2017

The Familiarity Factor: How Semantic Similarity Affects Associative Memory in Older and Younger Adults

Paige Warner
University of Richmond

The Familiarity Factor:

How Semantic Similarity Affects Associative Memory in Older and Younger Adults

by

Paige Warner

Honors Thesis

Submitted to

Department of Psychology
University of Richmond
Richmond, VA

April 28, 2017

Advisor: Dr. Jane Berry

Abstract

The role of semantic knowledge and familiarity on retrieval processes were investigated in a test of the associative deficit hypothesis (ADH), and the data were interpreted from the perspectives of fuzzy trace theory and source monitoring errors (Naveh-Benjamin, 2000). Younger and older adults (N=60) studied pairs of words for two recognition tests – an item test, for recognition of individual words, and an associative test, for recognition of word pairs. In the associative test, four word pairs were presented with a forced-choice response requirement (4AFC; Patterson & Hertzog, 2010). In addition to the studied, intact cue-target pair, three cue-target pairs were presented as lures. A similar framework was employed in the item test, but used individual words instead of pairs. The associative deficit hypothesis was not supported: Older and younger adults had comparable recognition accuracy for items and associates. However, the types of errors made in the associative test revealed significant age differences. Older adults had significantly higher false alarm rates (FAR) to recombined pairs compared to younger adults, and younger and older adults were more likely to make recombined false alarms than other types of false alarm. Additionally, older adults were more susceptible to making intact related false alarms than recombined related false alarm rate. Responses in the item test did not vary by false alarm type nor by age group. These results demonstrate the importance of familiarity in age-based associative memory errors and help uncover a possible source of the false alarm effect.

Keywords: older adults, associative deficit, memory, false alarms

The Familiarity Factor:

How Semantic Similarity Affects Associative Memory in Older and Younger Adults

Some aspects of memory decline in adulthood, while others remain intact. Memory performance often depends on the types of stimuli, as well as the testing conditions. For example, it is well documented that older adults perform more poorly than younger adults on paired-associate tasks. In order to try to explain this specific decline in episodic memory, Naveh-Benjamin (2000) hypothesized that the associations between units weaken with age, which leads to difficulty in recalling individual units as a linked pair. The associative deficit hypothesis (ADH) has been tested against various moderators and mediators, demonstrating its robust nature. In fact, when manipulating variables such as source, context, order, and location of stimuli, the associative deficit remains for older adults (Old & Naveh-Benjamin, 2008).

A unique aspect of the associative deficit effect is the tendency for older adults to say they recognize a pair of words in the associative test when, in fact, the pair was not presented previously (Berry, Williams, Usubalieva, & Kilb, 2013; Fox, Baldock, Freeman, & Berry, 2016; Bender, Naveh-Benjamin, & Raz, 2010; Light, Patterson, Chung, & Healy, 2004). This age-related deficit to commit false alarms for associative pairs is not observed on the item test, indicating that there is something special about the act of linking two items together and remembering them. Surprisingly, little research has been done to investigate the cause of the false alarm effect. Do older adults have a real memory experience for these pairs, and truly believe they have seen them before, or are they responding ‘yes’ due to a lack of specific memory for the correct pairing? In order to try to answer these questions regarding the cause of the false alarm effect, we examined two sources of memory error: gist errors from the perspective of fuzzy trace theory and source monitoring errors (Brainerd & Reyna, 2002).

Fuzzy trace theory states that when forming a memory, people store a verbatim trace (the surface form of the word) and a gist trace (episodic interpretation) in parallel (Brainerd & Reyna, 2002). While both types of memory are used throughout the lifetime, as people make more connections and associations, they tend to rely more heavily on gist traces. Brainerd (2013) reported that as age increases, so do the occurrences of both false recognition and false recall because adults focus more on the idea rather than the thing itself. Remembering the gist of something is usually adequate in everyday situations and helps save energy and mental resources. However, when being asked to remember specific words and pairs, the gist alone may not be as accurate as a verbatim trace. And if older adults rely more on gist memory than younger adults, then words that are closely related in meaning should create even more occurrences of false recognition or false alarms.

Source monitoring refers to the process of identifying the origin – or source – of a memory trace. A source monitoring error occurs when the origin of a memory is attributed to something other than the original source (Johnson, Hashtroudi, & Lindsay, 1993). Older adults are more likely than younger adults to make these sorts of errors and often experience difficulty identifying the origin of a piece of information (Dehon & Brédart, 2004). For example, a younger adult may remember that she heard the news of Trump's election on CNN, but an older adult may not remember whether he heard it on CNN or NBC. In the associative deficit paradigm, a source monitoring deficit would appear as an older adult's inability to monitor source information accurately, for example, not knowing whether the word "door" was paired with "cat" or with "spoon" in the original presentation of pairs.

The associative deficit paradigm tests participants using intact pairs (words presented together during the study phase) and recombined pairs (words presented during the study phase

but not together). Recombined pairs presented at test may lead to source monitoring errors. Because all the words are studied during the study phase but presented in some intact and some unique pairings at test, it is possible that older adults may be experiencing difficulty determining the source of the words. Here, the source is the intact pair, against which a recombined pair should be compared to determine source accuracy. Indeed, the false alarm effect may be a reflection of older adults struggling to recall exactly which words were presented together and becoming overwhelmed with feelings of familiarity.

Previous research has linked the associative deficit with source monitoring errors, especially in the context of binding information. When trying to recall sequences of actors and actions, older adults experience more difficulty distinguishing the old events with novel, similar events (Kersten, Earles, Curtayne, & Lane, 2008). These age-related source monitoring differences are not limited to events and are also seen in memory for words. For example, source memory appears to play a larger role in remember-know judgments for younger adults than for older adults, who do not typically use source memories to distinguish these judgments (Boywitt, Kuhlmann, & Meiser, 2012). In a remember-know task, a "remember" judgment corresponds to a detailed and conscious recollection by the participant, while a "know" judgment is made when there is familiarity present but not a definitive memory.

Further evidence for a source-monitoring deficit comes from neuroimaging studies. Younger adults typically make fewer source monitoring errors in word recognition tasks, with ERP patterns demonstrating a larger difference in activation between correct and incorrect trials (Dywan, Segatlowitz, & Webster, 1998). In contrast, older adults showed greater ERP reactivity to incorrect trials, which may indicate a loss in cognitive resources and the ability to inhibit incorrect responses. A recent FMRI study also demonstrated that when encoding pairs, but not

items, the pattern of brain activity was not as predictable or specific for older adults as for younger adults (Saverino et al., 2016). This delocalization of brain activity may help explain why older adults tend to exert more effort with less successful results. Additionally, during delocalized encoding, other memories or concepts may be stimulated, leading to false recognition and false alarms.

In a commonly used methodology to study false memories, the DRM paradigm presents participants with a list of words that are all related to a critical non-presented lure word, for example, bed, rest, and awake as associates for the lure sleep (Roediger & McDermott, 1995). For a recognition task, the false alarm rate for the critical lure is almost the same as the hit rate for the words actually present on the study list, illuminating the inaccuracy and malleability of memory. Specifically relating to age, Dehon and Brédart (2004) reported that older adults were more likely to say they had actually heard the critical lure. While this may be due to source monitoring errors, another possible explanation lies in fuzzy trace theory. Older adults may recall the gist of the category (e.g., items related to sleep) but not the items verbatim (e.g., sleep was not presented in the list of studied items).

In the typical associative deficit paradigm, participants study pairs of words and are instructed to try to remember both the items (such as apple, chair, frog, scarf) and the word pairs (apple-chair, frog-scarf) in order to prepare for two recognition memory tests. For one test, the item test, participants try to recognize the words from the study list, while rejecting new, unstudied words. For the other test, the associative test, participants attempt to identify intact pairs (words presented together during the study phase) and reject recombined pairs (words presented during the study phase but not paired together). The current study deviated from this typical paradigm by using a four alternative forced choice (4AFC) recognition test format instead

of the more commonly used yes-no recognition test format. This format may provide more insight into how test format affects the associate deficit observed in older adults (Patterson & Hertzog, 2010).

In 4AFC tests, participants are required to choose the correct answer from a set of four alternate choices or to make their best guess if the correct answer does not come to mind. Patterson and Hertzog (2010) used this paradigm to test the associative deficit hypothesis. They presented an intact pair as well as three rearranged foils comprised of words from the study list but studied with different pairs. They found that the 4AFC format eradicated the associative deficit in older adults, implying that the associative deficit might be specific to test-format (e.g., yes-no recognition of individually presented pairs) and not generalizable to all memory test formats (e.g., free recall). Their results pose an interpretive problem for possible mechanisms underlying the associative deficit effect.

In order to probe this problem further, the current study utilized the 4AFC format and manipulated the semantic strength of lure pairs in the recognition test. By making the lures semantically related to each other and to other pairs on the list, including the correct answer, the memory test was made more difficult than that used by Patterson and Hertzog (2010). Thus, we introduced both semantic familiarity and list familiarity into the recognition test. By doing so, we predicted that the associative deficit that was reversed by the 4AFC used in the Patterson and Hertzog study would be reintroduced in our study.

Both gist errors and source monitoring errors rely on some form of familiarity, but which type of error will older adults commit more often? We hypothesized that the associative deficit will become even more pronounced in older adults when introducing items with similar meanings into the pairs because as people age, they tend to rely more on meaning and have

developed a larger semantic network (Brainerd, Reyna, & Ceci, 2008). If older adults were to choose the intact-related lure pair during the associative test, this would provide evidence for the distal familiarity explanation of fuzzy trace theory. However, if older adults tend to choose the recombined lure pair, they are making source-monitoring errors, which suggests that they are susceptible to the proximal familiarity explanation. The final type of error, the recombined-related lure pair, is the most dissimilar from the others, possibly indicating a loss of the original memory trace for intact pairs. For the item test, we did not expect to find differences between item error types because older adults often perform as well as younger adults on the item test and do not commit more item false alarms than younger adults (Berry et al., 2013).

Method

Participants

Twenty-eight younger adults (20 female) were recruited through the University of Richmond Psychology Department, and they were offered course credit for their participation. Thirty-two older adults (19 female) were recruited from the Richmond area through a newspaper advertisement and received $15 for participation. All of the participants were in good health (M = 8.23, SD = 1.29) and reported no history of Alzheimer's or other memory disorders. The majority of the participants were Caucasian (49 participants), followed by Asian/Asian American (5), African American (2), Latino (2), and other (2). Older adults had higher vocabulary scores (M = 28.20, SD = 3.86) than did younger adults (M = 24.84, SD = 3.26), $t(58)$=-3.615, p = 0.001, as measured by the Ekstrom, French, Harman, and Dermen Synonyms Test (1976), a vocabulary test. Younger adults exhibited higher processing speeds (M = 69.64, SD = 10.56) than did older adults (M = 43.94, SD = 11.78), $t(58)$=8.85, p = .001, as revealed by the WAIS-R Digit-Symbol Substitution Task (DSST, Weschler, 1981).

Materials

The stimuli consisted of semantically unrelated word pairs, as well as semantically related counterparts to the words in each pair. These unrelated word pairs and semantic counterparts were used previously in a study of the associative deficit and semantic relatedness (Patterson, Light, Ocker, & Olfman, 2009). The stimuli were originally chosen from the University of South Florida word association, rhyme, and word fragment norms (Nelson, McEvoy, & Schreiber, 1998). In their database, the mean relatedness of the original word to its semantic counterpart was 0.26, while the mean relatedness of the semantic counterpart to the original word was 0.20. The mean frequency of the words was 48.38, and only words with fewer than 3 syllables were chosen as stimuli.

Both semantically related and unrelated lures were presented in the item and associative tests. More specifically, in the associative test, three recombined cue-target pairs were presented as lures in addition to the intact cue-target pair. The first lure type combines the cue from the intact pair with a target from another studied pair (recombined pair). The second lure type combines the cue with a target that is semantically related to the target in the intact pair (intact related pair). The third lure type combines the cue with a target that is semantically related to the target in the recombined pair (recombined related pair). For example, if participants saw the pairs "error-glove" and "mix-cellar" during the study phase, they would see these four options in the associative test: "error-glove" (intact pair), "error-cellar" (recombined pair), "error-mitten" (intact related pair), and "error-basement" (recombined related pair).

The item test followed a similar pattern, with the target word in addition to three lures. The first lure type was a completely new, unstudied word (novel word). The second lure type was a word that is semantically related to the target (target related word), and the third was a

semantically similar word to the novel lure (novel related word). For example, if participants studied "error-glove," they would see these choices on the item test: "error" (target word), "herb" (novel word), "mistake" (target related), and "spice" (novel related). This manipulation of semantically related lure types allowed us to examine the role of semantic knowledge (Patterson et al., 2009) on false alarm rates to items and pairs at the time of retrieval.

The memory task consisted of three blocks, each with 30 randomized, unrelated pairs on the study list. From these 180 words, 45 words were used in the item test and 135 words were used in the associative test. The item test in each block contained 15 trials (45 total trials), with 15 randomly selected words from the study phase along with the three lures described previously. The associative test also had 15 trials per block (45 total trials), with the intact cue-target pair and three other rearranged foils that had the same cue but different targets, as described above. The order of the four options in each test type was randomized, as well as the order of the item and associative tests. All portions of the memory task were programmed and run using ePrime 2.0.

Study Design

This study employed a mixed design, with age (young or old) as the primary between-subjects factor and test type (item or associative test) as the within-subjects factor. However, because we were interested in the types of errors that older and younger adults make during an associative test, we also used error type (recombined, intact related, recombined related) as a within-subjects factor. A similar structure was used for the item test, with differently labeled error types (novel, target related, and novel related). These mixed analyses of error type may provide insight into the underlying memory mechanisms affecting associative memory in younger and older adults.

Procedure

Participants were asked to come into the lab for an experimental session that lasted approximately one hour. After completing a background information questionnaire, participants were informed about the procedure for the memory task and completed a practice block in order to familiarize them with the procedure. The practice block contained six pairs in the study list, three item test trials, and three associative test trials. After practice, participants studied 30 pairs of words one at a time for either three seconds (for younger adults) or six seconds (for older adults). The presentation times by age group follows the procedure of Patterson and Hertzog (2010) in the 4AFC design. After all 30 pairs were presented, participants completed a computerized pattern comparison distractor task to prevent rehearsal (Salthouse, 1991). They were asked to indicate whether two abstract drawings were the same or different and were given 20 seconds to complete as many as possible.

Since the order of the item and associative tests was counterbalanced, half of the participants took the item test first, while the other half took the associative test first after the study phase. For both types of tests, four options appeared on each quadrant of the computer screen. Participants used four keys on the keyboard corresponding to the different quadrants to indicate which word or pair they recognized from the study phase. If participants were uncertain about the correct answer, the experimenter would instruct them to make their best guess for trials of which they are unsure. Each block consisted of a study phase with 30 word pairs, an item test with 15 trials, and an associative test with 15 trials. When all three blocks were completed, participants completed a post-test questionnaire asking about their memorization strategies, performance outcomes confidence, and perceived task difficulty for the item and pair tasks.

Lastly, a vocabulary test and a digit symbol substitution task were administered. Participants were debriefed and given either $15 or course credit upon completion.

Results

Recognition memory performance

A 2 (age group: young, old) x 2 (test type: item, associative) mixed analysis of variance (MANOVA) was conducted to test for an age-related associative deficit. The main effect for age was not significant, $F(1,58) = 0.706$, $p > 0.05$, indicating that older and younger adults obtained similar recognition accuracy, calculated as the proportion of correct responses, collapsed across the two types of tests. The main effect for test type approached significance, $F(1,58) = 3.732$, $p = 0.058$, with higher accuracy, surprisingly, on the associative test relative to the item test. These results are uncommon in the associative deficit literature, where older adults typically perform more poorly than younger adults overall and, in particular, on the paired associate tests. The interaction between age group and test type was also not significant, $F(1,58) = 2.601$, $p > 0.05$, see Figure 1. These findings replicate those of Patterson and Hertzog (2010), who reported the absence of an associative deficit on their four alternative forced-choice (4AFC) test format.

Response Types for the Associative Test

Although older and young adults were comparable on overall recognition accuracy, we also analyzed types of errors made by older and younger adults. When participants did not choose the correct pair, we labeled this error as a "false alarm" and calculated the false alarm rate (FAR), the proportion for each error type. We conducted a 2 (age group: young or old) x 3 (error type: recombined, intact related, recombined related) MANOVA. The main effect for error type was significant, $F(2, 116) = 44.630$, $p < 0.01$. When an error was made in the associative test, it was most commonly as a response to a recombined pair, followed by a response to an intact

related pair, and last, to a recombined related pair. The interaction between error type and age group was also significant, $F(2,116) = 6.743$, $p < 0.01$, see Figure 2.

In order to characterize this interaction, we conducted paired samples *t*-tests within each age group. For younger adults, the only differences occurred between false alarm rates for recombined and intact related pairs, $t(27)=-4.239$, $p < 0.01$, and recombined and recombined related pairs, $t(27)=4.446$, $p < 0.01$. The FAR for recombined pairs was higher than both intact related and recombined related pairs, but the difference between intact related and recombined related pairs was not significant, $t(27)=0.130$, $p > 0.05$. For older adults, significant differences were observed between all three types of errors. A paired *t*-test revealed a higher FAR to recombined than intact related pairs, $t(31)=-4.683$, $p < 0.01$, and a higher FAR to recombined than recombined related pairs, $t(31)=7.786$, $p < 0.01$. Unlike younger adults, older adults also experienced significantly more false alarms to intact related than recombined related $t(31)=3.641$, $p < 0.01$. To summarize, when older adults chose the wrong answers (i.e., committed false alarms), they were most likely to choose a recombined pair, followed by an intact related pair, and then, a recombined related pair. This pattern provides some evidence that older adults may be more sensitive to different types of response pairings than younger adults and implies that older adults may be experiencing more familiarity-based memory processes.

These results showed that both older and younger adults have the highest FAR for recombined pairs compared to the other two types of errors. However, an independent samples *t*-test revealed that older adults experience more false alarms than younger adults specifically for recombined pairs, $t(58)=-2.204$, $p = 0.03$. Independent samples *t*-tests for the other error types (intact related and recombined related) between the age groups were not significant.

Response Types for the Item Test

We examined accuracy and error types for the item test. Older and younger adults had equivalent recognition accuracy scores for the item test. We then analyzed age differences in errors types committed on the item test, using a 2 (age group: young or old) x 3 (error type: novel, target related, or recombined related) MANOVA. The main effect for error type approached significance, $F(2,116)=2.398$, $p < 0.10$, with the target related lure having the largest false alarm rate, followed by the novel related lure, and then the novel lure. The main effect for age group was not significant, $F(1,58)=0.053$, $p > 0.05$, and the interaction between age group and error type was not significant, $F(2,116)=0.241$, $p > 0.05$, see Figure 3. These results support our hypothesis that younger and older adults would have comparable recognition accuracy scores for items. The results for age-related differences on error types for items (i.e., no age differences on error types) are different than those observed for pairs, which yielded significant age differences.

Discussion

The results of this study confirmed our hypothesis that older adults may be relying on familiarity-based processes when attempting to recognize previously viewed stimuli. However, these processes may not be specifically due to semantic relatedness for paired-associate tasks. We did not observe an associative deficit on our 4AFC task with semantically-related lure types. Yet, older adults did experience more false alarms with intact related pairs than the recombined related pairs, a pattern not seen in younger adults. These differences were only observed on the associative test, further supporting the idea that there is something special about memory for paired stimuli versus single items. Although we did not observe an overall difference in accuracy, the differences in types of errors may be indicative of deeper cognitive processes.

Our findings provide more support for a source monitoring explanation of error patterns on tests of the associative deficit because both younger and older adults struggled with the recombined pairs compared to the other lures (Johnson et al., 1993). However, there was a significant discrepancy between the FAR for intact related and recombined related pairs for only older adults. Therefore, older adults may have been more sensitive to semantic familiarity, which provides some support for fuzzy trace theory (Brainerd & Reyna, 2002). Further research could elucidate whether the manipulation of semantic similarity becomes more salient after eliminating the recombined option, which may be currently masking a semantic familiarity effect.

Our results are consistent with Patterson and Hertzog (2010). The 4AFC test format reversed the associative deficit, which highlights the importance of the test format to the ADH. This may be due to more support at retrieval since the participants know that one of the answers is correct. Future research should employ additional test format manipulations to analyze the broader impact of test format on the associative deficit hypothesis. To fully capture a psychological phenomenon, it is important that the experimental results arise from the phenomenon itself and not from the way it is tested. If the associative deficit is only replicable in specific situations with specific methods and stimuli, it may be due to an issue with yes-no recognition for pairs rather than an issue of binding two items together.

One of the limitations of this study was using undergraduate students as our younger adult sample because they are typically more educated and from a higher socioeconomic status than a more representative sample of younger adults. Additionally, we only tested recombined pairs and semantically similar pairs in the associative test, so all of the answer choices were familiar in some way. In order to test whether the false alarm effect is driven by feelings of familiarity versus a lack of memory for the original pairing, a completely novel pair would need

to be included as a lure in the associative test. If older adults tended to choose this option as often as the other lure types, it would provide evidence for a lack of memory of the original pairing. We also only employed one type of test (the 4AFC), so it is unclear how the manipulation of semantic similarity would affect the results in the typical associative deficit paradigm.

Future research should further explore the types of errors that we observed in older adults on the associative test. An eye-tracking study could illuminate age-related differences in gaze patterns and time spent dwelling on stimuli at encoding and retrieval, perhaps revealing different patterns for item and associative trials. Another way of examining age differences in recognition memory is through fMRI imaging (Saverino et al., 2016). Older adults may recruit more areas of the brain to study and retrieve words, which may lead to stimulation of the other areas and therefore, more false alarms. Also, using different test types and manipulating distal and proximal familiarity could reveal more specific tendencies about the role of familiarity in associative memory, especially for older adults. Finally, using a free recall test format might reveal unique types of memory errors or accuracy on the associative task.

Conclusions

This study makes a novel contribution to the growing body of research on the associative deficit through investigating the role of familiarity in associative memory. Although we did not replicate the associative deficit in older adults, the analysis of error types within the associative test revealed an interesting pattern for younger and older adults. Older adults were significantly more likely than younger adults to choose recombined pairs and also demonstrated an increased sensitivity to intact related pairs. These results primarily provide support for source monitoring errors as the root of the false alarm effect, but provide some evidence for fuzzy trace theory as well. Taken together, our results demonstrate the importance of familiarity in the memory

processes of older adults. Future research may help parse out the exact nature of this relationship and shed light onto how test type affects the associative deficit.

References

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*(4), 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

Boywitt, C. D., Kuhlmann, B. G., & Meiser, T. (2012). The role of source memory in older adults' recollective experience. *Psychology and Aging, 27*(2), 484-497. doi:10.1037/a0024729

Brainerd, C. J. (2013). Developmental reversals in false memory: A new look at the reliability of children's evidence. *Current Directions in Psychological Science, 22*(5), 335-341. doi:10.1177/0963721413484468

Brainerd, C. J., & Reyna, V. F. (2002). Fuzzy-trace theory and false memory. *Current Directions in Psychological Science*, *11*(5), 164-169. doi:10.1111/1467-8721.00192

Brainerd, C. J., Reyna, V. F., & Ceci, S. J. (2008). Developmental reversals in false memory: A review of data and theory. *Psychological Bulletin, 134*(3), 343-382. doi:10.1037/0033-2909.134.3.343

Dehon, H., & Brédart, S. (2004). False memories: Young and older adults think of semantic associates at the same rate, but young adults are more successful at source monitoring. *Psychology and Aging, 19*(1), 191-197. doi:10.1037/0882-7974.19.1.191

Dywan, J., Segalowitz, S. J., & Webster, L. (1998). Source monitoring: ERP evidence for greater reactivity to nontarget information in older adults. *Brain and Cognition, 36*(3), 390-430. doi:10.1006/brcg.1997.0979

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.

Fox, M. C., Baldock, Z., Freeman, S. P., & Berry, J. M. (2016). The role of encoding strategy in younger and older adult associative recognition: A think-aloud analysis. *Psychology and Aging*, *31*(5), 471-480. doi:10.1037/pag0000111

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*(1), 3-28. doi:10.1037/0033-2909.114.1.3

Kersten, A. W., Earles, J. L., Curtayne, E. S., & Lane, J. C. (2008). Adult age differences in binding actors and actions in memory for events. *Memory & Cognition, 36*(1), 119-131. doi:10.3758/mc.36.1.119

Light, L. L., Patterson, M. M., Chung, C., & Healy, M. R. (2004). Effects of repetition and response deadline on associative recognition in young and older adults. *Memory & Cognition, 32*(7), 1182-1193. doi:10.3758/bf03196891

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*(5), 1170-1187. doi:10.1037/0278-7393.26.5.1170

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. http://www.usf.edu/FreeAssociation/.

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*(1), 104-118. doi:10.1037/0882-7974.23.1.104

Patterson, M. M., & Hertzog, C. (2010). The effects of age in four alternative forced-choice item and associative recognition tasks. *Psychology and Aging, 25*(1), 235-238. doi:10.1037/a0016046

Patterson, M. M., Light, L. L., Van Ocker, J. C., & Olfman, D. (2009). Discriminating semantic from episodic relatedness in young and older adults. *Aging, Neuropsychology, and Cognition, 16*(5), 535-562. doi:10.1080/13825580902866638

Roediger, H. L., & Mcdermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(4), 803-814. doi:10.1037//0278-7393.21.4.803

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

Saverino, C., Fatima, Z., Sarraf, S., Oder, A., Strother, S. C., & Grady, C. L. (2016). The associative memory deficit in aging is related to reduced selectivity of brain activity during encoding. *Journal of Cognitive Neuroscience, 28*(9), 1331-1344. doi:10.1162/jocn_a_00970

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

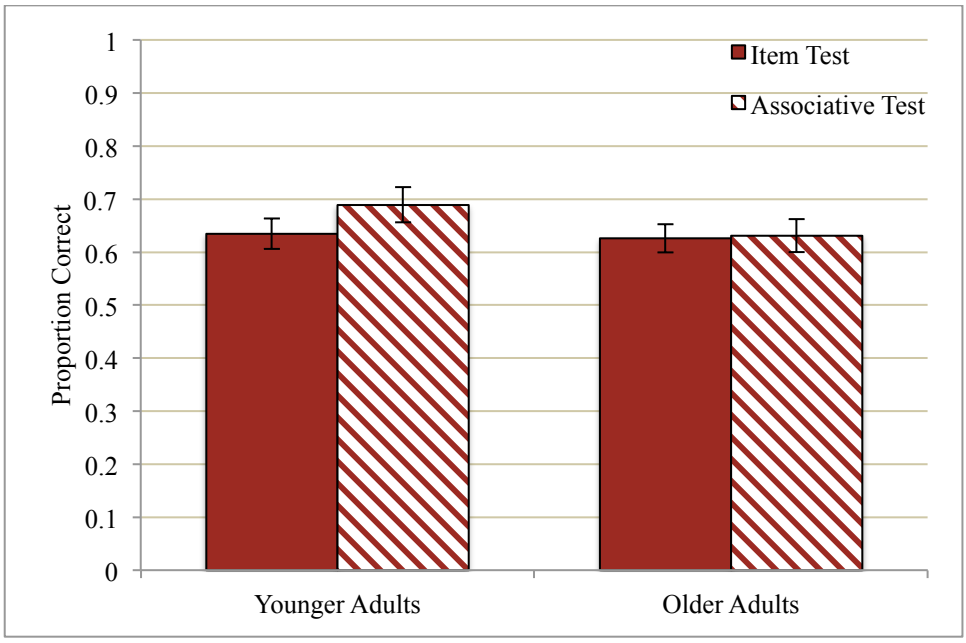

Figure 1. Accuracy by Age Group and Test Type. No age differences were found with regard to overall accuracy. Accuracy for the associative test was surprisingly better than the item test, and no interaction between age group and test type was observed.

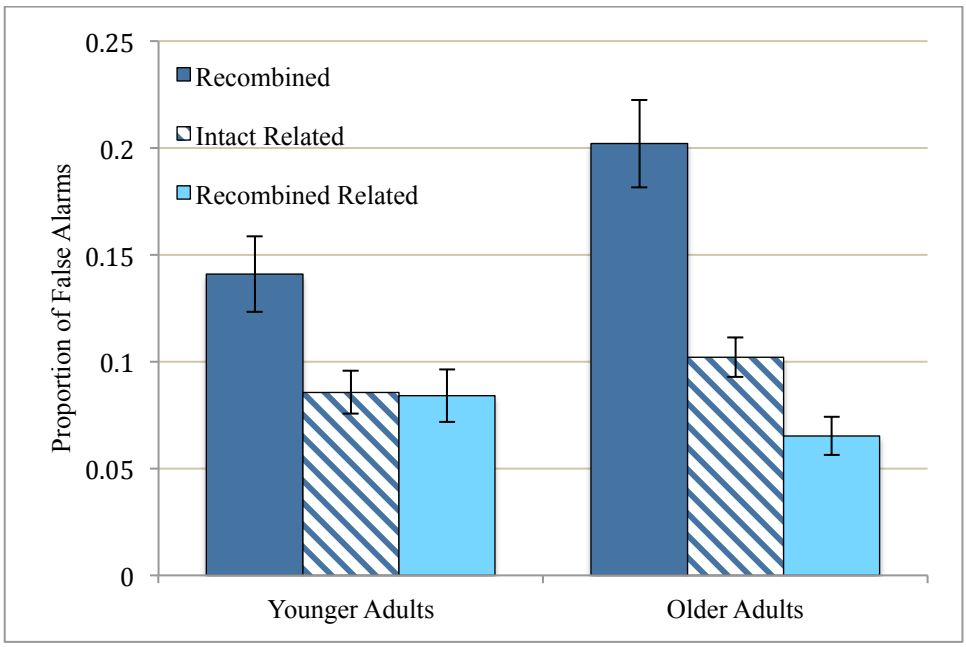

Figure 2. Types of Errors by Age Group in the Associative Test. Overall, there was a main effect for error type, with recombined pairs being the most common, followed by intact related and recombined related pairs. The FAR for recombined pairs was significantly higher for older adults than younger adults. For older adults, all three mean proportions of false alarms were different from each other, while for younger adults, the proportions were only different between recombined and recombined related pairs. This indicates that older adults may be more sensitive to semantic familiarity with the intact related pairs.

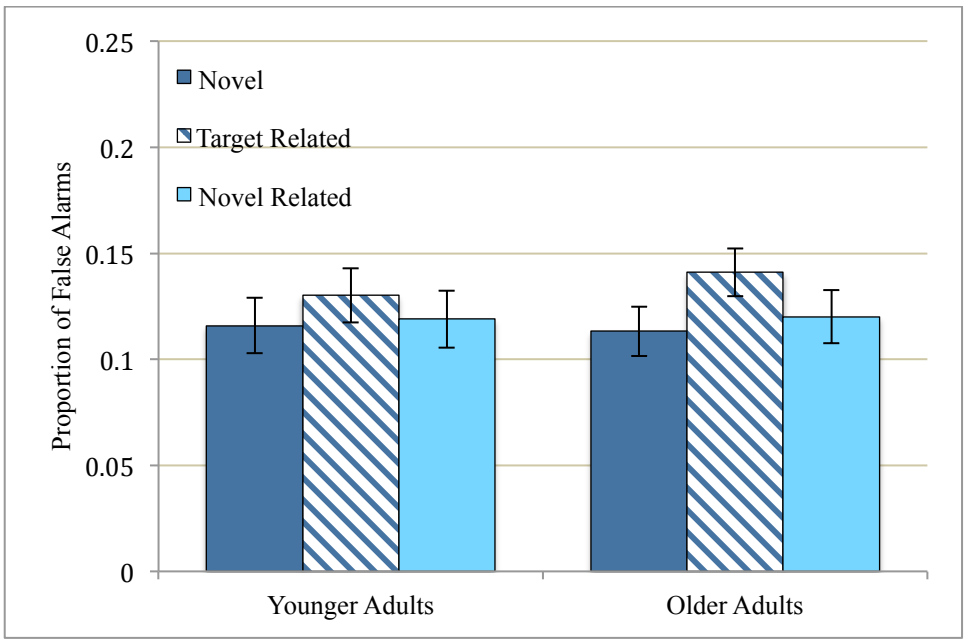

Figure 3. Types of Errors by Age Group in the Item Test. Neither age differences nor error type differences were found. The interaction effect between age group and error type was also not significant.