occurred because the response deadlines imposed in the FOK and Choice conditions led to a slightly less conservative response criterion in the first recognition test block.

Table 2 is omitted from this formatted document.

Accuracy

Table 2 also reports marginal means and standard errors for recognition memory accuracy aggregated into Age × Match × Block cells. Older adults were less accurate overall, $F(1, 144) = 68.89$, $p < 0.001$ (marginal $M_{Old} = 78.6$, marginal $M_{Young} = 92.8$; $d = 0.86$). Memory performance was worst for rearranged pairs, $F(1, 144) = 30.02$, $p < 0.001$, and this effect interacted with age, $F(1, 144) = 16.59$, $p < 0.01$. Rearranged pairs produced larger mean age differences ($d_{intact} = 0.72$, $d_{rearranged} = 1.0$). Recognition accuracy improved over blocks, $F(2, 288) = 102.50$, $p < 0.01$), with greater improvements by older adults, $F(2, 288) = 16.35$, $p < 0.01$. The reliable Age × Match × Block interaction, $F(2, 288) = 8.33$, $p < 0.01$), was attributable to slower initial learning by older adults that was exacerbated for rearranged pairs, followed by a narrowing of these differences with practice.

Conditional probability of retrieval given accurate recognition memory

In our previous work we showed reliable age differences in the likelihood of choosing to retrieve in a standard NP trial, given a correct recognition memory response on a preceding recognition memory test with this probe. This index was taken as a measure of retrieval reluctance (e.g., Touron & Hertzog, 2004b).

We used recognition memory performance to compute these conditional probabilities of retrieval, separately for intact and rearranged trials. Given the nature of the design, it was possible to compute the conditional probability separately for (1) subsequent standard NP trial retrieval reports in all three conditions (Control, FOK, Choice), and (2) retrieval reports on FOK trials or retrieval choices on Choice trials. Data were available for blocks 5 and 9 only (when NP trials followed blocks of recognition memory tests in blocks 4 and 8, respectively).

For standard NP trials, the Age × Condition × Match × Block (Block 5 versus 9) general linear model analysis revealed robust main effects of Age [$F(1, 144) = 48.99$, $p < 0.001$, $d = 1.10$], Block [$F(1, 144) = 54.53$, $p < 0.001$, $d = 0.23$], and Match [$F(1, 97) = 8.89$, $p < 0.01$, $d = 0.23$] (see Figure 5). Older adults were less likely to retrieve after successful recognition, the probability of a retrieval improved between blocks 5 and 9, and the retrieval probability was also lower for rearranged versus intact pairs. However, Match also interacted with age, $F(1, 144) = 4.67$, $p < 0.05$, with larger age differences on rearranged pairs ($M_{young,intact} = 0.83$, $M_{young,rearr} = 0.82$, $_{Mold,intact} = 0.51$, $M_{old,rearr} = 0.43$). Match also interacted with Condition, $F(2, 144) = 5.16$, $p < 0.01$, with differences being larger for the control and FOK conditions relative to the Choice condition, $p < 0.01$. The effects involving condition suggested a possible reactive effect of requesting choices on retrievals after a correct recognition response.

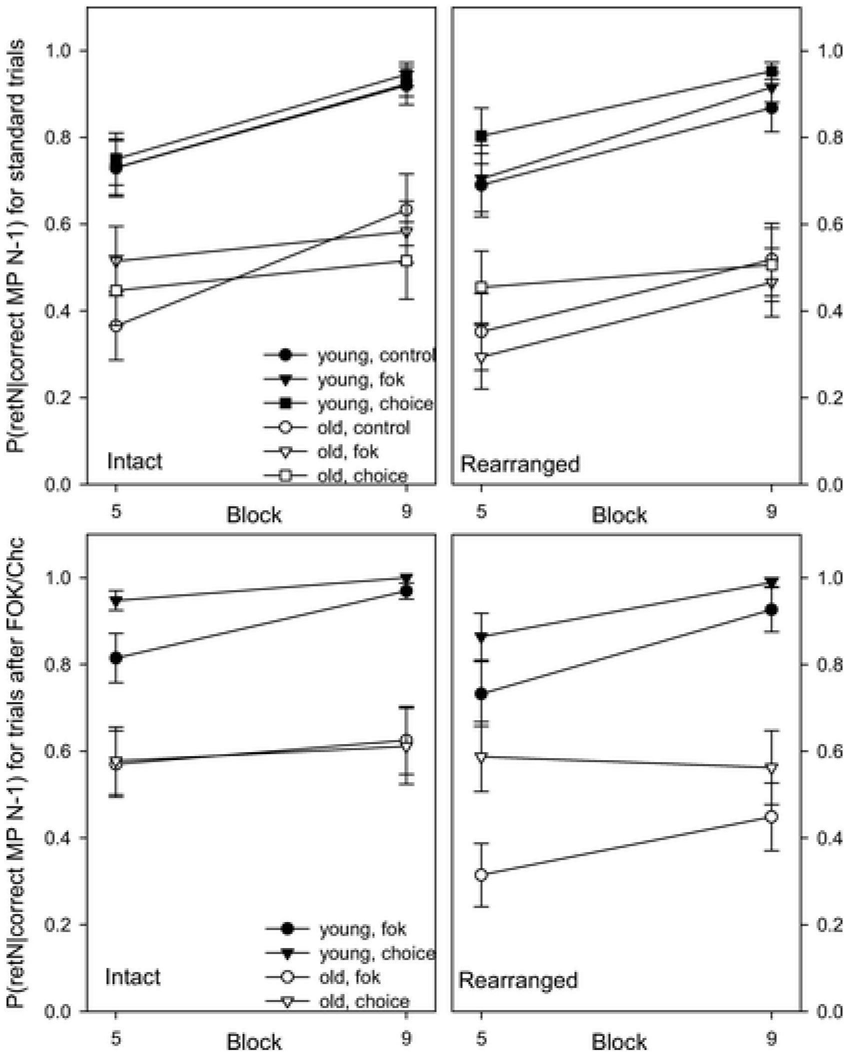


Figure 5. Probability of retrieval for NP trials after accurate recognition memory on test trials on a given item, as a function of block by age and condition.

The conditional probability of a retrieval given a correct prior recognition was also analyzed after FOK and Choice trials. Again, robust age differences were observed on these trials, $F(1, 97) = 41.82$, $p < 0.001$, $d = 1.10$, with older adults showing lower likelihood of choosing to retrieve on FOK trials or Choice trials. There were also reliable main effects of Block [$F(1, 97) = 13.94$, $p < 0.001$, $d = 0.23$] and Match [$F(1, 97) = 16.50$, $p < 0.001$, $d = 0.24$]. The conditional probabilities increased over blocks and were lower for rearranged pairs. The Match effect interacted with Condition, $F(1, 97) = 7.38$, $p < 0.01$, and the associated three-way interaction with Age was reliable, $F(1, 97) = 5.41$, $p < 0.05$. Retrieval was less likely for rearranged pairs in the FOK condition, and this effect was driven entirely by the older adults (see Figure 5).

Taken together, these conditional probability outcomes indicate that older adults were far less likely to choose the retrieval strategy after a successful recognition, particularly when the new trial involved a rearranged pair. Interestingly, it appeared that whether an item was intact or rearranged had little effect in the choice condition but a larger effect in the FOK condition, showing again the distinction between FOK and Choice. One potential explanation of this effect is that older adults in the Choice condition are more likely to make a risky retrieval choice on rearranged items by responding on the basis of associative familiarity, whereas older adults in the FOK condition were more likely to later resolve their initial uncertainty on rearranged items by scanning the available look-up table. Nevertheless, the larger point is that older adults manifest retrieval reluctance throughout practice in all conditions relative to young adults.

Confidence judgments

We analyzed recognition CJs with the same model as used for memory performance. Condition yielded no main effect nor any associated interaction, $F < 1$. Table 2 reports the marginal means and standard errors for CJs, aggregated in the Age × Match × Block cells. Older adults reported lower confidence after recognition memory test responses, $F(1, 141) = 66.82$, $p < 0.01$, marginal $M_{Old} = 67.1$, marginal $M_{Young} = 94.2$, $d = 1.20$. As with accuracy, confidence was lower for rearranged pairs, $F(1, 141) = 21.51$, $p < 0.001$, and this difference interacted with age, $F(1, 141) = 10.11$, $p < 0.01$. As memory improved over blocks, so did confidence, $F(2, 292) = 48.19$, $p < 0.01$, but no other interactions were reliable, $p > 0.05$.

We also evaluated the resolution of CJs for recognition memory accuracy, as assessed by Goodman-Kruskal gamma correlations (Nelson, 1984). These within-person ordinal correlations assess whether increases in confidence are associated with increases in recognition memory performance. Gammas were much lower for older adults compared with younger adults, $F(1, 121) = 30.23$, $p < 0.01$ ($M_{Old} = .30$, $M_{Young} = 0.71$, $d = 0.72$) and were lower for rearranged pairs than intact pairs, $F(1, 121) = 5.0$, $p < 0.05$, $M_{rearr} = 0.43$, $M_{intact} = 0.58$, $d = 0.27$. No other effect was reliable, although the Age × Match interaction approached significance, $F(1, 121) = 3.07$, $p < 0.08$. The trend indicated little difference in the correlation of CJs with memory outcomes between intact and rearranged pairs for the younger adults but a substantial difference between types of pairs for the older adults ($M_{old, intact} = 0.43$, $M_{yng, intact} = 0.72$, $M_{old, rearr} = 0.17$, $M_{yng, rearr} = 0.69$).

What causes this age difference in the accuracy of recognition memory CJs? It appears that older adults have particular difficulty assessing rearranged (or mismatched) NPs. For rearranged pairs, older adults manifested (1) lower recognition memory performance, (2) lower CJs, and (3) lower CJ resolution. The yes-no nature of the associative recognition test used here may increase the importance of recollection, including recall-to-reject or accept strategies (Cohn, Emrich, & Moscovitch, 2008). Older adults are known to be less likely to experience recollection during recognition tests (Light et al., 2000). Older adults are also more prone to misrecollection effects (i.e., high confidence recognition memory errors; Dodson, Bawa, & Krueger, 2007; Shing, Werkle-Bergner, Li, & Lindenberger, 2009). However, age differences in recognition memory CJ resolution are not consistently found (e.g., Hines, Touron, & Hertzog, 2009; Kelley & Sahakyan, 2003), and the conditions under which they occur have not yet been clearly understood. In the present case, the fact that the NP task does not focus on memory, making it in effect an incidental memory paradigm, may also contribute to the age differences in the accuracy of retrieval monitoring.

Age differences in the accuracy of CJs in the NP task contradicts Touron and Hertzog's (2004b; Experiment 3) previous finding of age equivalence in CJ resolution in the NP task. In that experiment, recognition memory tests were continuously presented throughout practice. Including continuous recognition memory tests also dramatically speeds the retrieval shift in older adults (e.g., Rogers & Gilbert, 1997). It could do so because (1) testing enhances learning (e.g., Roediger & Karpicke, 2006) or (2) memory testing raises explicit awareness in older adults that they have learned the items (Touron & Hertzog, 2004b). Given that Touron and Hertzog (2004b; Experiment 1) found that probed and unprobed items resulted in similar retrieval shift rate enhancements, we argue the effect is more likely to be attributable to a general awareness of the possibility of retrieval-based success. Unlike Touron and Hertzog (2004b), the present study deferred recognition memory trials until Block 4, which may have hindered older adults' awareness that levels of learning were sufficient to afford use of the retrieval strategy.

Consistent with this argument, at the end of practice older adults report much lower confidence in their ability to base NP responses on memory retrieval. Rated confidence in successful reliance on memory for NP trials was lower for older adults ($M = 51.5\%$, $SE = 2.0$) than for younger adults ($M = 88.2\%$, $SE = 2.1$), $d = 1.46$, $F(1, 145) = 160.51$, $p < 0.001$. Older adults' lower confidence in retrieval strategy use could be a direct outcome of inaccurate trial-level confidence in their recognition memory, a reflection of lower memory self-efficacy, in general (see Hertzog & Hultsch, 2000), or both. In any case, it is clear that older adults are not confident about their memory accuracy in the NP task.

The functional impact of confidence on NP strategy use

The pattern of effects for CJs was meaningful in light of the differences between the FOK and Choice conditions in retrieval strategy use. Age differences in mean CJs and CJ resolution opened the possibility that age differences in retrieval shift would be related to low and or inaccurate confidence in the ability to correctly recognize the noun pairings.

To see whether CJs covaried with later retrieval strategy choices, we first compared CJs at Block T (4 or 8) with subsequent strategy reports on the next block (Block T + 1), aligning these continuous CJs with the conditional probability of retrieving, given accurate recognition memory, reported earlier. That is, we computed the mean CJ at Block T as a function of whether individuals scanned or retrieved on the next standard NP trial. The age difference in CJs was still reliable ($M_{young} = 89.3$, $M_{old} = 64.9$), $F(1, 143) = 52.85$, $p < 0.01$. Of greater interest, confidence was higher for items for which retrieval was next selected, relative to items which were later scanned ($M_{retrieval} = 81.4$, $M_{scan} = 69.1$), $F(1, 117) = 26.75$, $p < 0.01$. The mean CJ increased as a function of blocks, $F(1, 143) = 43.21$, $p < 0.01$ without affecting the difference between scanning or retrieving trials, with the overall degree of confidence increased more for older adults, $F(1, 143) = 3.96$, $p < 0.05$. There was no reliable Age × Retrieval interaction, $F < 1$.

These outcomes suggest that CJs have a functional impact on the retrieval selection behaviors of both age groups. Older adults were more likely to retrieve given higher levels of confidence in the accuracy of recognition memory responses, even though this effect tended to be smaller than observed for younger adults. However, older adults manifested lower accuracy of CJs, as measured by the gamma correlations. It appears that older adults' reduced ability to discriminate correct from incorrect recognition responses, along with the influence of confidence on retrieval choices, influences their retrieval avoidance.

As noted earlier, Lamson and Rogers (2008) argued that older adults were, for the most part, making valid choices about whether to retrieve or calculate in their multiplication problems. Our findings challenge this argument, at least as concerns the NP task. Lamson and Rogers based their conclusion on consistent patterns of mean differences in trial accuracy and the likelihood of retrieval choices; as shown here, these variables can dissociate. We argue, instead, that older adults' deficient retrieval monitoring early in NP practice, as manifested by lower CJ resolution, leads to an avoidance of the retrieval strategy, even when, later in practice, the level of incidental associative learning would support accurate and fast NP trial responses based on memory. Of course, additional experiments that measured both CJs and FOKs in the Lamson and Rogers' task would be needed to further test this hypothesis.

Decisions by probe accuracy

We also computed the probability of a retrieval choice or high confidence FOK given a correct response to the previous recognition memory test for the same item. This analysis was conducted for blocks 5 and 9, comparing choice and FOK responses to the item recognition probe accuracy in the preceding block. Table 3 reports these probabilities. Choice or FOK behavior after a correct recognition probe varied by age, $F(1, 97) = 10.49$, $p < 0.01$, $d = 0.56$, and the age difference interacted with condition, $F(1, 97) = 6.68$, $p < 0.01$. There was also a difference in the conditional probabilities as a function of match, $F(1, 97) = 26.75$, $p < 0.01$, $d = 0.35$, and this effect interacted with age, $F(1, 97) = 5.34$, $p < 0.05$ and condition, $F(1, 97) = 14.93$, $p < 0.01$. No other interaction was significant. To facilitate interpretation, we analyzed the mean conditional probabilities separately for the Choice and FOK conditions.

Table 3 is omitted from this formatted document.

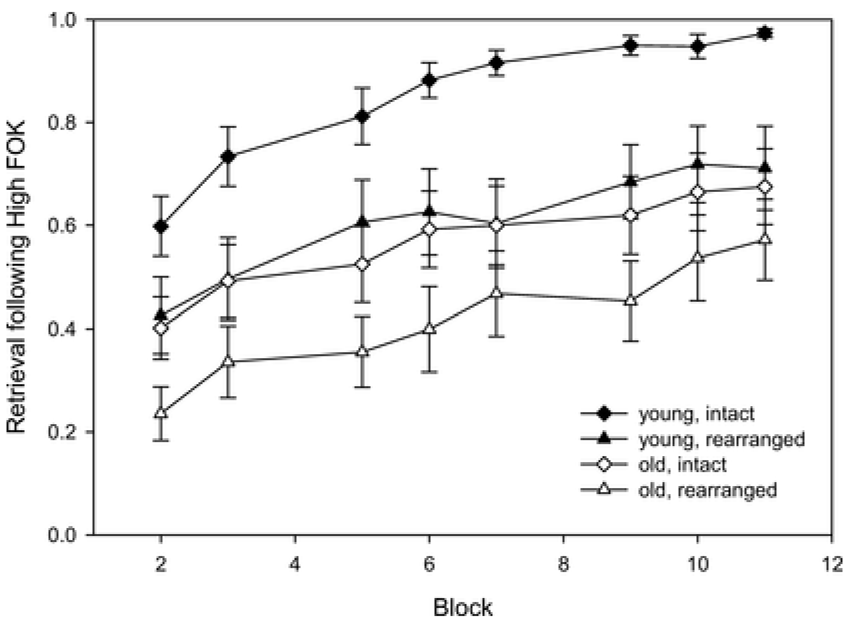
Older adults were less likely to make a retrieval choice after a correct probe, $F(1, 48) = 14.3$, $p < 0.01$, but the conditional probability of retrieval choice did not vary by Match or Block or with the interactions, ps > 0.15.

The conditional probability of a high FOK given correct probe response did not vary in the age main effect $F < 1$ but did vary by Match and with the Match × Age interaction, $F(1, 49) = 28.82$,

$p < 0.01$ and $F(1, 49) = 4.25$, $p < 0.05$, respectively. Young adults were more likely to report a high FOK after a correct probe response for intact items but not for rearranged items. Further comparisons were not reliable, $ps > 0.25$. The small increase in conditional probability for block 9 was not reliable, $F(1, 49) = 2.21$, $p > 0.10$, and the age by block interaction was again unreliable, $F < 1$.

Decisions by retrieval reports

For the FOK condition, we also computed the probability of a retrieval report for a trial immediately after experiencing a high confidence FOK for that item. Figure 6 shows the retrieval after high FOKs as a function of Match (intact or rearranged) and age. Older adults were less likely to report retrieval for trials when they had provided a high FOK, $F(1, 49) = 8.92$, $p < 0.01$ ($M_{yng} = .73$, $M_{old} = .49$, $d = 0.62$). The probability of retrieval was higher for intact ($M = 0.71$) than rearranged ($M = 0.51$) pairs, $F(1, 49) = 30.77$, $p < 0.01$, $d = 0.53$. The probability of retrieval after a high FOK increased reliably over blocks, $F(7, 343) = 26.53$, $p < 0.01$). The Age × Block × Match interaction was not reliable [three-way interaction $F(7, 343) = 1.51$, $p > 0.10$].



Taken together, the pattern of retrieval choices after FOKs is consistent with the retrieval reluctance hypothesis and is difficult to reconcile with a pure associative deficit. Older adults were less likely to use the retrieval strategy when they had a high FOK.

In general, the evidence for retrieval reluctance late in practice is consistent with our earlier demonstrations that monetary incentives hasten older adults' rates of retrieval shift (e.g., Touron, Swaim, & Hertzog, 2007). As noted earlier, such effects are inconsistent with a pure associative learning deficit. Touron, Swaim, & Hertzog (2007) provided monetary incentives after several blocks of standard NP trials. The outcomes of this study suggest that it would be interesting to see whether providing monetary incentives from the beginning of the experiment would reduce older adults' retrieval reluctance by providing a strong incentive to rely on memory retrieval even when they are uncertain about the accuracy of memory-based responding. On the other hand, it is also possible that the punishment experienced by failing to achieve monetary goals because retrieval failures early in practice would further inhibit older adults from attempting to rely on memory retrieval in the NP task (see West, Thorn, & Bagwell, 2003).

Cued Recall and Judgments of Learning

After training there were reliable age differences in cued recall for the NP stimuli, $F(1, 145) = 138.75$, $p < 0.001$. Younger adults had a much higher proportion cued recall of the NPs ($M = 0.88$, $SE = 0.02$) than older adults ($M = 0.50$, $SE = 0.02$), $d = 0.97$. This effect did not interact with Condition, $F(2, 145) = 1.21$, $p > 0.25$. In contrast to the results with CJs, the resolution of JOLs with recall, as measured by gamma correlations, did not differ between the age groups, $F < 1$, $M_{young} = 0.62$, $SE = 0.13$ versus $M_{old} = 0.60$, $SE = 0.08$, $d = 0.03$, a finding consistent with the aging literature on JOL accuracy (e.g., Hertzog, Dunlosky, Powell-Moman, & Kidder, 2002). This outcome shows that the age differences in the resolution of CJs reported earlier does not reflect a general memory monitoring deficit by older adults. It also demonstrates that age differences in resolution are not obligatory, given differences in mean levels of either cognition or metacognitive judgments.

Limitations and Future Directions

There were some limitations of the present study. First, we used only a soft deadline (which changed the font color of the stimuli) for FOKs and Choices. Perhaps partly as a result, the FOK and Choice RTs were relatively long, and one could argue we didn't assess FOKs at the point in time when retrieval choices are made in standard NP trials. We piloted a hard deadline, however, in which participants had to respond by the deadline or were informed they had timed out. We abandoned this approach after difficulty in setting a deadline that functioned well early in practice (when associative learning was most fragile). Future research with a revised deadline procedure could produce different results. Second, we treated FOK and Choice as between-subjects variables to minimize the complexity of our task and task instructions. It would be informative to evaluate situations where the same individuals were asked to generate FOKs or Choices (on different trials) to see whether the within-subjects manipulation would mirror the between-subjects effects seen here. Third, our sample size was not large enough to fully evaluate individual differences in variables that might have predicted difficulties with retrieval shift as a function of variables such as rearranged item CJ accuracy. Further work correlating between-person variation in monitoring with retrieval choices in the standard, FOK, and Choice conditions could help determine whether these differences are normative or specific to a subgroup of older adults, as well as what other characteristics might predict retrieval avoidance in older adults.

Conclusion

These data indicate that an associative deficit plays a strong role in older adults' delayed retrieval shift early in NP practice (as argued by Cerella, Onyper, & Hoyer, 2006), as manifested in slower rates of associative learning and lower accuracy by older adults when they chose to use the retrieval strategy. However, the data provide new evidence that deficits in metacognitive monitoring inside the NP task also contribute to slowed retrieval shift attributable to an avoidance of the retrieval strategy. Older adults had lower confidence in the accuracy of the recognition memory test responses as well as poor resolution of CJs with memory, indicating that they were less aware of whether their memory responses were accurate. Retrieval strategy choices in subsequent NP trials correlated with CJs. Finally, older adults were less likely to choose the retrieval strategy after a high FOK. The overall pattern of the data indicate that continued retrieval avoidance late in practice reflects, for at least some older adults, a top-down strategic choice to scan rather than to retrieve, with this choice being influenced by inaccurate monitoring of retrieval accuracy. Including the FOK and Choice conditions early in practice allowed us to demonstrate that the associative deficit and retrieval avoidance both operate to produce slower retrieval shift in older adults.

References

Ackerman, P. L., & Woltz, D. J. (1994). Determinants of learning and performance in an associative memory/substitution task: Task constraints, individual differences, volition, and motivation. *Journal of Educational Psychology*, *86*, 487– 515. doi: 10.1037/0022-0663.86.4.487

Bajic, D., & Rickard, T. C. (2009). The temporal dynamics of strategy execution in cognitive skill learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 113– 121. doi: 10.1037/a0013647

Botwinick, J. (1984). *Aging and behavior* (3rd ed.). New York: Springer.

Bourne, L. E., Raymond, W. D., & Healy, A. F. (2010). Strategy selection and use during classification skill acquisition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 500– 514. doi: 10.1037/a0018599

Cerella, J., Onyper, S., & Hoyer, W. J. (2006). The associative-memory basis of cognitive skill learning: Adult age differences. *Psychology and Aging*, *21*, 483– 499. doi: 10.1037/0882-7974.21.3.483

Cohn, M., Emrich, S. M., & Moscovitch, M. (2008). Age-related deficits in associative memory: The influence of impaired strategic retrieval. *Psychology and Aging*, *23*, 93– 103. doi: 10.1037/0882-7974.23.1.93

Cohn, M., & Moscovitch, M. (2007). Dissociating measures of associative memory: Evidence and theoretical implications. *Journal of Memory and Language*, *57*, 437– 454. doi: 10.1016/j.jml.2007.06.006

Connor, L. T., Balota, D. A., & Neely, J. H. (1992). On the relation between feeling of knowing and lexical decision: Persistent subthreshold activation or topic familiarity?*Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 544– 554. doi: 10.1037/0278-7393.18.3.544

Dodson, C. S., Bawa, S., & Krueger, L. E. (2007). Aging, metamemory, and high-confidence errors: A misrecollection account. *Psychology and Aging*, *22*, 122– 133. doi: 10.1037/0882-7974.22.1.122

Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). *Manual for kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.

Hertzog, C., Dunlosky, J., & Sinclair, S. M. (2010). Episodic feeling-of-knowing resolution derives from the quality of original encoding. *Memory & Cognition*, *38*, 771– 784. doi: 10.3758/MC.38.6.771

Hertzog, C., & Hultsch, D. F. (2000). Metacognition in adulthood and aging. In Salthouse, T. & Craik, F. I. M. (Eds.) *Handbook of Aging and Cognition* (2nd ed., pp. 417– 466). Mahwah, NJ: Erlbaum.

Hertzog, C., Dunlosky, J., Powell-Moman, A., & Kidder, D. P. (2002). Aging and monitoring associative learning: Is monitoring accuracy spared or impaired?*Psychology and Aging*, *17*, 209– 225. doi: 10.1037/0882-7974.17.2.209

Hertzog, C., Raskind, C. L., & Cannon, C. J. (1986). Age-related slowing semantic information processing speed: An individual differences analysis. *Journal of Gerontology*, *41*, 500– 502.

Hertzog, C., Touron, D. R., & Hines, J. C. (2007). Does a time monitoring deficit contribute to older adults' delayed shift to retrieval during skill acquisition?*Psychology and Aging*, *22*, 607– 624.

Hines, J. C., Touron, D., & Hertzog, C. (2009). Metacognitive influences on study time allocation in an associative recognition task: An analysis of adult age differences. *Psychology and Aging*, *24*, 462– 475. doi: 10.1037/a0014417

Kelley, C. M., & Sahakyan, L. (2003). Memory, monitoring, and control in the attainment of memory accuracy. *Journal of Memory and Language*, *48*, 704– 721. doi: 10.1016/S0749-596X(02)00504-1

Koriat, A., & Levy-Sadot, R. (2001). The combined contributions of the cue-familiarity and accessibility heuristics to feelings of knowing. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *27,* 34– 53.

Lamson, N., & Rogers, W. A. (2008). Assessing age-related patterns in strategy selection on a mathematical problem-solving task. *Journal of Gerontology: Psychological Sciences,63B*, P146–P155.

Light, L. L., Prull, M. W., La Voie, D. J., & Healy, M. R. (2000). Dual-process theories of memory in old age. In T. J.Perfect & E. A.Maylor (eds.) , *Models of cognitive aging* (238–300). London: Oxford University Press.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*, 492– 527. doi: 10.1037/0033-295X.95.4.492

Metcalfe, J., Schwartz, B. L., & Joaquim, S. G. (1993). The cue-familiarity heuristic in metacognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *19*, 851–861. doi: 10.1037/0278-7393.19.4.851

Microsoft Visual Studio (2007). *Microsoft. Net Framework Version 3.5 SP1*. Microsoft Corporation.

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. doi: 10.1037/0278-7393.26.5.1170

Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin*, *95*, 109– 133. doi: 10.1037/0033-2909.95.1.109

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H.Bower (ed.) , *The psychology of learning and motivation* (*Vol. 26*, pp. 125– 141). New York: Academic Press.

Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review*, *104*, 266– 300. doi: 10.1037/0033-295X.104.2.266

Reder, M., & Ritter, F. E. (1992). What determines initial feeling of knowing? Familiarity with question terms, not with the answer. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 435– 451. doi: 10.1037/0278-7393.18.3.435

Rickard, T. C. (1997). Bending the power law: A CMPL theory of strategy shifts and the automatization of cognitive skills. *Journal of Experimental Psychology: General*, *126*, 288– 310. doi: 10.1037/0096-3445.126.3.288

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, *17*, 249– 255. doi: 10.1111/j.1467-9280.2006.01693.x

Rogers, W. A., & Gilbert, D. K. (1997). Do performance strategies mediate age-related differences in associative learning?*Psychology and Aging*, *12*, 620– 633. doi: 10.1037/0882-7974.12.4.620

Rogers, W. A., Hertzog, C., & Fisk, A. D. (2000). An individual differences analysis of ability and strategy influences: Age-related differences in associative learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 359– 394. doi: 10.1037/0278-7393.26.2.359

Schunn, C. D., Reder, L. M., Nhouyvanisvong, A., Richards, D. R., & Stroffolino, P. J. (1997). To calculate or not to calculate: A source activation confusion model of problem familiarity's role in strategy selection. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 3– 29. doi: 10.1037/0278-7393.23.1.3

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, *84*, 127– 190. doi: 10.1037/0033-295X.84.2.127

Shing, Y. L., Werkle-Bergner, M., Li, S-C., & Lindenberger, U. (2009). Committing memory errors with high confidence: Older adults do but children don't. *Memory*, *17*, 169– 179. doi: 10.1080/09658210802190596

Singer, M., & Tiede, H. L. (2008). Feeling of knowing and duration of unsuccessful memory search. *Memory & Cognition*, *36*, 588– 597. doi: 10.3758/MC.36.3.588

Spieler, D. H, Mayr, U., & LaGrone, S. (2006). Outsourcing cognitive control to the environment: Adult age differences in the use of task cues. *Psychonomic Bulletin & Review*, *13*, 787– 793.

Touron, D. R., & Hertzog, C. (2004a). Distinguishing age differences in knowledge, strategy use, and confidence during strategic skill acquisition. *Psychology and Aging*, *19*, 452– 466. doi: 10.1037/0882-7974.19.3.452

Touron, D. R., & Hertzog, C. (2004b). Strategy shift affordance and strategy choice in young and older adults. *Memory & Cognition*, *32*, 298– 310.

Touron, D. R., & Hertzog, C. (2009). Age differences in strategic behavior during a computation-based skill acquisition task. *Psychology and Aging*, *24*, 574– 585. doi: 10.1037/a0015966

Touron, D. R., Swaim, E., & Hertzog, C. (2007). Moderation of older adults' retrieval reluctance through task instructions and monetary incentives. *Journal of Gerontology: Psychological Sciences*, *62B*, 149– 155.

West, R. L., Thorn, R. M., & Bagwell, D. K. (2003). Memory performance and beliefs as a function of goal setting and aging. *Psychology and Aging*, *18*, 111– 125. doi: 10.1037/0882-7974.18.1.111

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*, 441– 517. doi: 10.1006/jmla.2002.2864