View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by University of Richmond

University of Richmond
UR Scholarship Repository

Psychology Faculty Publications | Psychology

8-2016

The Role of Encoding Strategy in Younger and Older Adult Associative Recognition: A Think-Aloud Analysis

Mark C. Fox

Zachary Baldock

Sara P. Freeman

Jane M. Berry
University of Richmond, jberry@richmond.edu

Follow this and additional works at: http://scholarship.richmond.edu/psychology-faculty-publications

Part of the Applied Behavior Analysis Commons

This is a pre-publication author manuscript of the final, published article.

Recommended Citation

Mark C.Fox, Zachary Baldock, Sara P. Freeman, and Jane M. Berry. "The Role of Encoding Strategy in Younger and Older Adult Associative Recognition: A Think-Aloud Analysis." *Psychology and Aging*, *31*(5) (August 2016), 471-480. http://dx.doi.org/10.1037/pag0000111

This Post-print Article is brought to you for free and open access by the Psychology at UR Scholarship Repository. It has been accepted for inclusion in Psychology Faculty Publications by an authorized administrator of UR Scholarship Repository. For more information, please contact scholarshiprepository@richmond.edu.

The Role of Encoding Strategy in Younger and Older Adult Associative Recognition: A Think-Aloud Analysis

Mark C. Fox

Zachary Baldock

Sara P. Freeman

Jane M. Berry

Department of Psychology, University of Richmond

Revised and resubmitted, *Psychology and Aging*

Final copy published in *Psychology and Aging*, 31(5) (August 2016), 471-480.

Authors' Note

Correspondence should be sent to Mark C. Fox, Department of Psychology, University of Richmond, Richmond, VA 23173. Email: mfox23@richmond.edu.

The authors acknowledge the MacEldin Dunn Trawick Postdoctoral Fellowship (first author) and Professorship in Psychology (fourth author) for support of this study.

Abstract

Older adults have especially poor recognition memory for word pairs, and recent research suggests this associative deficit manifests primarily in older adults' higher rates of false alarms compared to younger adults. This could result from older adults either failing to generate meaningful (deep) mediators at study, or failing to benefit from having generated deep mediators at test. Younger and older adults performed a recognition memory task for words and word-pairs. A think-aloud analysis of their spontaneous encoding strategies (e.g., repetition, shallow mediators, and deep mediators) revealed that generation of deep mediators did not differ between younger and older adults, and was associated with high hit rates for items and associates in both age groups. However, generation of deep mediators was inversely related to false alarm rates in younger adults but not older adults. A trial-level analysis of encoding strategies and recognition responses revealed that younger adults benefited from having generated deep mediators when presented with corresponding recombined pairs at test as shown in their lower false alarm rates. In contrast, older adults who generated deep mediators during study (e.g., to *blanket-figure*) did not benefit from having done so when they encountered the corresponding recombined pairs at test (*blanket-summer* and *district-figure*): Their false alarm rates to pairs at test were unrelated to generation of deep mediators at study. These results suggest that many older adults have difficulty retrieving their mediators when presented with recombined pairs at test, older adults' mediators are not distinct enough to individuate intact pairs from recombined pairs at test, or some combination of both.

KEY WORDS: older adults, associative deficit, strategies, think-aloud, false-alarm

Word count: 6,926

The Role of Encoding Strategy in Younger and Older Adult Associative Recognition: A Think-Aloud Analysis

Older adults' impaired recognition memory for pairs of items such as words is a robust effect that generalizes across stimulus types (Berry, Williams, Usubalieva, & Kilb, 2013; Naveh-Benjamin, 2000; Naveh-Benjamin et al., 2009). After being presented with pairs of items (associates) such as *blanket-figure* at study, younger and older adults are similarly able to determine at test whether individual words (e.g., *blanket, pistol, dancer,* etc.) are old or new. However, older adults are less able than younger adults to determine at test whether pairs (e.g., *blanket-summer, pistol-dancer, district-figure*, etc.) are old or new.

This associative deficit is often attributed to an age-related deficit in binding information in memory (Naveh-Benjamin, 2000). Although this binding deficit is generally assumed to occur even when younger and older adults use the same strategies (Kuhlmann & Touron, 2012), some have suggested that impaired associative recognition of older adults is at least partly attributable to differences in encoding strategies, in particular, that older adults do not generate deep associations between items or fail to benefit from having generated them (Bender & Raz, 2012; Dunlosky, Hertzog, & Powell-Moman, 2005; Hertzog, Fulton, Mandviwala, & Dunlosky, 2013; Naveh-Benjamin, Brav, & Levy, 2007).

Deep associations or mediators are associations between two items in a pair that utilize the meanings of both items. For example, a participant presented with the pair *blanket-figure* may concoct an elaborate mental image of a human form wrapped in a blanket. Deep mediators can be distinguished from non-associative encoding strategies such as repetition of the pair. They can also be distinguished from relatively shallow mediators that utilize more concrete features of stimuli such as their letters (Craik, 2002). One reason deep mediators are especially beneficial is that deep encodings tend to be more distinct than shallow encodings (Gallo, Meadow, Johnson, & Foster, 2008). A deep mediator retrieved at test is, in theory, able to individuate intact pairs (e.g., *blanket-figure*) from recombined lures (*district-figure*). For example, the mental image of a human form wrapped in a blanket – a deep mediator – applies to only a small range of possible pairs that may be presented at test. In contrast, the observation that both *blanket* and *figure* contain the letter *e* or that both have two syllables – shallow mediators – applies to many possible pairs. When the intact pair, *blanket-figure* is presented at test, the deep mediator of a figure wrapped in a blanket, if retrieved, is distinct enough to verify that the pair is old. In contrast, the shallow observation that both items have two syllables, if retrieved, is not sufficient to verify that the pair is old as many recombined pairs (including *district-figure*) will contain two words with two syllables.

Craik (1977; Craik & Simon, 1980) theorized that older adults rely on shallower encoding strategies than younger adults, which means that they generate fewer deep mediators and rely more on shallow mediators or no mediators. An alternative possibility is that older adults do generate deep mediators at study but do not benefit from having done so at test (Hertzog?). This could be the case if, for example, older adults are less able than younger adults to retrieve their mediators (Dunlosky & Hertzog, 1998). Alternatively, older adults may not benefit from having generated mediators if they fail to “decode” them at test, that is, if they retrieve their mediators but are nevertheless unable to use them as a basis for determining whether a pair is old or new (Dunlosky et al., 2005; Hertzog et al., 2013). Thus, the nature of the relationship between aging, strategy, and retrieval is an important issue that goes beyond the question of whether younger and older adults happen to use similar strategies.

Generation of Mediators in Younger and Older Adults

Studies examining spontaneous use of strategies, via self-reports of strategy, reveal relatively few differences between younger and older adults. Dunlosky and Hertzog (1998) found that although encoding strategies differ from person to person within age groups, and these differences do contribute to individual differences in memory performance, they do not contribute substantially to differences between age groups in memory performance. In particular, there was no evidence that younger adults are more prone than older adults to generating meaningful mediators. Likewise, Kuhlmann and Touron (2012) reported no differences between younger and older adults in the strategies used to remember associations between words and fonts or words and locations.

These studies suggest that impaired performance in older adults arises in part because mediators generated at study don't benefit memory performance at test (Dunlosky et al., 2005; Kuhlmann & Touron, 2012). For example, Hertzog et al. (2013) found that even when older adults retrieve their mediators at test, the fidelity of the retrieved mediators is often compromised relative to the original mediators (i.e., they are retrieved in a gist form relative to how they were originally verbalized). Mediators retrieved in this gist form are not as effective for determining whether a presented pair is old or new. Hertzog et al. also found that even when older adults retrieved verbatim mediators from study, they were less likely to produce the correct target in a cued recall test.

Accounting for Older Adults' Higher Rates of False Alarms

It is noteworthy that research on the use of mediators by younger and older adults has been conducted primarily in the context of cued recall (Dunlosky et al., 2005; Hertzog et al., 2013) rather than recognition (but see Patterson & Hertzog, 2010). Understanding how the failure to make use of mediators at retrieval affects recognition performance may require examining the relationship between particular strategies at study and the responses elicited at test in recognition memory specifically. We suggest that the strategies associated with false alarms in particular may provide clues to a potential mechanism for an associate learning deficit as response distributions reveal that older adult impairments are driven primarily by higher rates of false alarms (Bender, Naveh-Benjamin, & Raz, 2010; Brubaker & Naveh-Benjamin, 2014) rather than misses.

The idea that older adults fail to bind information effectively cannot be reconciled with the tendency of older adults to commit false alarms without positing some additional difference between younger and older adults in the processes giving rise to responses at test. After all, failure to bind information should give rise to misses rather than false alarms in the absence of an accompanying difference in process, such as a difference in standard of proof for declaring pairs old. The necessity of such a difference in process has been acknowledged, described, for example, as a liberal response bias in older adults (Bender et al., 2010), but the nature of this difference itself has not been identified in any detail.

One reason why existing research has not accounted for false alarms is that performance has traditionally been examined at a level of analysis that overlooks step-by-step mechanisms occurring at the trial level. Aggregating observations across trials to create person-level variables removes the sequences of cause and effect as they occur at the trial level. Identifying a mechanism to account for how the failure to bind information could lead to false alarms requires a different approach—one that examines performance as a sequence of processes, beginning with a particular strategy at study and culminating with a particular response at test.

In this paper, we examine younger and older adults' encoding strategies with think-aloud verbal reporting (Ericsson & Simon, 1980; Fox, Ericsson, and Best, 2011), a method that makes it possible to observe encoding strategies as they occur during study while minimally affecting thought processes. This method enables a detailed analysis of the relationship between encoding strategies and associative recognition performance in younger and older adults. The emphasis of this approach is on mechanisms rather than variables; that is, the goal is not to determine whether variation in strategies accounts for variation in performance per se, but rather to identify the specific sequences of processes that occur, beginning with a particular strategy at study and culminating with a specific response at test in both age groups. Such an approach not only elucidates the relationship between age group and strategies, but also identifies a mechanism to explain why older adult deficits in associative recognition manifest.

Observing Encoding Strategies with Think-Aloud Verbal Reports

Think-aloud (Ericsson & Simon, 1980, 1993) is a verbal reporting method that enables researchers to test theories of the psychological processes mediating performance of a cognitive task without relying on participants' introspections about their own thought processes (Fox et al., 2011). In think-aloud studies, participants are asked to give verbal expression to their "inner voice" as it occurs spontaneously while completing the task. These overt verbalizations are recorded and then classified according to *a priori* criteria as instances of various psychological processes or strategies. Think-aloud verbalizations (e.g., "The figure was surrounded by a blanket") make it possible to identify a particular strategy without interrogating the participant about a particular strategy (e.g., Ariel, Price, & Hertzog, 2015, Exp. 1b). Think-aloud verbalizations are instances of the strategy itself as distinct from participants' retrospective reports of the strategy. Ericsson and Simon (1980) provide a detailed theoretical account of the difference between think-aloud and introspection. Studies show that think-aloud is generally non-reactive under the conditions specified by Ericsson and Simon (1980), eliciting performance that is roughly equal to that observed in silent conditions (Ericsson & Simon, 1993; Fox et al., 2011). The present study featured a silent control group, which made it possible to verify that thinking-aloud did not alter recognition performance.

A think-aloud analysis of encoding strategies requires a system for classifying encoding strategies that can be implemented within the constraints imposed by both the associative deficit paradigm and the think-aloud method. Previous research by the first author (Fox & Charness, 2010) and pilot testing in our lab suggested that think-aloud reports generated during the study phase of the associative deficit paradigm can be expected to provide sufficient information to categorize verbalizations into three categories: repetition, shallow mediator, and deep mediator (Craik, 2002).

Repetition is a common and relatively shallow strategy for memorizing words and pairs of words that amounts to mentally rehearsing the stimulus repeatedly. Participants using this strategy repeat the stimulus over and over again in their minds.

Shallow mediators make use of the orthographic, phonemic, or other relatively concrete properties of the stimuli. For example, a participant presented with the pair, *summit-tourist*, may notice that both words have two syllables or both contain the letter *i* (or both). Similarly, he or she may combine parts of the words into a new word such as “sut” (an actual example from our data). In either case, the participant has managed to associate the two words in a way that is meaningful only with respect to relatively concrete properties of the stimulus, but not with respect to the meanings of the two words.

Deep mediators are associations that actually make use of the meanings of the two words. The image of a human form wrapped in a blanket (*blanket-figure*) or the idea of a coachman navigating through the darkness of night with a flashlight (*carriage-flashlight*) fall into this category. Participants may occasionally provide relatively detailed verbal traces of meaningful mediators, but it is more often the case that even very elaborate pictorial associations may elicit only brief descriptions (e.g., “there was a flashlight on the carriage.”). More specific criteria for classifying verbalizations into these categories are provided in the Method section.

We expected deep mediators to be better predictors of hit rates and false alarm rates (inversely) than other mediators. The theory that older adults rely on shallower processing (Craik, 1977; Craik & Simon, 1980) leads to the prediction that older adults should generate fewer deep mediators than younger adults. In contrast, the theory that older adults generate deep mediators but fail to benefit from having done so (Dunlosky & Hertzog, 1998; Dunlosky et al., 2005; Hertzog et al., 2013) leads to the prediction that younger and older adults generate similar mediators, and a weaker relationship between mediators and memory performance in older adults than in younger adults. In either case, trial-level data on strategies of younger and older adults make it possible to elucidate the relationship between strategy use and recognition at the process level..

Method

Design and Participants

The experiment utilized a mixed design with age group (young/old) and verbalization (think aloud/silent) as between-subjects factors, and test type (item, associative) as a within-subjects factor. There were 97 participants, 45 older adults (31 female) aged 58-91 (*M* = 73.18 years, *SD* = 7.88) and 52 younger adults (35 female) aged 18-24 (*M* = 18.85 years, *SD* = 1.07). Younger adults were students recruited from introductory psychology classes who received course credit for participation. Older adults were recruited from the surrounding community through newspaper ads and a database of participants from previous studies. They received $15 for participation. Demographic data and standardized scores on processing speed and vocabulary measures are reported in Table 1.

Materials

The stimuli consisted of common-noun words from the English Lexicon Project. Word stimuli had a mean length of 6.46 letters (range = 4-10, *SD* = 1.14), were 1-2 syllables, and were of medium-to-high frequency in the language, meaning they reached an average log-transformed HAL frequency of 8.68 (Balota et al., 2007; Lund & Burgess, 1996). We constructed and implemented 60 total pairs of words with additional words chosen to serve as practice stimuli and as lures in the item recognition tests. Lures were chosen to match target stimuli in word length and number of syllables, as well as HAL frequency. Word pairs were constructed so that they would avoid integrative relations (Badham Estes, & Maylor, 2012) and simple associations (e.g., *shoe-foot*). For the experimental stimuli, pairs were divided evenly into two blocks of 30 pairs of words. Lists were equated on measures of length and frequency. Four configurations of stimuli were created.

Two blocks of study and test measures were administered. Each block comprised a study phase followed by an item test and an associative test. Assignment of stimuli to test (item or associative) was counterbalanced across configurations. Item and associative test order was counterbalanced between participants; half of the participants were tested for item recognition first, and half for associative recognition first.

Procedure

Participants were brought into a silent testing room and seated at a computer. The study was introduced, informed consent obtained, and a demographic questionnaire administered.

Participants assigned to the think-aloud condition were familiarized with the think-aloud protocol (Ericsson & Simon, 1993). They were instructed to “speak out loud your inner voice as you would if you were alone in a room talking to yourself” but asked that they not “explain to me what you are doing or what you think you are doing,” nor that they “force yourself to speak if you have nothing to say.” An audio recorder placed on the table was used to collect verbal reports. Participants were asked to think aloud during the study phase and the recognition tests and were reminded to do so before each task. If participants fell silent for a lengthy period of time during either study or test, experimenters issued up to two additional verbal reminders to “please remember to think aloud.” Prompting was minimized in this fashion because continuous prompting threatens to change the task for think-aloud participants, leading them to focus on generating speech rather than on the primary memory task, thereby rendering the data invalid. One drawback of minimal prompting is that it may lead to incomplete data (participants do not verbalize on many trials). However, the assumption that trials accompanied by verbalization are representative of all trials can be tested by comparing performance on trials where strategies are verbalized to trials for which no strategies are verbalized. Participants assigned to the silent condition received no instructions to think-aloud and were not recorded.

The practice session, study phase, and item and associative recognition tests were programmed and run on E-Prime version 2.0. Participants worked on the Salthouse (1991) pattern comparison tasks as distractor tasks between the study and recognition test portions of each block.

The practice session was used to familiarize participants with the word stimuli and the associative recognition paradigm, and to practice thinking aloud for participants in that condition. To practice for the study phase, participants viewed six pairs of words and were instructed to try to memorize the words and the pairs to prepare for the upcoming practice tests. Each pair was presented on the computer screen for 5 seconds. After study, participants completed three pattern comparison items to practice the distractor task. Each pattern comparison item entailed indicating whether two patterns were the same or different. Then, to practice for the test phase, three pairs were presented one-by-one and three words were presented one-by-one. Participants were instructed to press “yes” on the keyboard if they had seen the word or pair at study, and press “no” if they had not seen it.

The study phase of block 1 consisted of 30 pairs of words and was followed by the pattern comparison task to prevent rehearsal. Participants had 20 seconds to complete as many of the pattern comparison items as possible. Participants then began either the item or associative recognition test. For the item test, participants were presented with 40 words, half of which came from the study list, and half of which were new (lures). For the associative test, participants were presented with 20 pairs, 10 of which were presented intact from the study list, and 10 of which were recombined by combining words that had appeared in previously studied pairs into new pair configurations (lures). No new words were used to create other possible new pair types (e.g., old-new, new-new pairs). The same procedure was used for block 2. After block 2, participants the digit-symbol substitution task and the vocabulary task. Participants were debriefed, thanked, and compensated or credited for their participation.

Coding of Think-Aloud Reports

Verbal reports from participants in the think-aloud condition were transcribed by three members of the lab and then coded on a trial by trial basis, with *trial* referring to each 5s period during which a word pair was presented. Each trial was assigned to one of five categories: (1) repetition, (2) shallow mediator, (3) deep mediator, (4) reading without repetition, and (5) nothing/missing/unclear. If part or all of the stimulus was verbalized at least twice, even if the second verbalization occurred during a subsequent trial, then a trial qualified as repetition. For a trial to be categorized as a shallow mediator, an association must have been verbalized that made use of the alphabetic, phonemic, or orthographic properties of both words comprising the stimulus. When an association was verbalized that made use of the meanings of both words comprising the stimulus, then a trial was categorized as a deep mediator. The three categories were treated as ordinal such that any trial that could be classified in multiple categories was classified in the highest of those multiple categories. For example, if a participant repeated an item from an early trial during a later trial, and the early trial had been coded as deep mediator, then the early trial was coded as a deep mediator. Two nominal categories were used to account for the remainder of responses (reading without repetition and missing). Two research assistants coded verbal reports. Their independent ratings of the 53 participants' two blocks of 30 pairs per trial (for a total of over 3,000 trials) yielded Cohen's $k = .91$. Discrepant ratings were resolved by consensus. Examples of repetition, shallow, and deep mediator strategies appear in Table 2.

Results

Recognition Performance

Hit rates (HRs) are derived from correct responses to old items and intact pairs, and false alarm rates (FARs) are derived from correct responses to (that is, rejections of) new items and recombined pairs. In the following analyses, data were collapsed across Blocks 1 and 2 because inclusion of block as a within-subjects factor in the three-way ANOVAs of interest, designed to test our hypotheses, yielded nonsignificant four-way interaction effects for HRs and FARs. Additionally, block did not interact with age group; the pattern of results within each block for the HRs and FARs by age group was unchanged by inclusion of block as a factor.

Although we distinguish between HRs and FARs to examine the role of encoding strategy, we first test for the associative deficit as it is usually observed by creating an overall recognition memory performance variable of HR minus FAR (Berry et al., 2013; Naveh-Benjamin, 2000). A 2(Age Group: Younger/Older) x 2(Verbalization: Silent/Think-Aloud) x 2(Test: Items/Associates) ANOVA revealed no main effect of Verbalization (Silent: *M* = .62, *SE* = .03; Think-Aloud: *M* = .63, *SE* = .03), but revealed main effects of Age Group (Younger: *M* =.72, *SE* = .03; Older: *M* = .53, *SE* = .03), $F(1, 91) = 20.82$, $p < .001$, $\eta_p^2 = .186$, and Test (Items: *M* = .67, *SE* = .02; Associates: *M* = .58, *SE* = .03), $F(1, 91) = 17.57$, $p < .001$, $\eta_p^2 = .162$, and an interaction between Age Group and Test, $F(1, 91) = 7.11$, $p = .009$, $\eta_p^2 = .072$. Younger adults had better overall recognition than older adults and, consistent with Ericsson and Simon’s (1980) predictions, thinking aloud had no effect on performance. Importantly, the general pattern of the associative deficit was observed as younger adults had somewhat better recognition for items than older adults (Younger: *M* = .74, *SE* = .03; Older: *M* = .60, *SE* = .03), but considerably better recognition for associates than older adults (Younger: *M* = .70, *SE* = .04; Older: *M* = .46, *SE* = .04).

In accord with our thesis that encoding strategies may have different effects on old/intact (hit outcomes) versus new/recombined (false alarm outcomes) stimuli at test, dependent variables of HR and FAR were analyzed separately. A 2(Age Group: Younger/Older) x 2(Verbalization: Silent/Think-Aloud) x 2(Test: Items/Associates) mixed ANOVA revealed no main effect of Age Group (Younger: *M* = .81, *SE* = .02; Older: *M* = .77, *SE* = .02), Verbalization (Silent: *M* = .79, *SE* = .02; Think-Aloud: *M* = .79, *SE* = .02), or Test (Items: *M* = .79, *SE* = .02; Associates: *M* = .79, *SE* = .02), and no interactions for HR. Hit rate as a function of Age Group and Test is shown in Figure 1.

In contrast, the same analysis for FAR revealed a main effect of Age Group (Younger: *M* = .09, *SE* = .02; Older: *M* = .24, *SE* = .02), $F(1, 91) = 31.88$, $p < .001$, $\eta_p^2 = .26$, a main effect of Test (Items: *M* = .12, *SE* = .01; Associates: *M* = .21, *SE* = .02), $F(1, 91) = 38.72$, $p < .001$, $\eta_p^2 = .30$, and an interaction between Age Group and Test, $F(1, 91) = 21.88$, $p < .001$, $\eta_p^2 = .19$. Once again, there was no main effect of Verbalization or interactions involving this variable (Silent: *M* = .17, *SE* = .02; Think-Aloud: *M* = .16, *SE* = .02). In keeping with other studies (Bender et al., 2010; Brubaker & Naveh-Benjamin, 2014), the data reveal that the associative deficit reflects a tendency of older adults to commit false alarms to recombined pairs, rather than a tendency to reject intact pairs. Younger and older adults had comparable rates of false alarms to new items (Younger: *M* = .08, *SE* = .01; Older: *M* = .10, *SE* = .02), but older adults had a considerably higher rate of false alarms to recombined pairs than younger adults (Younger: *M* = .16, *SE* = .01; Older: *M* = .32, *SE* = .03). False alarm rate as a function of Age Group and Test is shown in Figure 2.

Encoding Strategies as Revealed by Think-Aloud Reports

The absence of any effects of thinking aloud on performance satisfies the minimum necessary condition for inferring the validity of think-aloud verbal reports (Ericsson & Fox, 2011). Nonparametric statistics were used to analyze strategies revealed by think-aloud reports as the strategy variables were neither normally distributed nor expected to have a quantitative structure (Grice, 2011). Rank-based non-parametric tests are conservative, minimizing the likelihood of type-I error by minimizing the effects of outliers. Frequencies of use for repetitions, shallow mediators and deep mediators are presented in Table 3.

We tested for differences between younger and older adults in number of times each strategy was observed with Mann-Whitney tests, which compare the mean ordinal ranks of younger and older adults for number of uses (see Table 3). The test for repetitions was significant, $U = 176.00$, $p = .005$, as younger adults used repetition more frequently (range: 0-59) than older adults (range: 0-57). No effect of age group was observed for shallow mediators (younger range: 0-30; older range: 0-13) or deep mediators (younger range: 0-48; older range: 0-53). Younger adults were no more likely to use deep mediators than older adults.

Next we examined the relation between encoding strategies and HRs and FARs for items and pairs. If a particular strategy increases the likelihood of recognizing older items or pairs, it should be correlated with HR. Conversely, if a strategy reduces the likelihood of mistaking a new item or recombined pair for an old item or intact pair, it should be negatively correlated with FAR.

Neither number of repetitions nor number of shallow mediators was correlated with HR or FAR for items or pairs. Participants who use either of these strategies according to think-aloud reports performed no better or worse than participants who did not use these strategies. In the case of shallow mediators, this lack of correlation may be due at least in part to the low number of shallow mediators observed. In contrast to other strategies, number of deep mediators was correlated with HR for items, $r_s = .43$, $p = .001$, and associates, $r_s = .64$, $p < .001$. Number of deep mediators was not correlated with false alarms for items. However, number of deep mediators was inversely correlated with FAR for associates, $r_s = -.37$, $p = .004$. Participants who generated deep mediators at study tended to correctly identify recombined pairs as new.

Although both younger and older adults generate deep mediators, it is possible that older adults do not make use of encoding strategies at test as effectively as younger adults do. To test this hypothesis, we examined correlations between use of the various strategies and HRs and FARs within age groups. As Table 4 shows, number of deep mediators was correlated with HR for both items and associates in younger adults (items: $r_s = .50$, $p = .004$; associates: $r_s = .53$, $p = .002$) and older adults (items: $r_s = .37$, $p = .043$; associates: $r_s = .75$, $p < .001$). However, number of deep mediators was correlated with FAR for associates only in younger adults, $r_s = -.51$, $p = .003$, and not older adults, $r_s = -.13$, $p = .284$. Fisher's r-to-z transformation revealed a marginally significant difference between the sizes of these two coefficients, z = 1.449, p = 0.074 (Preacher, 2002, http://www.quantpsy.org/corrtest/corrtest.htm).

The absence of a correlation between number of deep mediators and false-alarm rate in older adults does not necessarily imply that older adults do not benefit from having generated deep mediators during study when presented with recombined pairs at test. It is possible that there simply isn't sufficient variation in older adult FAR to observe a between-subjects correlation. If older adults are less able than younger adults to use deep mediators generated during study to reject recombined pairs, then it should be possible to observe their failure to use deep mediators within participants, that is, at the trial level. Older adults should commit false alarms when presented with recombined pairs after having generated a deep mediator for corresponding pairs during the study phase.

To test this hypothesis, we did a trial-level analysis to determine how often using deep mediators to encode specific pairs (e.g., *blanket-figure*) results in false alarms when participants are presented with recombined pairs containing the same words (e.g., *blanket-summer* and *district-figure*). The hypothesis was supported. Younger adults were found to commit a false alarm after having encoded with a deep mediator on only six out of 208 (three-percent of) possible trials. In contrast, older adults committed false alarms after having encoded with a deep mediator on 23 out of 106 (22% of) trials. A multilevel logistic regression (with participant modeled as a random effect) revealed that this difference was statistically significant, $t(145) =$

2.91, $p = .004$. Younger adults who generated deep mediators at study nearly always declared relevant recombined pairs new at test, whereas older adults who generated deep mediators at study were more likely to declare recombined pairs old at test.

Results of the think-aloud analysis reported above are predicated on the assumption that trials for which focal strategies could be inferred (repetitions, shallow mediators, and deep mediators) are representative of performance in general, that is, that strategy trials are representative of the “no-strategy” trials shown Table 3. To put it differently, to accept these analyses is to assume that participants did in fact use similar strategies on the no-strategy trials but failed to verbalize them while doing so. One way to test this assumption is to conduct a within-subjects analysis comparing performance from all trials for which a strategy could be inferred to performance for all no-strategy trials. If similar strategies were used in both cases, performance should be similar. In particular, because analyses are aimed at making inferences about relations between age group and strategies, the variable representing strategy versus no strategy should not interact with age group.

To test the hypothesis of similar performance for strategy and no-strategy trials we created a trial-level data file to pair strategies observed at encoding (for, say, *blanket-figure*) with responses to relevant items and associates at test, as this allowed us to extract performance data for both types of trials. Analyses of both hits and false alarms can be conducted for associates because both of the words comprising a pair at test were presented at study. In contrast, for items, only an analysis of hits is possible. An analysis of false alarms for items could not be conducted because an item that was not presented at study has no encoding strategy at study. Thus, analyses for items and associates were conducted separately.

A series of 2(Age Group: Younger/Older) x 2(Strategy: Strategy/No-strategy trials) mixed ANOVAs revealed no effect of strategy or interactions between age group and strategy for item HR or associate FAR. For associate HR, there were no interactions between strategy and age group, although a marginal main effect of strategy was observed, $F(1, 31) = 4.30$, $p = .05$, $\eta_p^2 = .12$, as participants tended to have slightly higher HRs on trials in which strategies were verbalized (Strategy: $M = .88$, $SE = .03$; No strategy: $M = .83$, $SE = .04$). This suggests that during at least some of the no-strategy trials, some participants may have used none of the strategies identified, leading to somewhat lower performance. However, the absence of any interaction between age group and strategy suggests that the conclusions about strategy and age revealed by the analyses above are representative of performance in general.

Discussion

We examined two possible accounts of the relationship between encoding strategy and associative recognition in younger and older adults in an effort to understand why older adults perform worse than younger adults on associative recognition tests. One hypothesis was that low recognition performance is caused by older adults failing to generate deep mediators at study. The other was that older adults do generate deep mediators at study but nevertheless fail to benefit from having done so at test (Dunlosky & Hertzog, 1998). Our results provide support for the second hypothesis. A think-aloud analysis of spontaneous encoding strategies revealed that generation of deep mediators did not differ between younger and older adults, and predicted high HR for items and associates in both age groups. However, the absence of deep mediators was found to predict FAR for associates in younger adults only, and not older adults. Indeed, the magnitudes of the relation between deep mediators and pair recognition (HR) and rejection (FAR) were virtually equivalent in younger adults but quite divergent in older adults.

Our findings are consistent with previous research on strategies in the context of cued recall (Dunlosky & Hertzog, 2001; Dunlosky et al., 2005; Hertzog et al., 2013). In particular, we found no evidence that older adults are less prone than younger adults to generating meaningful mediators. Moreover, Dunlosky and Hertzog (1998) examined spontaneous encoding strategies, and found, as we did, that although encoding strategies differ from person to person, and these differences do contribute to individual differences in memory performance, they do not contribute substantially to differences between age groups in memory performance.

Between-subjects analyses revealed no differences in recognition performance, HR, and FAR between silent and think aloud conditions. These results are consistent with Fox et al.'s (2011) conclusion that think-aloud reports do not alter performance, and satisfy the most important condition for inferring that verbal reports are valid (Ericsson and Simon, 1980). The use of think-aloud methods to measure strategic behaviors at encoding and their effects on recognition memory at test is novel and represents a unique method for getting at explanatory mechanisms for age differences in recognition memory.

Use of Deep Mediators by Younger and Older Adults at Test

Our most notable finding occurred at a trial-level analysis of encoding strategies and recognition performance. Specifically, older adults who generated deep mediators at study had higher hit rates for intact pairs at test than older adults who did not. However, and more notably, older adults who generated deep mediators at study did *not* benefit from having done so when presented with relevant recombined pairs at test. This finding offers a preliminary glimpse at possible mechanisms responsible for the prominence of false alarms in older adult associative recognition. A disproportionate effect of false alarms has been observed in other studies that compare younger and older adults directly (e.g., Bender et al., 2010) and in simulations of the associative deficit in younger adults (Brubaker & Naveh-Benjamin, 2014), suggesting that trials in which participants are presented with recombined pairs seem to be especially important to understanding the associative deficit. The susceptibility of older adults to false alarms is such that even when the intact pair is presented with three recombined pairs in a four alternative forced choice format, older adults are still more likely than younger adults to choose a recombined pair, i.e., commit a false alarm (Patterson & Hertzog, 2010).

The finding that older adults who generate deep mediators at study (for, say, *blanket-figure*) tend to commit false alarms when presented with recombined pairs (*blanket-summer*) points to deficiencies in either mediator retrieval or mediator decoding in a framework provided by Dunlosky et al. (2005). The first possibility is that the individual items in recombined pairs (*blanket* and *summer*) are not sufficient to elicit retrieval of mediators generated at study for *blanket-figure* and *district-summer* for older adults. If this is the case, then older adults *must* rely on a different standard of proof than younger adults for declaring pairs old because, all things equal, the inability to retrieve a mediator should lead one to declare recombined pairs new, not old. One example of such a difference in standard would be if younger adults do not declare pairs old unless they retrieve a mediator, but that older adults are willing to declare pairs old due to the mere familiarity of the individual items (see Jacoby, 1991). In fact, this difference might be expected if some older adults are seldom or never able to retrieve mediators when presented with recombined pairs as one cannot make use of a metacognitive cue that is never available.

The second possibility is a decoding deficiency. By this account, individual items in recombined pairs *do* elicit retrieval of mediators in older adults, but older adults then have trouble decoding the mediators. One way this deficiency could arise is if older adults' mediators are less distinct than younger adults' mediators such that the mediators do not make it possible to individuate intact pairs. For example, the very specific mental image of a human figure wrapped in a blanket would correctly lead one to declare the recombined pair *blanket-summer* new, but the more general (less distinct) idea of a human merely possessing a blanket may not be sufficient to declare *blanket-summer* new because some people sit on blankets outside during the summer and some wrap themselves in beach towels. If older adults' mediators are more general or less distinct than younger adults' mediators, this may manifest as a decoding deficiency that leads to false alarms. This possibility is compatible with Hertzog et al.'s (2013) finding that older adults' retrieved mediators are often gist representations of the mediators verbalized at study.

We suggest that a binding deficit, clearly defined, cannot by itself account for the high false alarm rate of older adults because failure to bind items should lead to *misses* rather than *false alarms* in the absence of difference in process. We have focused on two mechanisms—the failure to retrieve mediators and failure to decode mediators—as processes by which impairments in older adults' memory could reflect higher rates of false alarms.

Limitations and Future Directions

One limitation of our study is a possible floor effect in the associative test FAR data for younger adults. Indeed, 40% of younger adults had FAR scores at floor (.01) compared to 4.4% of older adults. Increasing the difficulty of the task by shortening the presentation time (e.g., Brubaker & Naveh-Benjamin, 2014) might partly mitigate floor effects in younger adults but would also likely increase FAR in older adults. Indeed, Brubaker and Naveh-Benjamin have shown that very short presentation rates (1.5s) increase FAR in younger adults relative to longer presentation rates (6s). Our participants had 5s to study pairs. It is possible that this presentation rate was more conducive to encoding for younger than older adults, and contributed to their low FAR for pairs. The 5s rate (and around it, e.g., 4.5s, 5.5s, 6s) is frequently employed in the associative deficit paradigm (e.g., Bender et al., 2010; Bender & Raz, 2012; Berry et al., 2013; Kilb & Naveh-Benjamin, 2007; Naveh-Benjamin, 2000, Experiments 2 and 3; Naveh-Benjamin & Kilb, 2012; Overman & Becker, 2009; Overman & Stephens, 2013).

The question of the effect of presentation rate on FAR is relatively new, so new, in fact, that it does not appear as a potential moderator in the Old and Naveh-Benjamin (2008) meta-analysis of the associative deficit. To our knowledge, only one study has conducted a systematic analysis of its effect in the associative deficit recognition paradigm, and only in younger adults (Brubaker & Naveh-Benjamin, 2014). Decreasing presentation rate may reduce or eradicate FAR floor effects in younger adults but create ceiling effects in older adults. Moreover, shorter presentation rates would impoverish verbalization content and undermine our focus on strategic behavior as represented by think-aloud data. It would also limit direct comparisons of our results to those based on comparable presentation rates. Yet, with shorter presentation rates for *both* age groups, we might be able to identify those individuals – both young and old – who are best at producing the highest quality mediators under the most demanding encoding conditions, thereby arriving at one of the boundary conditions for the associative deficit hypothesis.

A different approach to addressing the problem of floor effects in FAR in younger adults is to manipulate type of lure pairs presented at test. In our study, and most other research on the associative deficit hypothesis, lure pairs at test are recombinations of item stimuli comprising study pairs (but see Castel & Craik, 2003, who used old-new and new-new pairs as well).

Systematically varying the degree of semantic relatedness of lure pairs might also induce increases in FAR to pairs at test in younger adults when lures are semantically (or otherwise meaningfully) related to *either* of the individual items comprising study pairs. For example, if the study pair is *blanket-figure*, high semantic lure pairs at test could take the form of *towel-figure* and *blanket-human*. While such a manipulation might increase FAR in younger adults (and move them off floor), it could also produce FAR ceiling effects in older adults.

We used a yes-no recognition test format. Participants saw intact or recombined pairs and judged whether or not they recognized the pairs. Use of a 4-alternative-forced-choice recognition test format (Patterson & Hertzog, 2010) also implicates false alarms--not misses of bound pairs--as the core of the associative deficit in older adults. Future research ought to compare these two test formats directly.

Conclusion

A think-aloud analysis of encoding strategies used by younger and older adults for paired associates learning shows that strategies vary substantially both within and between age groups. Yet many older adults fail to benefit from encoding strategies that are conducive to high performance on recognition tests in younger adults. Although generating deep mediators during study confers high HRs in both age groups—younger and older adults are both more likely to recognize intact pairs after having encoded them with deep mediators—younger adults who generate deep mediators at study are more likely than their older adult counterparts to correctly declare recombined pairs new. These results suggest that many older adults have difficulty retrieving their mediators when presented with recombined pairs at test, older adults’ mediators are not distinct enough to individuate intact pairs from recombined pairs at test, or some combination of both. Further elucidation of the relationship between encoding strategies and recognition judgments may be attainable through additional process-oriented studies.

References

Ariel, R., Price, J., & Hertzog, C. (2015). Age-related associative memory deficits in value-based remembering: The contribution of agenda-based regulation and strategy use. *Psychology and Aging, 30,* 795-808. doi: 10.1037/a0039818

Badham, S. P., Estes, Z., & Maylor, E. A. (2012). Integrative and semantic relations equally alleviate age-related associative memory. *Psychology and Aging, 27,* 141-152. doi: 10.1037/a0023924

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B.,…Treiman, R. (2007). The English lexicon project. *Behavior Research Methods, 39,* 445-459. doi:10.3758/BF03193014

Bender, A. R., Naveh-Benjamin, M., & Raz, N. (2010). Associative deficit in recognition memory in a lifespan sample of healthy adults. *Psychology and Aging, 25,* 940-948. doi: 10.1037/a0020595

Bender, A. R., & Raz, N. (2012). Age-related differences in recognition memory for items and associations: contribution of individual differences in working memory and metamemory. *Psychology and Aging, 27,* 691-700. doi: 10.1037/a0026714

Bender, A. R., Naveh-Benjamin, M., & Raz, N. (2010). Associative deficit in recognition memory in a lifespan sample of healthy adults. *Psychology and Aging, 25,* 940-948. doi: 10.1037/a0020595

Berry, J. M., Williams, H. L., Usubalieva, A, & Kilb, A. (2013). Metacognitive awareness of the associative deficit for words and names. *Aging, Neuropsychology, and Cognition: A Journal on Normal and Dysfunctional Development, 20,* 592-619. doi: 10.1080/13825585.2012.761670

Brubaker, M. S., & Naveh-Benjamin, M. (2014). The effects of presentation rate and retention interval on memory for items and associations in younger adults: A simulation of older adults' associative memory deficit. *Aging, Neuropsychology, and Cognition: A Journal on Normal and Dysfunctional Development, 21,* 1-26. doi: 10.1080/13825585.2013.772558

Craik, F. I. M. (2002). Levels of processing: Past, present . . . and future? *Memory, 10,* 305-318. doi: 10.1080/09658210244000135

Castel, A. D., & Craik, F. I. M. (2003). The effects of aging and divided attention on memory for item and associative information. *Psychology and Aging, 18*, 873-885. doi: 10.1037/0882-7974.18.4.873

Dunlosky, J., & Hertzog, C. (1998). Aging and deficits in associative memory: What is the role of strategy production? *Psychology and Aging, 13,* 597-607. doi: 10.1037/0882-7974.13.4.597

Dunlosky, J., & Hertzog, C. (2001). Measuring strategy production during associative learning: The relative utility of concurrent versus retrospective reports. *Memory & Cognition, 29,* 247-253. doi:10.3758/BF03194918

Dunlosky, J., Hertzog, C., & Powell-Moman, A. (2005). The contribution of mediator-based deficiencies to age differences in associative learning. *Developmental Psychology, 41,* 389-400. doi: 10.1037/0012-1649.41.2.389

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.

Ericsson, K. A., & Fox, M. C. (2011). Thinking aloud is not a form of introspection but a qualitatively different methodology: Reply to Schooler (2011). *Psychological Bulletin, 137*, 351-354. doi: 10.1037//a0022388

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review, 87*, 215-251. doi: 10.1037/0033-295X.87.3.215

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data.* (Rev. ed.). Cambridge, MA: MIT Press.

Fox, M. C., & Charness, N. (2010). How to gain eleven IQ points in ten minutes: Thinking aloud improves Raven's Matrices performance in older adults. *Aging, Neuropsychology, and Cognition, 17*(2), 191-204. doi.org/10.1080/13825580903042668

Fox, M. C., Ericsson, K. A., & Best, R. (2011). Do procedures for verbal reporting of thinking have to be reactive? A meta-analysis and recommendations for best reporting methods. *Psychological Bulletin, 137,* 316-344. doi: 10.1037//a0021663

Gallo, D. A., Meadow, N. G., Johnson, E. L., & Foster, K. T. (2008). Deep levels of processing elicit a distinctiveness heuristic: Evidence from the criterial recollection task. *Journal of Memory and Language, 58*, 1095-1011. doi: 10.1016/j.jml.2007.12.001

Grice, J. W. (2011). *Observation oriented modeling: Analysis of cause in the behavioral sciences.* Burlington, MA: Academic Press.

Hertzog, C., Fulton, E. K., Mandviwala, L., & Dunlosky, J. (2013). Older adults show deficits in retrieving and decoding associative mediators generated at study. *Developmental Psychology, 49*, 1127-1131. doi: 10.1037/a0029414

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory & Language, 30*, 513-541. doi:10.1016/0749-596X(91)90025-F

Kilb, A., & Naveh-Benjamin, M. (2007). Paying attention to binding: Further studies assessing the role of reduced attentional resources in the associative deficit of older adults. *Memory & Cognition, 35*, 1162-1174 doi: 10.3758/BF03193486

Kuhlmann, B. G., & Touron, D. R. (2012). Mediator-based encoding strategies in source monitoring in young and older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*, 1352-1364. http://dx.doi.org/10.1037/a0027863

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, 28*, 203-208. doi:10.3758/BF03204766

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

Naveh-Benjamin, M., Brav, T. K., & Levy, O. (2007). The associative memory deficit of older adults: The role of Strategy utilization. *Psychology and Aging, 22*, 202-208. doi: 10.1037/0882-7974.22.1.202

Naveh-Benjamin, M., & Kilb, A. (2012). How the measurement of memory processes can affect memory performance: The case of remember/know judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*, 194-203. doi: 10.1037/a0025256

Naveh-Benjamin, M. Shing, Y. L., Kilb, A., Werkle-Bergner, M. Lindenberger, U., & Li, S.-C. (2009). Adult age differences in memory for name-face associations: The effects of intentional and incidental learning. *Memory, 17,* 220-232. doi: 10.1080/09658210802222183

Old, S. R., & Naveh-Benjamin, M. (2008). Differential effects of age on item and associative measures of memory: A meta-analysis. *Psychology and Aging, 23,* 104-118. doi: 10.1037/0882-7974.23.1.104

Overman, A. A., & Becker, J. T. (2009). The associative deficit in older adult memory: Recognition of pairs is not improved by repetition. *Psychology and Aging, 24,* 501-506. doi: 10.1037/a0015086

Overman, A. A., Stephens, J. D. W. (2013). Synergistic effects of encoding strategy and context salience on associative memory in older adults. *Psychology and Aging, 28,* 654-665. doi: 10.1037/a0031441

Patterson, M. M., & Hertzog, C. (2010). The effects of age in four alternative forced-choice item and associative recognition tasks. *Psychology and Aging, 25*, 235-238. doi: 10.1037/a0016046

Salthouse, T. A. (1991). Mediation of adult age differences in cognition by reductions in working memory and speed of processing. *Psychological Science, 2,* 179-183. doi: 10.1111/j.1467-9280.1991.tb00127.x

Wechsler, D. (1981). *The Wechsler Adult Intelligence Scale—Revised*. San Antonio, TX: Psychological Corporation.

Table 1

Means, Standard Deviations, and Effect Sizes for Demographic Comparisons between Age Groups and Verbalization Groups.

Variable	Younger adults Silent (*N = 22)*	Younger adults Think-aloud (*N = 30)*	Older adults Silent (*N = 19)*	Older adults Think-aloud (*N = 26)*	Age effect size *(d)*	Verbalization effect size *(d)*
Years of education	13.00 (1.20)	13.02 (1.06)	16.10 (2.13)	16.12 (2.85)	1.58**	0.00
Self-rated health	8.55 (1.26)	8.47 (1.25)	7.89 (1.76)	7.73 (1.93)	-0.45*	-0.07
Self-rated vision	8.23 (2.14)	8.70 (1.51)	7.79 (1.75)	7.77 (1.82)	-0.40*	0.14
Self-rated hearing	8.50 (1.54)	8.70 (1.39)	7.00 (1.97)	7.82 (1.72)	-0.65*	0.31
Speed of processing[a]	71.14 (11.85)	69.50 (8.51)	44.16 (13.06)	48.60 (10.89)	-2.14**	0.06
Vocabulary[b]	24.61 (2.93)	24.45 (2.38)	27.89 (2.21)	29.80 (3.34)	1.58**	0.21

Note. Scales for self-rated health, vision, and hearing ranged from 0 (poor) to 10 (excellent). [a]Digit Symbol Substitution Test (DSST; Wechsler, 1981). [b]Ekstrom, French, Harman, and Dermen (1976) Synonyms Test.

*p < .05; **p < .01.

Table 2

Examples of Verbalizations Classified as Repetition, Shallow Mediator, and Deep Mediator.

Pair	Repetition	Shallow mediator	Deep mediator
Radar-picnic	"Radar-picnic, radar-picnic, radar-picnic."	"Radar-picnic, radar-picnic, darp [sic]."	"Radar-picnic. It ain't no picnic when you're in a submarine with a radar."
Essay-husband	"Essay-husband, essay-husband, essay-husband, essay-husband."	"Essay-husband, E-H."	"Essay-husband. You might want your husband to write an essay on why he's always wrong."
Blanket-figure	"Blanket-figure, banner-textile [the next pair in the list], blanket-figure."	"Blanket-figure, B-F."	"Okay, the blanket-figure, put a blanket over a dead figure, over a dead body."

Table 3

Frequencies of Strategies by Age Group

	Younger adults		Older adults	
Strategy	Frequency	Mean rank	Frequency	Mean rank
No strategy	848	22.38	1,029	33.02
Repetitions	468	31.72	121	20.85
Shallow mediators	51	27.22	22	26.72
Deep mediators	332	28.67	148	24.83

Note. No strategy = collapse of remaining two categories that are not theoretically relevant (reading without repetition, and nothing/missing/unclear.). Higher mean rank reflects higher frequency.

Table 4

Correlations between Performance Outcomes and Use of Encoding Strategies for Younger and Older Adults

	Repetitions	Shallow mediators	Deep mediators	Items HR	Items FAR	Associates HR	Associates FAR
Repetitions	--	.26	-.05	-.40*	-.05	-.25	.27
Shallow mediators	.33	--	-.03	.04	-.15	-.09	.20
Deep mediators	.36	.66**	--	.50**	-.23	.53**	-.51**
Items HR	.15	.23	.37*	--	-.25	.46**	-.33*
Items FAR	.21	-.07	.18	.08	--	-.38**	.59**
Associates HR	.31	.44*	.75**	.43**	.07	--	-.65**
Associates FAR	.14	-.08	-.13	-.09	.44**	-.05	--

Note. HR = hit rate; FAR = false alarm rate. Younger adult data above diagonal, older adult data below diagonal.

*$p < .05$; **$p < .01$.

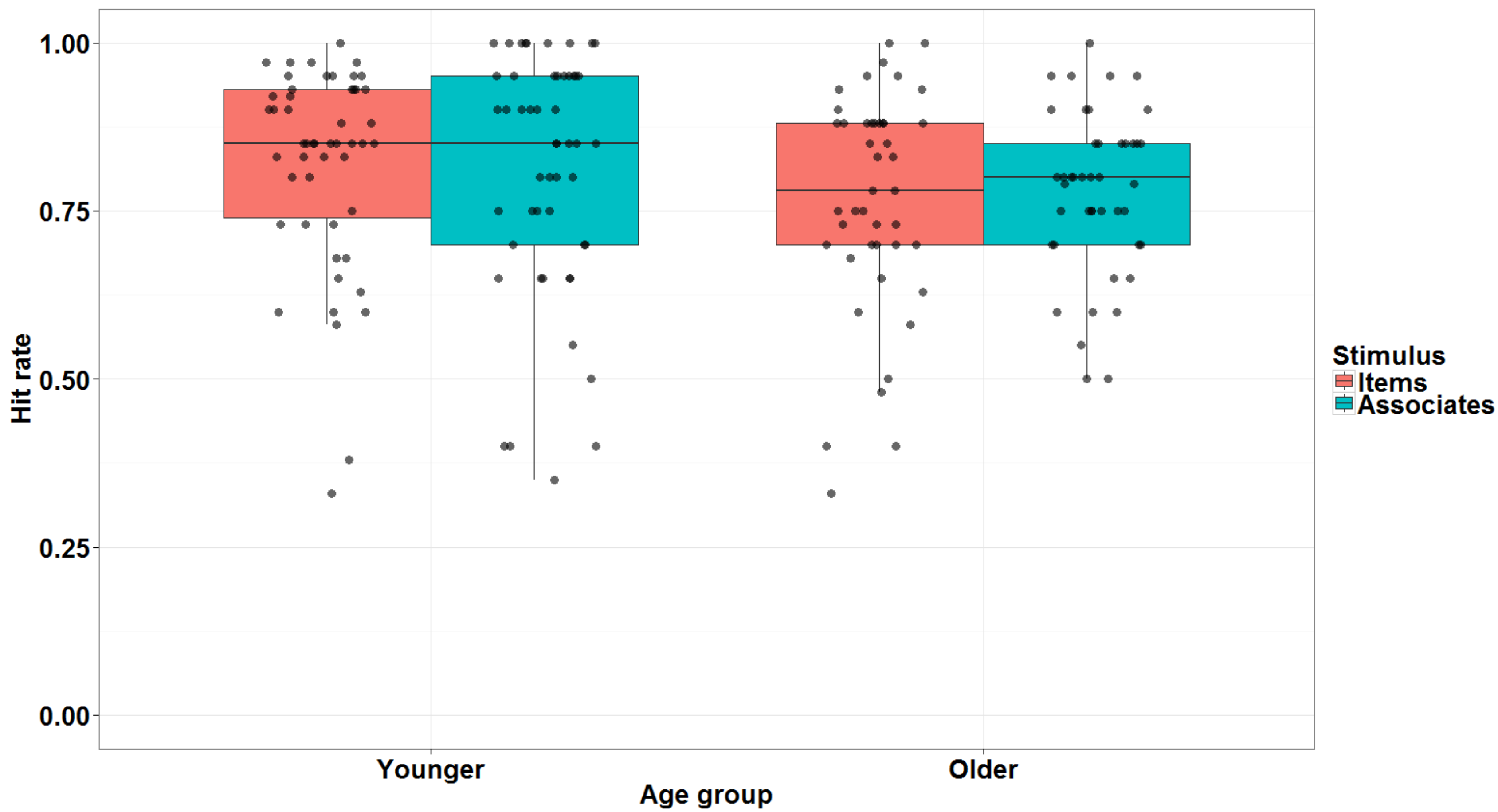

Figure 1. Hit rates with boxplots and scatterplots. Horizontal lines represent medians.

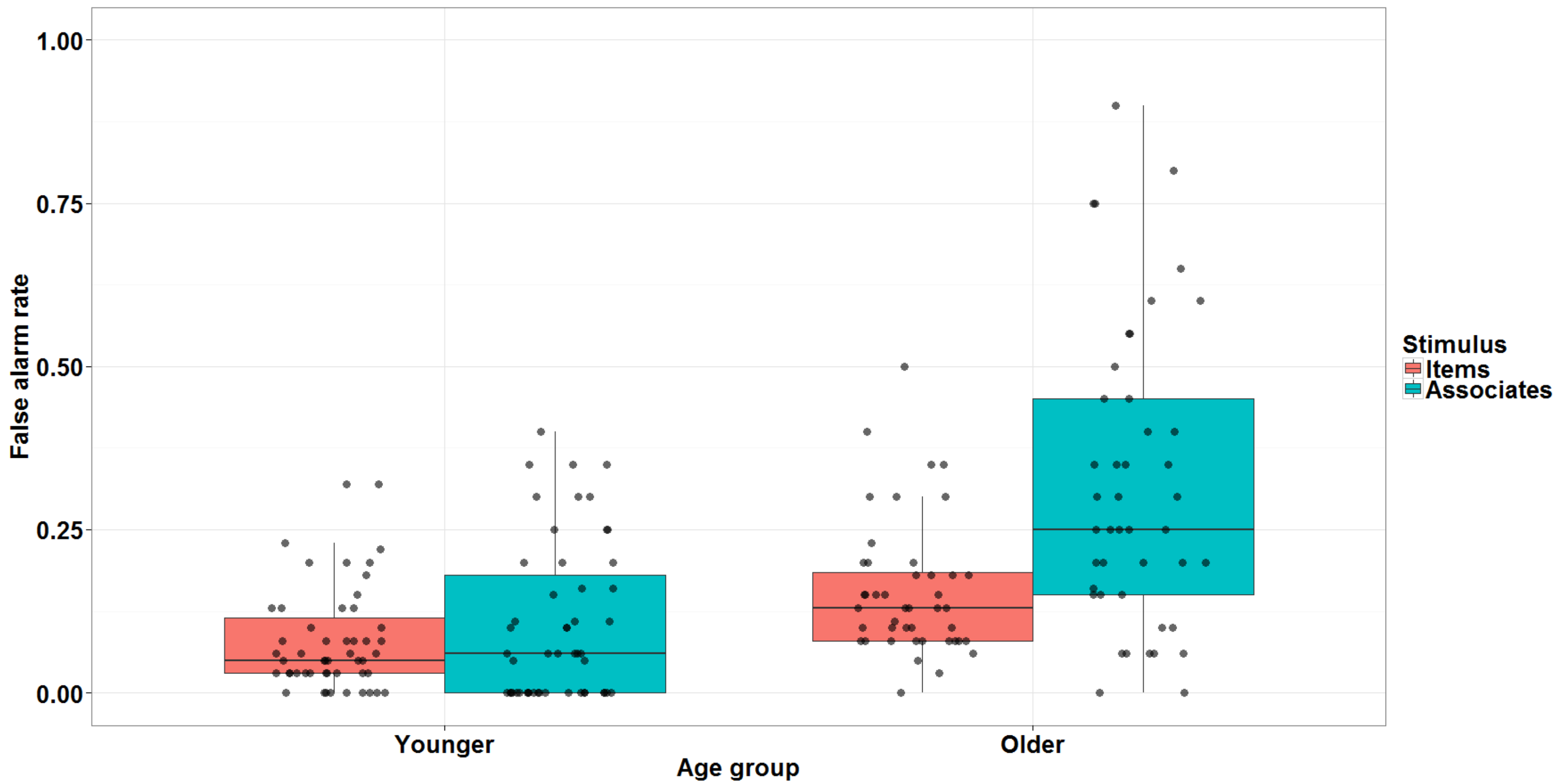

Figure 2. False alarm rates with boxplots and scatterplots. Horizontal lines represent medians.