

ATTEMPTS TO REDUCE THE HIGH FALSE ALARM RATE IN OLDER ADULTS'
ASSOCIATIVE MEMORY

A Thesis Presented to
The Faculty of the Graduate School
At the University of Missouri

In Partial Fulfillment
Of the Requirements for the Degree
Masters of Arts

By
HOPE C. FINE
Moshe Naveh-Benjamin, Thesis Supervisor

MAY, 2014

The undersigned, appointed by the dean of the Graduate School,
have examined the thesis entitled

ATTEMPTS TO REDUCE THE HIGH FALSE ALARM RATE
IN OLDER ADULTS' ASSOCIATIVE MEMORY

Presented by Hope C. Fine

A candidate for the degree of

Master of Arts

And hereby certify, that in their opinion, it is worthy of their acceptance

Dr. Moshe Naveh-Benjamin

Dr. John Kerns

Dr. Linda Day

Dr. Nelson Cowan

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Moshe Naveh-Benjamin for his insight and encouragement throughout this research process as well as my committee members, John Kerns, Linda Day, and Nelson Cowan. I would like to express gratitude to members of The Memory and Cognitive Aging Lab for their constructive remarks. I greatly appreciate all of the loving support my family has provided for me. Finally, I would like to thank the Department of Psychological Sciences for helping to fund this research.

TABLE OF CONTENTS

| | |
|--|-----|
| ACKNOWLEDGEMENTS..... | ii |
| TABLE OF CONTENTS..... | iii |
| LIST OF FIGURES | iv |
| ABSTRACT..... | vi |
| INTRODUCTION | 1 |
| Dual-Process Model of Recognition Memory at Retrieval | 2 |
| High False Associative Recognition Memory in Older Adults..... | 3 |
| The Effects of Item Repetition on Associative Recognition..... | 5 |
| Recall-to-Reject..... | 6 |
| Purpose of the Current Studies..... | 8 |
| EXPERIMENT 1 | 8 |
| Method | 10 |
| Participants | 10 |
| Design..... | 11 |
| Materials | 11 |
| Procedure..... | 15 |
| Results | 16 |
| Discussion | 25 |
| EXPERIMENT 2 | 28 |
| Method | 30 |
| Participants and Design..... | 30 |
| Materials | 30 |
| Procedure..... | 34 |
| Results | 35 |
| Discussion | 43 |
| GENERAL DISCUSSION | 54 |
| REFERENCES | 66 |
| APPENDIX..... | 70 |

LIST OF FIGURES

| | |
|--|----|
| Figure 1. Hypothesis Predictions..... | 10 |
| Figure 2a. Example of Stimuli during Training Phase for Experiment 1..... | 12 |
| Figure 2b. Example of Stimuli during Study Phase for Experiment 1..... | 13 |
| Figure 3. Example of Stimuli for Item (Left) and Associative Tests (Right) for Experiment 1..... | 14 |
| Figure 4. Mean Proportion Hits minus Proportion False Alarms in Experiment 1 with Error Bars Indicating Standard Errors..... | 19 |
| Figure 5. Mean Proportion Hits in Experiment 1 with Error Bars Indicating Standard Errors..... | 21 |
| Figure 6. Mean Proportion False Alarms in Experiment 1 with Error Bars Indicating Standard Errors..... | 23 |
| Figure 7. Mean Proportion Hits minus Proportion False Alarms Hypothesis Predictions in Experiment 1 with Error Bars Indicating Standard Errors..... | 24 |
| Figure 8. Mean Proportion False Alarms Hypothesis Predictions in Experiment 1 with Error Bars Indicating Standard Errors..... | 25 |
| Figure 9a. Example of Stimuli during Training Phase for Experiment 2..... | 31 |
| Figure 9b. Example of Stimuli during Study Phase for Experiment 2..... | 32 |
| Figure 10. Example of Stimuli for Item (Top) and Associative Tests (Bottom) for Experiment 2 | 34 |
| Figure 11. Mean Proportion Hits minus Proportion False Alarms in Experiment 2 with Error Bars Indicating Standard Errors..... | 38 |

| | |
|---|----|
| Figure 12. Mean Proportion Hits in Experiment 2 with Error Bars Indicating Standard Errors..... | 40 |
| Figure 13. Mean Proportion False Alarms in Experiment 2 with Error Bars Indicating Standard Errors..... | 43 |
| Figure 14a. Mean Proportion Hits minus Proportion False Alarms in Experiment 2 Follow-up Analyses with Error Bars Indicating Standard Errors..... | 48 |
| Figure 14b. Mean Proportion Hits in Experiment 2 Follow-up Analyses with Error Bars Indicating Standard Errors..... | 49 |
| Figure 14c. Mean Proportion False Alarms in Experiment 2 Follow-up Analyses with Error Bars Indicating Standard Errors..... | 51 |
| Figure 15. Mean Proportion Hits minus Proportion False Alarms Hypothesis Predictions in Experiment 2 Follow-up Analyses with Error Bars Indicating Standard Errors..... | 52 |
| Figure 16. Mean Proportion Hits minus Proportion False Alarms Hypothesis Predictions in Experiment 2 Follow-up Analyses with Error Bars Indicating Standard Errors..... | 53 |
| Figure 17. Mean Proportion Hits minus Proportion False Alarms Hypothesis Predictions (Experiment 1 and Experiment 2 Follow-up Analyses) with Error Bars Indicating Standard Errors..... | 63 |
| Figure 18. Mean Proportion False Alarms Hypothesis Predictions (Experiment 1 and Experiment 2 Follow-up Analyses) with Error Bars Indicating Standard Errors..... | 64 |

ABSTRACT

Older adults' seem to have a special difficulty binding components together in their episodic memory (Naveh-Benjamin, 2000). This finding, known as the age-related associative deficit, seems to be driven to a large degree by high false alarm rates in the associative test. These high false alarm rates could be due to the effects of high item familiarity and low recall-to-reject processing. The purpose of the current set of experiments was to examine if manipulating item familiarity and recollection could decrease the false alarm rate in the associative test, especially for older adults. We predicted that by decreasing item familiarity and increasing the ability to use recall-to-reject processes false alarm rates will be reduced. In Experiment 1, younger and older adults were tested for their item and associative recognition memory after viewing product-price pairs. Item familiarity was manipulated by having a pre-exposure phase of individual products and prices. Products were paired with either an underestimated price, a market value price, or an overestimated price. Recall-to-reject was manipulated via preexisting schematic knowledge where product-price pairings were either shown in the same price category or in a different price category between study and test. We hypothesized that older adults would use low familiarity and previous schematic knowledge of product-price pairings in order to learn product-price associations at encoding, and in particular, to reject recombined pairs at retrieval. The purpose of Experiment 2 was to strengthen the results of Experiment 1 by using a different set of stimuli, face-name pairs, and a somewhat different procedure. Names were paired with either a younger face or an older face. Again, we hypothesized that older adults' false alarm rates in the associative test may decline if they can rely on schematic knowledge at

encoding. In this study, schematic information is age of the face stimuli (young or old). Results indicate that low item familiarity (Experiment 2) and preexisting schematic knowledge (Experiments 1 and 2) increase older adults' ability to recall-to-reject recombined product-price and face-name pairs at retrieval, resulting in a reduced associative memory deficit.

INTRODUCTION

Memory abilities differentially decline with age depending on the type of information being processed. In particular, memory for general knowledge (semantic memory) stays relatively intact with age (Kausler & Puckett, 1980); whereas episodic memory or, remembrance of a specific, personal event, declines (Zacks, Hasher, & Li, 2000). For example, older adults may remember obscure factoids from years past, but forget what they had for breakfast that morning. Within general episodic memory, a specific age-related decline has been found in tasks that require memory for associations. Two hypotheses have been suggested for this age-related decline of associative memory: a binding deficit at encoding and a recollection process at retrieval.

The Associative Deficit at Encoding

Chalfonte and Johnson (1996) suggested a binding-deficit by demonstrating that older adults' poor memory performance for associations is not due to the inability to remember individual features, but to the failure to bind these features together in memory. In their experiment, younger and older adults studied an array of uniquely colored picture-objects each presented at a different location on a gridded background. Participants were given a recognition memory test for each individual feature (e.g. item, color, or location) and for bounded-features (e.g. item and color; item and location). Chalfonte and Johnson concluded that under intentional learning age-related differences are evident for associative memory, whereas recognition memory for the individual color and item features did not significantly differ between younger and older adults.

Further investigating the binding-deficit, Naveh-Benjamin (2000) proposed the associative deficit hypothesis (ADH) suggesting that the age-related binding deficit is not

caused solely by the inability to bind specific features but rather occurs due to a more general binding deficit for all types of information, such as the relation between two focal items. In addition, the ADH predicts that it is not just the inability to bind these features together at encoding, but also the lack of ability to retrieve these associations at test. In order to test this hypothesis, Naveh-Benjamin (2000) conducted a series of experiments in which participants were asked to study word and nonword pairs and were tested on their memory for items and associations. The results for younger and older adults' show a differential age-related decline in item and associative memory suggesting an associative deficit. Further research has replicated and extended these findings supporting the associative deficit hypothesis using several different types of item components (e.g. name-face, face-scene, person-activity) as well as different modalities (see a meta-analysis by Old & Naveh-Benjamin, 2008a) suggesting that the phenomenon is quite robust.

Dual-Process Model of Recognition Memory at Retrieval

A mediator of the associative deficit could be the tendency of older adults' to rely on automatic processes at retrieval. Yonelinas (2002) suggested a dual-process model of recognition memory comprised of recollection and familiarity retrieval processes. A conscious, attention demanding process, recollection requires contextual retrieval. Familiarity judgments are more automatic and can be made without remembering the specific episode at retrieval.

Different types of techniques have been used to measure familiarity and recollection. In a process dissociation procedure, participants are presented with an inclusion and an exclusion list and are told only to recollect those items from the

inclusion list. In this task, participants must overcome automatic retrieval processes to consciously remember which items were presented on the inclusion list (Jennings & Jacoby, 1997). Another task commonly used to distinguish between recollection and familiarity judgments is the remember/know task. Subsequently deciding that an item was previously shown in the study list, participants must assess whether they recollect that particular episode with perceptual detail (remember) or that they believe they saw the item previously but cannot remember the contextual instance (familiar; Gardnier, Ramponi, & Richardson-Klavehn, 1998; Tulving, 1985). This task allows the experimenter to dissociate between familiarity and recollection. A yes/no or old/new recognition task can also be used (Kilb & Naveh-Benjamin, 2011). A final technique commonly used to dissociate recollection from familiarity is receiver operating characteristic (ROC) modeling using confidence-rating judgments. After producing confidence judgments, a ROC plot is generated based on the cumulative probabilities as a function of proportion of hits (a studied item called “old”) to false alarms (a new item called “old”; Howard, Bessette-Symmons, Zhang, & Hoyer, 2006). Using all of these techniques, previous research has repeatedly found that recollection retrieval processes severally decline with age, whereas older adults’ ability to make familiarity judgments does not (Kilb & Naveh-Benjamin, 2011; Light, Prull, LaVoie, & Healy, 2000).

High False Associative Recognition Memory in Older Adults

An important contributor to the associative deficit of older adults is a high false alarm rate in the associative test (Castel & Craik, 2003; Cohn, Emrich, & Moscovitch, 2008; Kilb & Naveh-Benjamin, 2011). For example, Old and Naveh-Benjamin (2008b) presented younger and older adults at study with video clips of various different people

each performing a different action and were later tested on their item and associative recognition memory. There was an age-related decline in associative memory compared to item memory, with memory accuracy being measured as proportion of hits (accepting a target pair correctly) minus proportion of false alarms (wrongly accepting a rearranged lure pair). When analyzing the hits and false alarms separately, these results suggest that the age-related decline in associative memory was driven by the high false alarm rate and not as much by the reduced hit rate. Shing, Werkle-Bergner, and Lindenberger (2008) also found that age-related differences in the associative test were due to high false alarms more so than reduced hit rates even when given practice and a memory strategy.

The high false alarm rate in older adults' associative memory could be due to either an increase in item familiarity or a decrease in recall-to-reject (to be discussed). Older adults rely more on unconscious retrieval processes because of a decline in their strategic retrieval processing, particularly memory monitoring (Fandakova, Shing, & Lindenberger, 2013b) as well as on gist-based memory (Brainerd & Reyna, 1998). In the associative test, all of the item-components were previously seen in the study phase, however only half of the pairs are intact. The other half of the pairs shown at test is rearranged. A rearranged pair is comprised of two item-components that were not previously presented together and were presented with different pairs at study. Participants need to successfully monitor and use verbatim memory to determine which item-components were shown together at study (intact pair) and which were recombined. Since all of the item-components look familiar at test, older adults, relying on their unconscious retrieval processes, may erroneously endorse the recombined pairs as seen before.

The Effects of Item Repetition on Associative Recognition

The dual-process model assumes that familiarity affects associative recognition memory differentially for younger and older adults. One way familiarity can be manipulated is by repeating items or pairs at study. In a phenomenon known as the *ironic* or *mirror effect* (Jacoby, 1999), repetition of study items or pairs causes an increase in hits in the item test for younger and older adults. While the effects of repetition raises hit rates for the younger and older adults in the associative test as well, the effects of repetition decrease false alarms for younger adults but increase false alarms for older adults in the associative test (Light, Chung, Pendergrass, & Van Ocker, 2006; Light, Patterson, Chung, & Healy, 2004).

Light, Patterson, Chung, and Healy (2004) examined the effects of pair familiarity on memory recognition. Word pairs were presented either once or four times. Pair repetition increased hits for intact pairs on an associative recognition test for younger and older adults. False alarms decreased for rearranged pairs for younger adults with repetition, whereas for older adults' repetition of pairs increased false alarms. They concluded that repetition does not help older adults' associative memory. Kilb and Naveh-Benjamin (2011) challenged that conclusion by examining the effects of pure item and pure pair familiarity on associative memory recognition in younger and older adults. They predicted that item repetition should increase false alarms in the associative test whereas pure pair repetition (the effects of pair repetition after item repetition effects are controlled) should decrease false alarms. Item repetition increased hit rates for younger and older adults in the item and the associative test. When items were repeated at study, older adults significantly made more false alarm errors on the associative test than

younger adults. Pair repetition decreased older adults' false alarms in the associative test relative to item repetition suggesting that it is item familiarity that increases the false alarm errors made by older adults in the associative test. Moreover, Kilb and Naveh-Benjamin suggest that item familiarity, not pair familiarity, causes the high false alarm rate in older adults' associative memory. Since the repeated item-component appears more familiar during an associative recognition test, older adults' are likely to accept the rearranged pair because they remember seeing each of the individual components. This suggests that the associative deficit could be due to older adults' reliance on unconscious, automatic retrieval processes (Light, Prull, La Voie, & Healy, 2000).

In younger adults at short response deadlines, familiarity caused by repetition simulates older adults' false alarm rate in the associative test. However, if younger adults are given enough time to use conscious retrieval, repetition decreases false alarm rates (Light, Patterson, Chung, & Healy, 2004). Even when older adults are given an extra-long response deadline, repetition increases false alarms in the associative task. Together these findings suggest that younger adults use conscious retrieval if given enough time at retrieval, whereas older adults still rely on unconscious retrieval even when given enough time to respond.

Recall-to-Reject

Recall-to-reject is the ability to recognize during retrieval that two items were not paired together because you remember one being paired with another stimulus at encoding. The ability to use recall-to-reject can counteract any feelings of high familiarity since recall-to-reject uses recollection at retrieval. The use of the recall-to-reject strategy can be measured by a decline in false alarms to distractors (Brainerd,

Reyna, & Kneer, 1995). Many different methods have been used to investigate recall-to-reject, such as the conjoint recognition model (Jones & Jacoby, 2001) and the simplified conjoint recognition model (Schmid, Herholz, Brandt, & Buchner, 2010). Receiver operating characteristic analysis has also examined the role of recall-to-reject in associative recognition tasks (Rotello & Heit, 2000). Self-report measures, such as post-test questionnaires, can also be used to engage a sense of the usage of the recall-to-reject strategy (Gallo, 2004).

Manipulations of recall-to reject can be list length (Odegard, Lampinen, & Toggia, 2005), attention, (Jones & Jacoby, 2001), presentation duration and item duration (Leding & Lampinen, 2009), and response deadline (Cohn & Moscovitch, 2007). For example, dividing attention at study or at test lowers the use of the recall-to-reject strategy at retrieval by making it harder to access the original study pairings as well as reject those that are recombined; recall-to-reject requires conscious attention. The use of the recall-to-reject strategy at test when given longer study durations is greater because at longer durations participants can encode the information better. When given a shorter response deadline, there is not much time compared to a longer response deadline to use recollection and hence the recall-to-reject strategy.

Several of the studies that examined the use of the recall-to-reject strategy by older adults suggest that they have difficulties using it (e.g. Cohn, Emrich, Moscovitch, 2008; Gallo, Bell, Beir, & Schacter, 2006; Healy, Light, & Chung, 2005). Cohn et al. examined strategic retrieval demands (recall-to-reject and recall-to-accept) on older adults' associative memory. Recall-to-reject is a high-demanding conscious retrieval strategy, whereas another type of strategy, recall-to-accept, the ability to accept old items

or pairs rather than reject new ones, has lower demands by supplementing familiarity. Their results show that older adults' associative memory is severely impaired when given a task that place high demands on retrieval process, whereas no age-related differences were found on tasks that utilized recall-to-accept.

Purpose of the Current Studies

The purpose of the current set of experiments is to manipulate item familiarity and recall-to-reject opportunity in order to assess how they may affect encoding and retrieval processes involved in associative memory. By manipulating item familiarity through a pre-exposure or training phase, we predicted that older adults will show a smaller associative deficit evident by lower false alarm associative recognition for pairs that included items with no prior pre-exposure. This is so since the reduction of the familiarity of the two components in the association will make the recombined pairs less familiar to older adults, decreasing their reliance on familiarity. By manipulating the opportunity to rely on preexisting schematic information we expected to assist recall-to-reject processes in older adults, which would lead to decrease in their high false alarm rate and consequently to a reduction in their associative deficit.

EXPERIMENT 1

Castel (2005) examined younger and older adults' memory for prices of grocery products using a cued-recall task. He did not find any age-related differences when products were paired with a realistic or market value price. However, when products were paired with an unrealistic arbitrary price, age-related differences did emerge. This suggests that when older adults can rely on previous semantic knowledge of a general range of prices, memory performance for associations will improve.

In the first experiment, we tested younger and older adults' item and associative memory for product-price pairs while manipulating item familiarity (more familiar vs. less familiar) and recall-to-reject (price category match vs. price category mismatch, see below). More specifically, a repetition manipulation that makes products and prices more familiar, should lead to both younger and older adults' performing well on the item test, but older adults performing worse in the associative test (particularly displaying a high false alarm rate for recombined product-price pairs). As for the recall-to-reject manipulation, for the purpose of this study, a price category match is defined as a recombined product-price pair at test where the original product appears with a same-range price to that shown at study. For example, if the study phase dollar value was overpriced, the product will appear at test with another overpriced dollar value; a price category mismatch is defined as a recombined product-price pair at test where the original product appears with a very different price than that shown at study (for example, if the study phase dollar value was overpriced, the product will appear at test with an underpriced value or with a market value, see Method section for details). The price category mismatch manipulation of the price estimate should aid in the recall-to-reject strategy. For example, if a participant sees a product-price pair at study that is market value, and then sees that product paired at test with an overestimated price, the participant can use recall-to-reject at retrieval to remember that product was not paired with an overestimated price because he or she can remember that the product was paired with a market value price therefore using the recall-to-reject strategy. It is harder to use the recall-to-reject strategy in the match recombined pair condition because the product-price pairing at study is in the same schematic category at retrieval. Moreover, we

predicted that the best memory performance would be shown when items have low familiarity and provide for a high ability to use the recall-to reject strategy (lower left quadrant-see Figure 1) and the worst performance would be when items have high familiarity and low ability to use the recall-to-reject strategy (upper right quadrant-see Figure 1).

Figure 1. Hypothesis Predictions

| | | Recall-to Reject | |
|------------------|------|--|---|
| | | High | Low |
| Item Familiarity | High | High-Familiarity + High Recall-to-Reject | High-Familiarity + Low Recall-to-Reject |
| | Low | Low-Familiarity + High Recall-to-Reject | Low-Familiarity + Low Recall-to-Reject |

Method

Participants

The experiment was conducted with a total of 43 younger adults (17 males and 26 females) with a mean age of 19.6 ($SD = 1.97$) and a mean education level of 13.51 years ($SD = 1.76$). This experiment also included 40 older adults (16 males and 24 females) with a mean age of 74.23 ($SD = 6.28$), with a mean education level of 14.74 years ($SD = 1.58$). Older adults had significantly higher level of education ($M = 14.74$, $SD = 1.58$;

$t(81) = 3.32, p = .001$) than younger adults ($M = 13.51, SD = 1.76$). The younger adults were students from the University of Missouri who volunteered to receive partial course credit for their participation. Healthy older adults without any previous cognitive functioning disabilities were recruited from the local community and were compensated \$15 for their participation. The education level of the older adults did not exceed completion of a Masters Degree.

Design

In order to assess the specific hypothesis we are interested in we used a non-complete factorial design, which allow for several analyses, each of which assessed specific hypotheses. We used a 2 (age: younger adults and older adults) X 2 (test: item including both products and prices, and associative) X 2 (familiarity: more familiar and less familiar) X 2 (product-price schematic relationships in study and test: match and mismatch, only in the associative test). Test, familiarity, and product-price schematic relationships in study and test were within-subject factors.

Materials

This experiment was run using the E-prime program software. Stimuli were retail products taken from an online Google search. Products included grocery products, small appliances, large appliances, furniture, and technology items. Prices were assigned after collecting price pilot data. In the price pilot study, younger and older adults were asked to give a price range for a given product as well as decide whether a given price was underestimated, market value, or overestimated. For the study phase, each product was presented at the center of the 15 inch computer screen using the computer's default resolution set at 640 x 480. Products were 230 x 280 pixels. A label of what the product

was depicting was placed under the product in size 18-point Courier New for clarification. Prices were rounded to the whole dollar value using the ranges from the price pilot data. For each study phase, prices were centered above the product in size 22-point Courier New bolded font.

Figure 2a. Example of Stimuli during Training Phase for Experiment 1

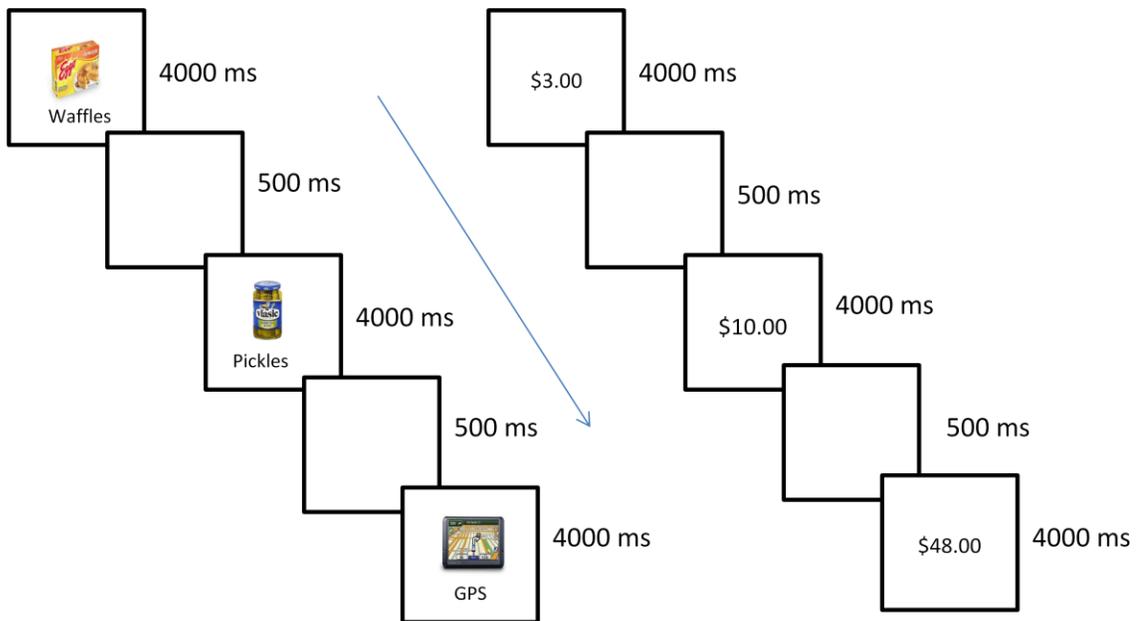
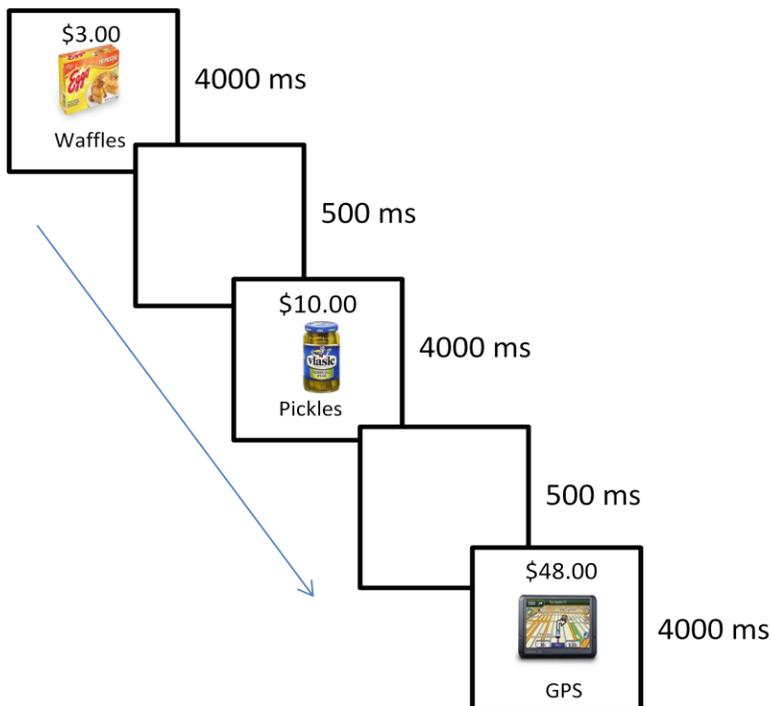


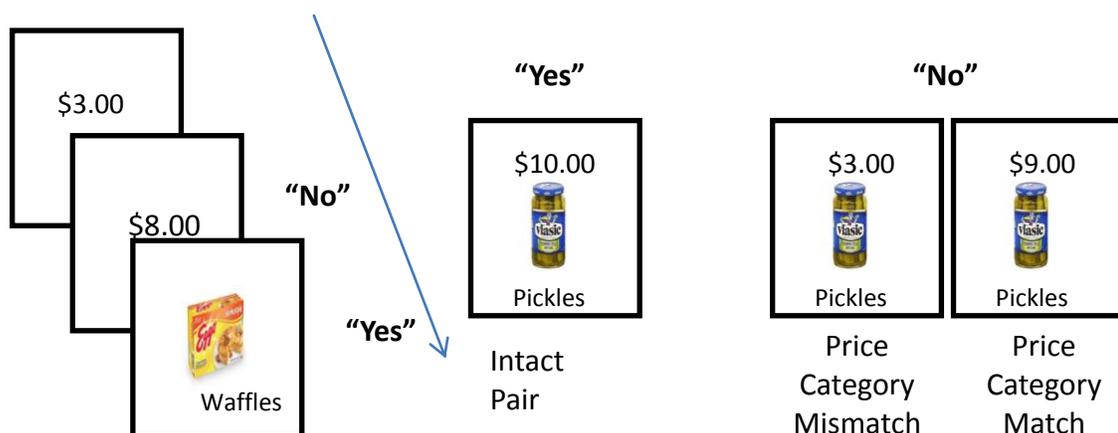
Figure 2b. Example of Stimuli during Study Phase for Experiment 1



There were 10 trial blocks. Five of the trial blocks contained a training phase directly before each study phase (in order to manipulate item familiarity). For each training phase products and prices were blocked and presented separately twice in a spaced manner (see Figure 2a). Trial blocks that did not have a training phase proceeded directly into the study phase. Each study block consisted of 12 product-price pairs presented twice in a spaced manner (see Figure 2b). The design of the study defined a price category match as one where the price at test is similar, approximately within 10% of the given price at study; whereas a price mismatch at test was defined as approximately 3 times greater or one-third of the original studied price. Since there was a wide range of stimuli used for this experiment and many of the smaller items were

estimated to be about the same price, the assignments of product-price pairs were asymmetrical with more larger products (i.e. technology and furniture) being paired with an underestimated price and smaller products being paired with an overestimated price. To ensure enough stimuli in each combination of independent variables, each of the lists had nine or 10 pairs designated to the associative test and two pairs designated to the item test as well as an additional two products and two prices not shown at study. Therefore, due to the design of the study, four of the lists had filler study-pairs in order to keep each list length the same. Throughout all 10 blocks there was an equal number of underestimated, market value, and overestimated product-price pairs, each was presented equally in the more familiar and less familiar conditions.

Figure 3. Example of Stimuli for Item (Left) and Associative Tests (Right) for Experiment 1. The diagram on the right shows an example of an intact pair, a price mismatch recombined pair, and a price match recombined pair (left to right).



Overall, collapsed across all 10 trial blocks, there were a total of 96 pairs used in the associative test. Forty-eight pairs were intact and 48 were recombined at test. Forty-eight of the pairs included products and prices shown during the training phase, making them more familiar, and 48 were not. Half of the recombined pairs were a price category match and half were a price category mismatch. There were an equal number (eight pairs) of underestimated, market value, and overestimated price matches in the more familiar and less familiar conditions, for a total of 24. There was also an equal number of price category mismatches. There were six types of price category mismatches (from study to test): underestimate to overestimate, underestimate to market value, market value to overestimate, market value to underestimate, overestimate to market value, and overestimate to underestimate. Each type was presented twice over the 10 study blocks in the more familiar (12) and less familiar (12) condition for a total of 24 times. More familiar items were presented a total of four times and less familiar were presented two times. Familiarity, test order, and list order were counterbalanced between subjects.

Procedure

After obtaining participants' informed consent, the task's instructions were given followed by a practice trial. The practice trial included a training phase. The participants were instructed to remember the products and prices from the training phase, but were told that they were only being tested on the product-price pairs presented during the study phase. During the training phase, products and prices were presented separately for 4 s, with an inter-stimulus interval (ISI) of 500 ms. Participants were told to repeat aloud the product or the price to strengthen encoding. A 30 s break was given between the pre-exposure and the study phase. After the pre-exposure phase, the study phase was given.

On trials that did not have a pre-exposure phase, the study phase started immediately after the instructions were given to remember the study list. For the study phase of each of the 10 lists, participants were told to study the products, prices, and their pairings equally in order to prepare for two later tests, for which they will have to decide if the stimulus was a studied item (target/intact pair) or a new item (distractor/recombined pair). Each product-price pair during the study phase was presented for 4 s, with an ISI of 500 ms. Participants were told to repeat each product-price pair aloud, first stating the product and then stating the price. An interpolated activity between the study and the test phase consisted of counting backwards by three's from the digit given by the experimenter. Participants were then asked to complete the item and the associative recognition tests for the list (see Figure 3). Since each list length was short, (including 12 pairs), the item recognition test included both product and price target and distractors. Subjects were told to press "V" with their left index finger for any item or pair that they saw previously in the study phase and to press "N" with their right index finger for any item or pair not previously shown during the study phase. Participants were also told that the percentage of targets and distractors is 50% each. Upon completion of the experiment, a post-test questionnaire and a debriefing were given.

Results

Memory accuracy measure, computed as proportion of hits minus proportion of false alarms, was calculated for each condition and averaged over participants of each age

group. The means and standard deviations are presented in Appendix A (Table A1) and can be seen in a graphical form (with standard errors) in Figure 4¹.

To assess the effects of item familiarity, a 2 (age: young vs. old) X 2 (test: item vs. associative) X 2 (item familiarity: more familiar vs. less familiar) ANOVA was conducted. The analysis indicated a main effect of age ($F(1, 81) = 5.02, p = .03, \eta_p^2 = .06$), with younger adults ($M = .67, SD = .12$) having better memory performance than older adults ($M = .61, SD = .11$). There was a main effect of test ($F(1, 81) = 26.17, p < .001, \eta_p^2 = .24$), with better memory performance for the item test ($M = .69, SD = .10$) than the associative test ($M = .59, SD = .18$). There was also a main effect of item familiarity ($F(1, 81) = 42.64, p < .001, \eta_p^2 = .35$), with more familiar items increasing memory performance to a greater degree ($M = .67, SD = .13$) than less familiar items ($M = .60, SD = .14$). The interaction of test and familiarity was significant ($F(1, 81) = 13.92, p < .001, \eta_p^2 = .15$) indicating that there was no difference in the associative test between more familiar ($M = .61, SD = .20$) and less familiar pairs ($M = .58, SD = .19; t(82) = 1.78, p = .08$), and a larger difference in the item test ($t(82) = 7.17, p < .001$) between more familiar ($M = .74, SD = .13$) and less familiar pairs ($M = .62, SD = .13$). No other interactions were significant.

The above-presented analysis did not reveal a significant age and test interaction, implying that older adults did not show a deficit in overall associative memory performance. One reason for this finding is that in the initial analysis presented above we also include our match-mismatch manipulation. In effect, this manipulation should

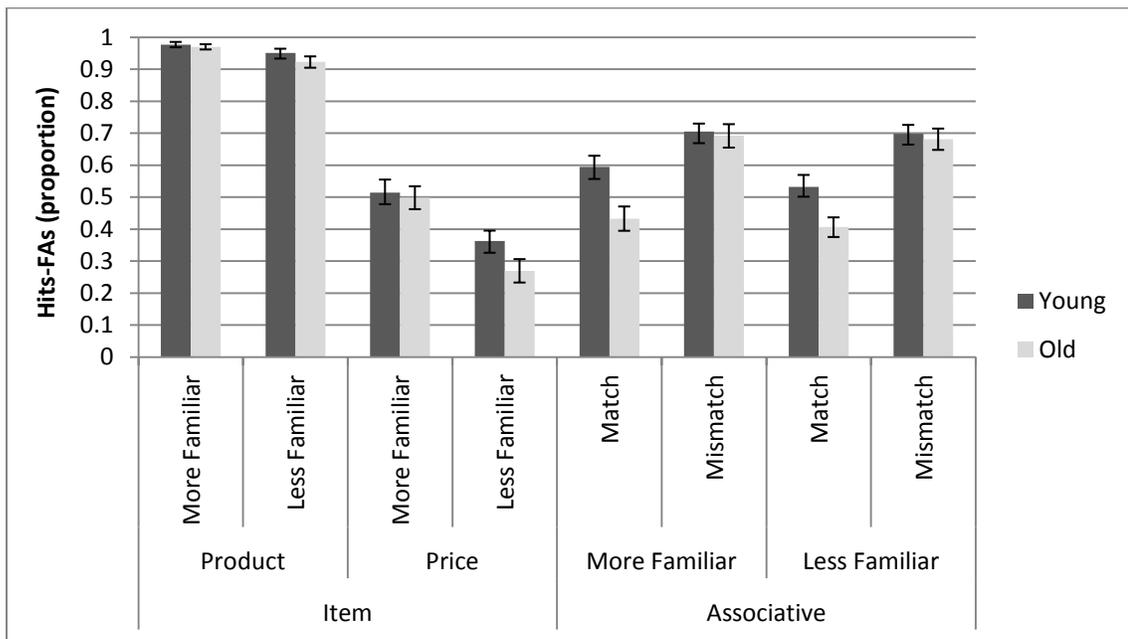
¹ Due to a programming error, three trials were thrown out for 10 younger and seven older adults and two were thrown out for one older adult. There was also a counterbalancing error in which there was two more market value price match more familiar than less familiar, but this was counterbalanced between participants.

decrease differences in the associative test, particularly for the mismatch condition, which could result in the absence of this interaction. Therefore, we ran an additional ANOVA comparing the item test and the match condition only in the associative test (the standard condition). An age-related associative deficit did occur as indicated by a significant age and test interaction ($F(1, 81) = 6.47, p = .01, \eta_p^2 = .07$), with older adults performing worse in the associative test relative to younger adults. Follow-up t-tests indicated older ($M = .67, SD = .10$) and younger adults ($M = .71, SD = .10$) performed similarly in the item test ($t(81) = 1.53, p = .13$), but that younger adults ($M = .55, SD = .21$) performed better than older adults ($M = .42, SD = .18$) in the associative test ($t(81) = 3.19, p = .002$)

A second analysis to assess the effects of change of schematic relationship and that of item familiarity was conducted on the results of the associative test only. The change of the schematic relationship is the change of price-estimate between study and test (match vs. mismatch). This analysis employed a 2 (age: young vs. old) X 2 (product-price schematic relationships in study and test: match vs. mismatch) X 2 (item familiarity: more familiar vs. less familiar). The ANOVA conducted showed a main effect of age ($F(1, 81) = 4.23, p = .04, \eta_p^2 = .05$), with younger adults ($M = .63, SD = .17$) having better memory performance than older adults ($M = .55, SD = .18$) and a main effect of the match-mismatch manipulation ($F(1, 81) = 177.04, p < .001, \eta_p^2 = .69$), with better memory performance in the price mismatch ($M = .69, SD = .18$) than the price match condition ($M = .49, SD = .21$). In addition, the only significant interaction was of change of schematic relationship and age ($F(1,81) = 13.82, p < .001, \eta_p^2 = .15$). Follow-up t-tests suggest that age differences were significant in the match condition ($t(81) =$

3.19, $p = .002$), but not significant in the mismatch condition ($t(81) = .55$, $p = .58$), suggesting that although older adults show a large associative memory deficit in the match condition ($M = .56$, $SD = .21$, and $M = .41$, $SD = .18$, for younger and older adults, respectively), their associative memory performance in the mismatch condition is similar to that of younger adults ($M = .70$, $SD = .16$, and $M = .68$, $SD = .21$, for younger and older adults, respectively).

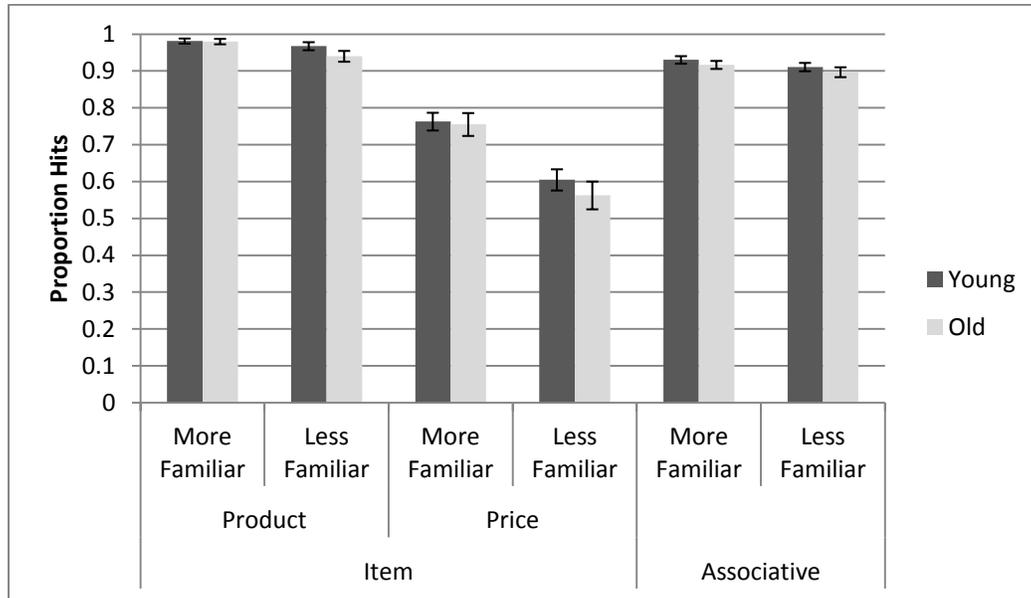
Figure 4. Mean Proportion Hits minus Proportion False Alarms in Experiment 1 with Error Bars Indicating Standard Errors



In order to determine whether it was the hits or the false alarms driving the above-presented results, similar ANOVAs to measure item familiarity and the change in the schematic relationship (match vs. mismatch) were run separately for the hits (Table A2 and Figure 5) and for false alarms (Table A3 and Figure 6) measures. For the analysis of hits, there was a main effect of test ($F(1, 81) = 83.28$, $p < .001$, $\eta_p^2 = .51$), with more hits

in the associative test ($M = .91, SD = .06$) than the item test ($M = .82, SD = .10$). This could be due to the relatively low performance in the price item test suggesting that older adults have trouble recalling individual arbitrary prices without contextual information. There was also a main effect of familiarity ($F(1, 81) = 88.02, p < .001, \eta_p^2 = .52$), with more hits when the items were made more familiar through the use of the pre-exposure phase ($M = .89, SD = .06$) than when they were less familiar and shown only in the study phase ($M = .81, SD = .10$). There was also a significant interaction of test and familiarity ($F(1, 81) = 37.4, p < .001, \eta_p^2 = .32$). Follow-up t-tests suggested significant differences in the item test hit rate between the more familiar ($M = .87, SD = .09$) and less familiar conditions ($M = .77, SD = .12; t(82) = 9.73, p < .001$), as well as in the associative test more familiar ($M = .92, SD = .08$) and less familiar conditions ($M = .90, SD = .07; t(82) = 2.79, p < .007$), although these familiarity differences were smaller in the latter test. None of the effects involving age were significant. In particular, there was no age by test interaction ($F(1, 81) = .07, p = .79, \eta_p^2 = .001$), with both younger ($M = .83, SD = .08; M = .92, SD = .06$, for item and associative tests, respectively) and older adults ($M = .81, SD = .11$ and $M = .90, SD = .07$) showing similar hit rates. This suggests that the hit rate is not driving the associative deficit in the current experiment.

Figure 5. Mean Proportion Hits in Experiment 1 with Error Bars Indicating Standard Errors

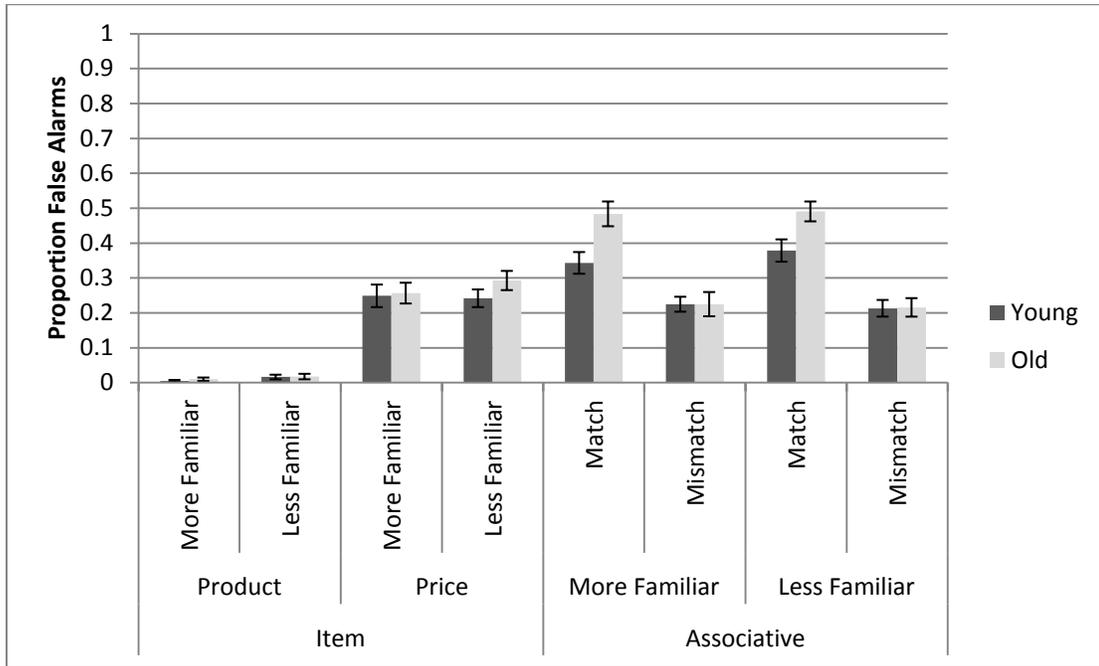


The ANOVA conducted on false alarms to assess the effects of item familiarity indicated a significant main effect of age ($F(1, 81) = 3.83, p = .05, \eta_p^2 = .05$), older adults making more false alarm errors ($M = .25, SD = .10$) compared to younger adults ($M = .21, SD = .10$), and a main effect of test ($F(1, 81) = 143.93, p < .001, \eta_p^2 = .64$), with more false alarms in the associative test ($M = .32, SD = .15$) than in the item test ($M = .14, SD = .08$). Similar to the analysis conducted on the overall memory performance, there was no significant age and test interaction. Again, we conducted a follow-up ANOVA comparing the false alarms in the item test and the match condition in the associative (the standard condition). The results of this analysis did show a significant age and test interaction ($F(1, 81) = 7.74, p = .01, \eta_p^2 = .09$), with older adults making more false alarms in the associative test. Follow-up t-tests confirmed these findings

suggesting that older adults ($M = .49, SD = .17$) made more false alarms than younger adults ($M = .36, SD = .19$) in the associative test ($t(81) = 3.15, p = .002$), but that in the item test younger ($M = .13, SD = .08$) and older adults ($M = .15, SD = .08$) performed similarly ($t(81) = 1.12, p = .27$).

The ANOVA that looked at the effects of change of schematic relationship and item familiarity on false alarms showed a main effect of age ($F(1, 81) = 3.87, p = .05, \eta_p^2 = .05$), with older adults making more false alarms ($M = .35, SD = .16$) than younger adults ($M = .29, SD = .14$). There was also a main effect of schematic relationship ($F(1, 81) = 177.04, p < .001, \eta_p^2 = .05$), with more false alarms being made in the match ($M = .42, SD = .19$) than the mismatch condition ($M = .21, SD = .15$). Similar to the previous analysis assessing the effects of item familiarity and match-mismatch manipulations on the proportion of hits minus false alarms, there was also a significant age and match-mismatch interaction ($F(1, 81) = 13.82, p < .001, \eta_p^2 = .15$). Follow-up t-tests revealed that the interaction was driven by significant age differences in the match condition ($t(81) = 3.15, p = .002$) and not in the mismatch condition ($t(81) = .21, p = .83$). These results suggest that older adults are benefitting more from the change of match ($M = .49, SD = .17$) to mismatch ($M = .22, SD = .18$) than younger adults ($M = .36, SD = .19; M = .22, SD = .13$).

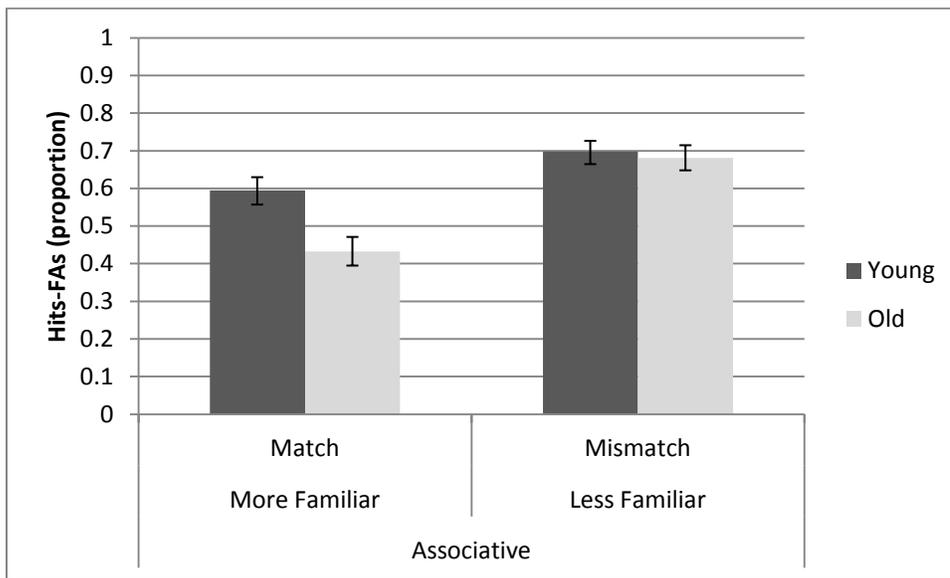
Figure 6. Mean Proportion False Alarms in Experiment 1 with Error Bars Indicating Standard Errors



Lastly, a final set of statistical analyses was conducted to assess our predictions of the contrast between the best associative memory performance in older adults, which should happen when items have low familiarity and the conditions allow for the use of recall-to-reject strategy (in the mismatch condition), and the worst associative memory performance in older adults, when items have high familiarity and the conditions do not allow for the use of a recall-to-reject strategy (in the match condition). An ANOVA (see Figure 7) using an overall memory accuracy (proportion of hits minus false alarms) revealed a significant interaction between age and the combination of less familiar-mismatch vs. more familiar-match conditions, ($F(1,81) = 9.09, p = .003, \eta_p^2 = .10$). Follow-up t-tests suggested that there are significant differences between younger ($M = .59, SD = .22$) and older adults ($M = .43, SD = .24$) in the condition with high item

familiarity and low recall-to-reject (more familiar-match condition; $t(81) = 3.09, p = .003$). However, when older adults could use recall-to-reject and when items were made less familiar (more familiar-mismatch) age-related differences were eliminated in overall memory performance ($t(81) = .69, p = .49; M = .70, SD = .19$, for younger adults, and $M = .67, SD = .22$ for older adults).

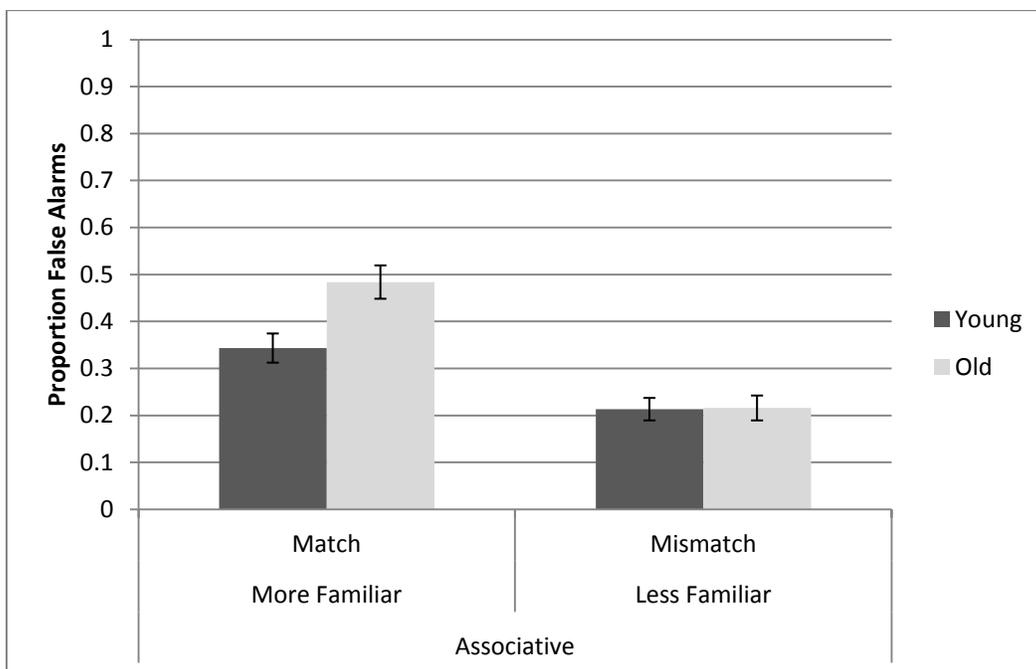
Figure 7. Mean Proportion Hits minus Proportion False Alarms for the Two Extreme Conditions in Experiment 1 with Error Bars Indicating Standard Errors



When conducting a similar ANOVA on false alarms (see Figure 8), a significant interaction was found between age and the combination of less familiar-mismatch vs. more familiar-match conditions, ($F(1,81) = 10.99, p = .001, \eta_p^2 = .12$). Again, additional analyses suggest that there were significant differences between younger ($M = .34, SD = .20$) and older adults ($M = .48, SD = .22$) when older adults could not use recollection (low recall-to-reject) and relied more so on familiarity processes (high familiarity; $t(81) =$

3.07, $p = .003$). Yet, in the less familiar-mismatch condition, in which older adults could rely more on recollection and less on item familiarity, younger ($M = .21$, $SD = .16$) and older adults ($M = .22$, $SD = .17$) did not show any significant differences ($t(81) = .39$, $p = .43$).

Figure 8. Mean Proportion False Alarms for the Two Extreme Conditions in Experiment 1 with Error Bars Indicating Standard Errors



Discussion

Two main analyses were conducted. One assessed the effects of item familiarity on item and associative recognition memory performance and the other examined the effects of schematic relationships, match vs. mismatch, and item familiarity on associative recognition memory performance. ANOVAs were conducted to analyze both the overall memory accuracy results (proportion of hits minus proportion of false alarms)

as well as the separate hits and false alarms measures, in order to determine whether it was the hits or false alarms driving memory performance.

The results of the analysis to measure overall memory performance (proportion of hits minus proportion of false alarms) revealed main effects of age, test, and item familiarity. Although there was a test by item familiarity interaction, there was no significant interaction of either of these variables with age. This would suggest that aging affects both the item and the associative tests equivalently (no age-related associative deficit) and that younger and older adults are affected by item familiarity equally. The lack of the age-related associative deficit could be caused by the short lists (12 product-price pairs) and the repetition (more familiar items repeated 4 times and less familiar items repeated twice). Both of these factors make the memory task easier for older adults; however, when analyzing the age and test interaction in the standard condition used in previous research (match condition only), older adults do display an associative deficit in overall memory performance, as expected. When analyzing the hits and false alarms separately, there were no age-related differences in the hit rate between the item and the associative test. However, there was an age and test interaction for the false alarms, with older adults showing higher false alarm rates in the associative test. This suggests that it is the false alarms that are driving the associative deficit in overall memory performance in this experiment.

The results of the second analysis assessing accuracy in the associative test revealed main effects of age (with older adults demonstrating poorer associative memory than younger ones), and match-mismatch effect, with participants demonstrating lower performance in the match than the mismatch condition; the effects of item familiarity

were not significant. The only significant interaction was of age by match-mismatch, indicating that older adults can use schematic support to overcome associative memory deficits. Specifically, older adults are taking a greater advantage of the change of match to mismatch relative to younger adults in improving their performance in the associative test. However, we did not find a significant item familiarity by age interaction effect. It is possible that our item familiarity manipulation was not strong enough, as shown by the lack of main effect of familiarity. In Experiment 2 we tried to strengthen this manipulation.

The results of the analysis that looked at the associative test only suggest that older adults have higher false alarm rates but not lower hits. Most importantly, older adults are able to use schematic support in the price category mismatch condition to lower their false alarm errors relative to the baseline match condition, as predicted. One potential mechanism by which such an improvement is attained is a recall-to-reject one. Whereas in the match condition, older adults are viewing a recombined product-price pair at test in which the pair is in the same schematic category as that at study (market value, overestimation, or underestimation), those product-price pairings in the mismatch condition change the schematic category between study and test (e.g., from underestimation to market value). Older adults are able to take advantage of this change of schematic relationship (mismatch condition) to use the recall-to-reject mechanism more efficiently. For example, by remembering at the test that the original product was priced at a low value at study, hence it couldn't have appeared with the current market value price. We also predicted that high item familiarity would increase the associative deficit. However, we are not seeing an age by familiarity interaction. As mentioned

above, one reason for this could be related to the specifics of the design of the experiment. This issue will be addressed in Experiment 2.

Finally, Castel (2005) found no age-related differences in memory performance for market value prices, however when a product was paired with either an under- or overestimated prices, younger adults displayed better recall for these prices. Our results are in line with these findings (see Appendix B), suggesting that older adults can use schematic support to help lower their associative memory deficit.

EXPERIMENT 2

The purpose of Experiment 2 was to assess whether a similar pattern of results will emerge with a different stimuli type, and to strengthen the results of Experiment 1 by using a somewhat different method. Because the length of each of the 10 lists of Experiment 1 is short (12 product-price study pairs), there is a potential ceiling effect found in the product item memory performance. In addition, the pre-exposure phase in Experiment 1 was shown immediately before each list. As a result, even though participants were instructed to focus just on the study phase, there is a possibility for confusion, as participants could figure out during the pre-exposure phase that they are being tested on those items. In addition, since there were many stimuli used in Experiment 1, only whole number values were used for prices, which could have been repeated between lists particularly for lower-priced items. This could have lead to interference effects. Finally, although we have counterbalanced all aspects across the experiment, this was not the case for each list. For example, there were not an equal number of associative pairs designated for the intact, recombined match, and recombined mismatch for each list.

In Experiment 2, we corrected for the above potential issues related to Experiment 1 and also extended the results to different types of stimuli and a different variant of the manipulation that potentially helps in recall-to-reject processes. We tested younger and older adults' item and associative memory using face-name pairs. Previous research has indicated that older adults show an associative memory deficit when using ecologically valid face-name pairs (Naveh-Benjamin, Guez, Kilb, & Reedy, 2004; Naveh-Benjamin, Shing, Kilb, Werkle-Bergner, Lindenberger, Li, 2009). We manipulated item familiarity (more familiar vs. less familiar) and change of schematic relationship (age match vs. age mismatch). As previously predicted, we hypothesize that repetition should increase memory performance for younger and older adults' on the item test, but reduce memory performance on the associative test for older adults. For this experiment, the manipulation to affect recall-to-reject processes involves a match vs. mismatch between the original age of the face that appears with a given name at study and at test. Previous research has suggested that people can be reasonably accurate when estimating the age of a face (see Rhodes, 2009 for a review). For the purpose of this study, an age match is defined as a recombined face-name pair in which the age of the person at test is similar to the age of the person shown at study (young-young or old-old), whereas an age-mismatch is a recombined face-name pair at test in which the age of the person at test is different than that at study (young-old or old-young; see Method section for details). The mismatch manipulation should aid in the use of recall-to-reject strategy. For instance, if a participant sees an older face-name pair at study and then at test the participant sees that name with a younger face, the participant can use recall-to-reject at retrieval to recollect that the name was originally paired with an older face and not with a younger face even if

the participant does not remember perceptual details of that older face. In the match condition, when an older face-name pair is recombined with a different older face it is harder for older adults to use recall-to-reject because the faces belong to the same schematic age category at retrieval. Again, we predicted that the best memory performance would be shown when items have low familiarity and there is a potential for recall-to-reject (in the mismatch conditions).

Method

Participants and Design

The experiment was conducted with a total of 43 younger adults (16 males and 27 females) with a mean age of 19.40 ($SD = 2.34$) and a mean education level of 12.72 years ($SD = 1.22$). This experiment also included 40 older adults (14 males and 26 females) with a mean age of 73.30 ($SD = 6.12$), with a mean education level of 14.49 years ($SD = 2.16$). Older adults had significantly higher level of education ($M = 14.49$, $SD = 2.16$; $t(60.61) = 4.54$, $p < .001$) than younger adults ($M = 12.72$, $SD = 1.22$). The younger adults and older adults were recruited and compensated in a similar manner as Experiment 1. The same non-complete factorial design, utilized in Experiment 1, was used for this experiment employing faces and names instead of products and prices.

Materials

This experiment was conducted with the E-Prime program software. Stimuli were faces (presented in color) and names (including a first and a last name). When presented in pairs during the study phase, the face was presented on the center of the screen with the name presented underneath it (see Figure 9a). During the training phase, each face and name were presented separately on the center of the 15 inch computer screen using

the computer's default resolution set at 640 x 480 pixels. Faces were 230 x 280 pixels. Names were presented in size 18-point Courier New font. Originating at the Max Planck Institute for Human Development, the faces adapted for this experiment were part of the FACES database (Ebner, Riediger, & Lindenberger, 2010) and Minear and Park (2004). There were an equal number of male and female as well as younger- and older-aged faces. The names were taken from MRC Psycholinguistic Database (Wilson, 1988), and Facebook.

Figure 9a. Example of Stimuli during Training Phase for Experiment 2

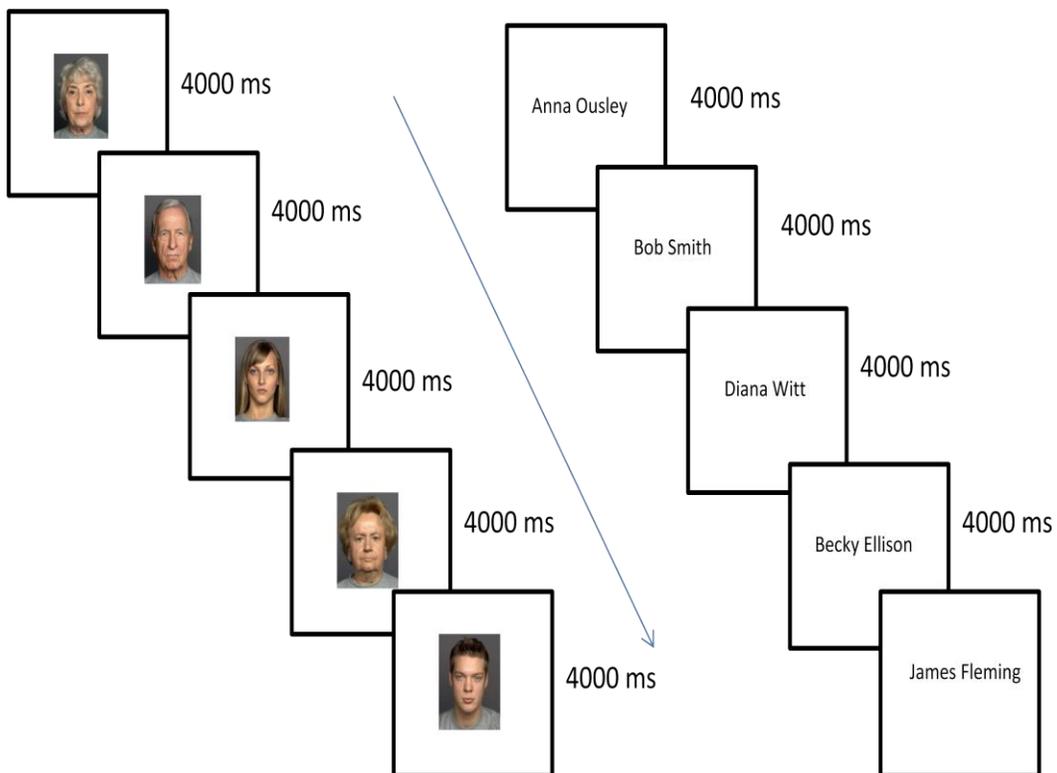
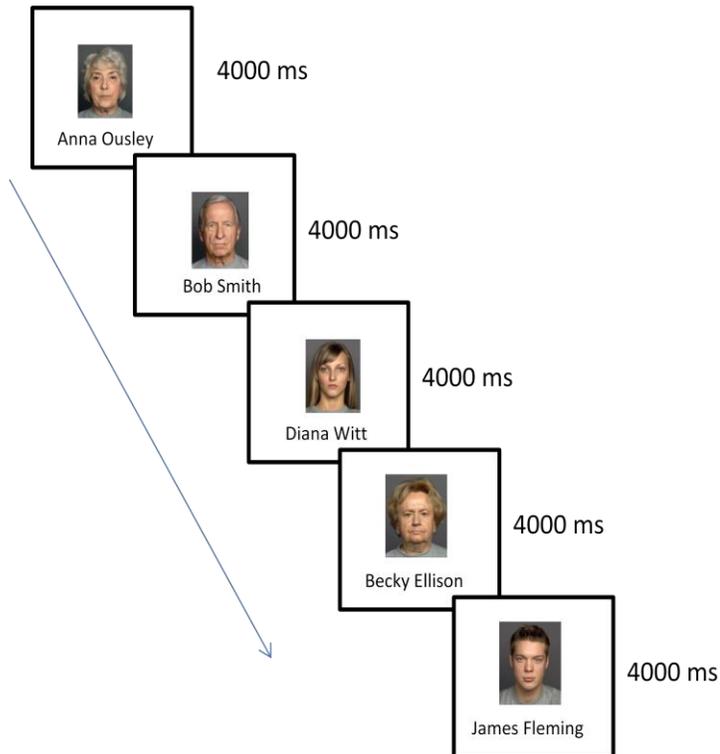


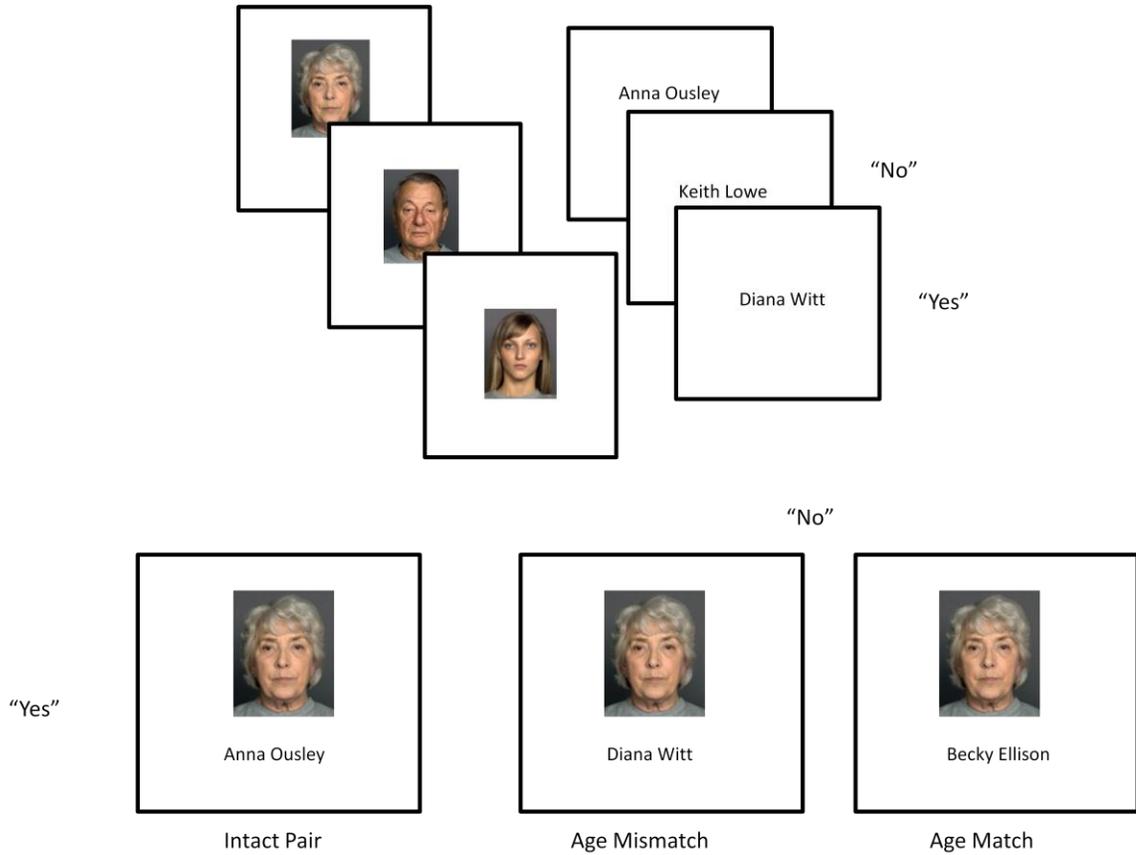
Figure 9b. Example of Stimuli during Study Phase for Experiment 2



In order to manipulate item familiarity, there were four pre-exposure or training blocks. Each block consisted of 24 faces and 24 names repeated twice in a blocked spaced manner (see Figure 9a). For the experimental phase, four study-test blocks were run. Each study phase consisted of 48 face-name pairs; in order to increase the level of performance each study pair was repeated twice in a spaced manner (see Figure 9b). Half of the studied pairs were also previously shown in the pre-exposure phase; therefore more familiar item-components were shown a total of four times and less familiar item-components were shown a total of two times. Following the study phase, four separate item recognition tests were given two for faces, each consisting of 16 items (eight from the study phase and eight new) and two for names, also with each consisting of 16 items (eight from the study phase and eight new). The items from the study phase for one of

the face recognition tests and for one of the name recognition tests were more familiar (appeared both in the training and in the study phases), while in the other two item recognition tests previously studied items were less familiar (appeared only in the study phase). In addition, two associative recognition tests were administered, each consisting of 16 pairs (eight from the study phase—intact and eight recombined). Half of the recombined pairs were age-match and the other half were age-mismatch. Unlike Experiment 1 which had three types of price-estimate categories (market value, underestimate, and overestimate), Experiment 2 only had two types of aged categories (younger faces and older faces). As mentioned earlier, an age-match pair is a younger or older adult face-name pair being recombined with a different younger or older adult face-name pair within the same age category of the face. An age-mismatch pair is a younger or older adult face-name pair being recombined with a different older or younger face-name pair between face age-categories (see Figure 10). Similar to the design of the item test, the studied items for one of the associative test were more familiar, while the studied items for the other test were less familiar.

Figure 10. Example of Stimuli for Item (Top) and Associative Tests (Bottom) for Experiment 2. The bottom diagram shows an example of an intact pair, an age-mismatch recombined pair, and an age-match recombined pair (left to right).



During the pre-exposure phase, faces and names were presented for 4 s. A 30 s break was given in between each block. The presentation of each face-name pair was also for 4 s. An interpolated activity between the study and test phase consisted of counting backwards by three's from the digit given by the experimenter. Familiarity, test order, and list order were counterbalanced between subjects.

Procedure

After getting participants' informed consent, the pre-exposure phase was presented to the participants. They were instructed to remember the individual faces and

names. They were not told that they are going to be explicitly tested on them. After the pre-exposure phase, the instructions for the main experiment were given. Participants were told that they will have four blocks of face-name pairs and that they should study the faces, names, and their pairings equally in order to prepare for the six later tests, in which they will have to decide if the stimuli was a studied item (target/intact pair) or a new item (distractor/recombined pair). They were told that the items and pairs that they were being tested on are those from the study phase portion of the experiment and not from the pre-exposure phase. Following the instructions, a practice block was given so that the participants could become acquainted with the study-test procedure. After viewing the study phase for approximately seven minutes, the interpolated activity was given. Participants were then asked to complete the six recognition tests using the same procedure as Experiment 1. For the item recognition test, either a face or a name was presented centrally on the screen in a blocked manner. Upon completion of the experiment, a post-test questionnaire and a debriefing were given.

Results

Overall memory accuracy was calculated by subtracting the proportion of false alarms from the proportion of hits. The means and standard deviations are presented in Table A4 and in graphical form with the standard errors in Figure 11.

To assess the effects of item familiarity, a 2 (age: young vs. old) X 2 (test: item vs. associative) X 2 (item familiarity: more familiar vs. less familiar) ANOVA was conducted. As expected, the analysis indicated a main effect of test was significant ($F(1, 81) = 175.70, p < .001, \eta_p^2 = .68$) with higher memory performance for the item test ($M = .63, SD = .14$) than the associative test ($M = .42, SD = .22$). There was also a

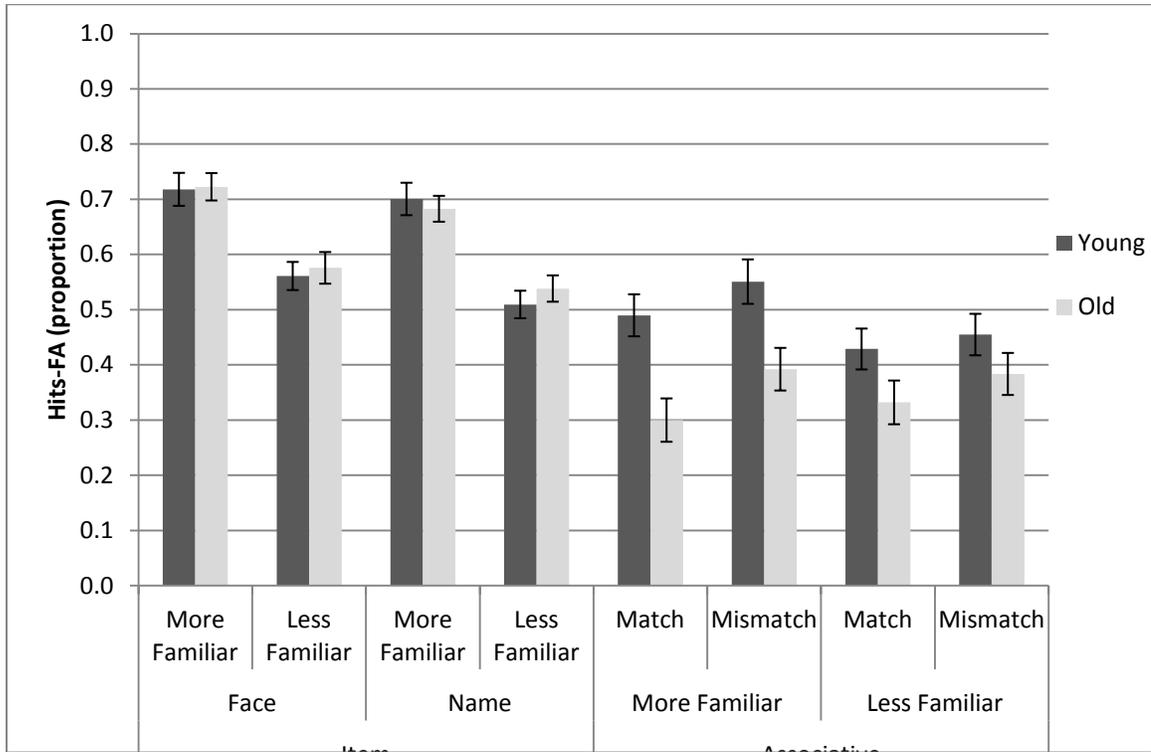
significant main effect of familiarity ($F(1, 81) = 78.93, p < .001, \eta_p^2 = .49$) with more familiar items ($M = .57, SD = .19$) being better remembered than less familiar items ($M = .47, SD = .17$). Importantly, there was a significant interaction between age and test ($F(1, 81) = 18.70, p < .001, \eta_p^2 = .19$). Follow-up t-tests indicated that younger adults ($M = .48, SD = .22$) performed better than older adults ($M = .35, SD = .21$) in the associative test ($t(81) = 2.71, p = .01$), but not in the item test ($t(81) = .24, p = .81, ns; M = .62, SD = .15$ and $M = .63, SD = .13$, for younger and older adults, respectively). This key interaction is indicative of the associative deficit. The interaction of age and familiarity was also significant ($F(1, 81) = 7.43, p = .01, \eta_p^2 = .08$). Follow-up t-tests indicated that the interaction was driven by age differences for more familiar items ($t(81) = 2.24, p = .03$) with younger adults ($M = .61, SD = .20$) performing better than older ($M = .52, SD = .17$); however, there were no significant differences between younger ($M = .48, SD = .17$) and older adults ($M = .46, SD = .17$) in the less familiar condition ($t(81) = .85, p = .40$). Finally, there was also a significant interaction of test and familiarity ($F(1, 81) = 34.36, p < .001, \eta_p^2 = .30$). Follow-up t-tests suggested that the interaction was driven by larger familiarity differences in the item test ($t(82) = 14.24, p < .001$) than in the associative test ($t(82) = 1.84, p = .07$). No other main effects or interactions were significant.

Experiment 2 reveals an age-related difference in associative memory. However, we still conducted a follow-up ANOVA to test the standard associative deficit, as done in Experiment 1 (looking only at the match condition). As expected, the results of this analysis indicated a significant age and test interaction, ($F(1, 81) = 6.47, p < .01, \eta_p^2 = .21$). Follow-up t-tests confirmed that this interaction is driven by age-related differences in the associative test, with younger adults ($M = .46, SD = .23$) performing better than

older adults ($M = .32$, $SD = .22$; $t(81) = 2.88$, $p = .005$) but not so in the item test ($t(81) = .242$, $p = .81$, ns; $M = .62$, $SD = .15$ and $M = .63$, $SD = .13$, for younger and older adults, respectively).

As done in Experiment 1, a second analysis looked at the associative test only and investigated the effects of change of schematic relationships and that of item familiarity. The design was a 2 (age: young vs. old) X 2 (age of face schematic relationship in study and test: match vs. mismatch) X 2 (item familiarity: more familiar vs. less familiar). The ANOVA conducted showed a main effect of age ($F(1, 81) = 7.37$, $p = .01$, $\eta_p^2 = .08$), with younger adults ($M = .48$, $SD = .22$) having better memory performance than older adults ($M = .35$, $SD = .21$). There was also a main effect of the match-mismatch manipulation ($F(1, 81) = 16.92$, $p < .001$, $\eta_p^2 = .17$). Memory performance was better in the age mismatch ($M = .45$, $SD = .23$) than in the age match condition ($M = .39$, $SD = .24$). There was also a marginal main effect of familiarity ($F(1, 81) = 3.25$, $p = .08$, $\eta_p^2 = .04$), with memory performance being better when items were more familiar ($M = .44$, $SD = .25$) than less familiar ($M = .40$, $SD = .23$). Finally, there was a significant two-way interaction of age and familiarity ($F(1,81) = 5.93$, $p = .02$, $\eta_p^2 = .07$). A follow-up paired sample t-test indicated that the interaction was driven by differences in familiarity for younger adults ($t(42) = 3.54$, $p = .001$) with more familiar items ($M = .52$, $SD = .24$) benefiting memory performance relative to less familiar ones ($M = .44$, $SD = .23$). A t-test that examined familiarity in older adults was not significant ($t(39) = .39$, $p = .70$; $M = .35$, $SD = .23$ for more familiar and $M = .36$, $SD = .23$ for less familiar). No other interactions were significant.

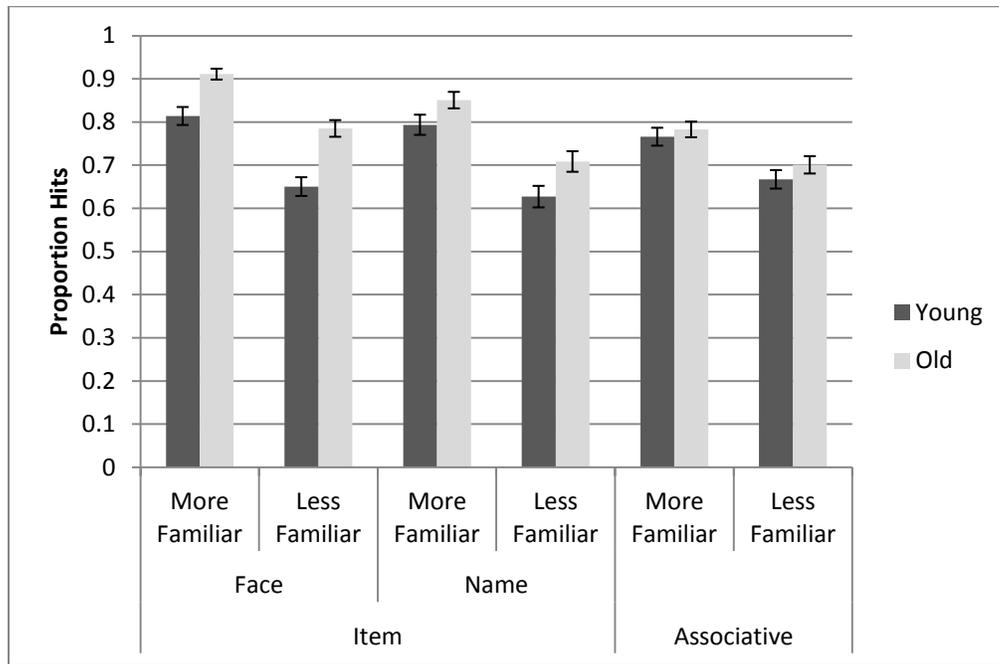
Figure 11. Mean Proportion Hits minus Proportion False Alarms in Experiment 2 with Error Bars Indicating Standard Errors



In order to determine whether it was the hits or the false alarms driving the above-presented results, similar ANOVAs to measure the effects of item familiarity and the change of schematic relationships, were run separately for the hits (Table A5 and Figure 12) and for the false alarms (Table A6 and Figure 13) measures. The analysis of hits revealed a main effect of age ($F(1, 81) = 7.02, p = .001, \eta_p^2 = .08$), with older adults ($M = .78, SD = .08$) having a higher hit rate than younger adults ($M = .72, SD = .11$). A main effect of test was found ($F(1, 81) = 13.87, p = .001, \eta_p^2 = .15$), with a higher hit rate in the item test ($M = .77, SD = .11$) than the associative test ($M = .73, SD = .11$). There was a main effect of familiarity ($F(1, 81) = 221.82, p < .001, \eta_p^2 = .73$), with higher proportion of hits when items were more familiar ($M = .84, SD = .12$) than when items were less

familiar ($M = .69$, $SD = .13$). There was a significant age by test interaction ($F(1, 81) = 8.83$, $p = .004$, $\eta_p^2 = .10$). Follow-up t-tests revealed that this interaction is driven by age-related differences in the item test hit rate ($t(81) = 4.03$, $p < .001$; $M = .81$, $SD = .09$ for older adults and $M = .72$, $SD = .12$ for younger adults), and not by differences in the associative test ($t(81) = .98$, $p = .33$; $M = .74$, $SD = .10$ for older adults and $M = .72$, $SD = .13$ for younger adults). Moreover, this interaction is not indicative of the associative deficit because age-related differences are in the item test not the associative test, suggesting that hits are not driving the associative deficit. There was also a significant test by familiarity interaction ($F(1, 81) = 14.74$, $p < .001$, $\eta_p^2 = .15$). Follow-up t-tests suggested that there were familiarity differences in both the item ($t(82) = 14.56$, $p < .001$) and the associative tests ($t(82) = 7.61$, $p < .001$), with greater differences in the item ($M = .84$, $SD = .12$ and $M = .69$, $SD = .13$, for more and less familiar, respectively) than in the associative test, ($M = .77$, $SD = .13$, and $M = .68$, $SD = .13$, for more and less familiar, respectively). No other interactions were significant.

Figure 12. Mean Proportion Hits in Experiment 2 with Error Bars Indicating Standard Errors



The analysis of false alarms using, test, age, and item familiarity revealed that older adults are making more false alarms ($M = .29$, $SD = .14$) than younger adults ($M = .17$, $SD = .11$; $F(1, 81) = 19.06$, $p < .001$, $\eta_p^2 = .19$). There was also a main effect of test ($F(1, 81) = 202.47$, $p < .001$, $\eta_p^2 = .71$), indicating a higher proportion of false alarms in the associative test ($M = .31$, $SD = .18$) than the item test ($M = .14$, $SD = .11$). In addition, there was a main effect of familiarity ($F(1, 81) = 9.19$, $p = .003$, $\eta_p^2 = .10$), with more false alarms being made when items were more familiar ($M = .24$, $SD = .15$), than when items were less familiar ($M = .21$, $SD = .14$). The key two-way interaction of age and test ($F(1, 81) = 8.33$, $p = .01$, $\eta_p^2 = .09$) was significant. Follow-up t-tests revealed that there were age-related differences in the item test ($t(77.64) = 3.69$, $p < .001$) and even larger ones in the associative test ($t(81) = 4.34$, $p < .001$), with older adults making

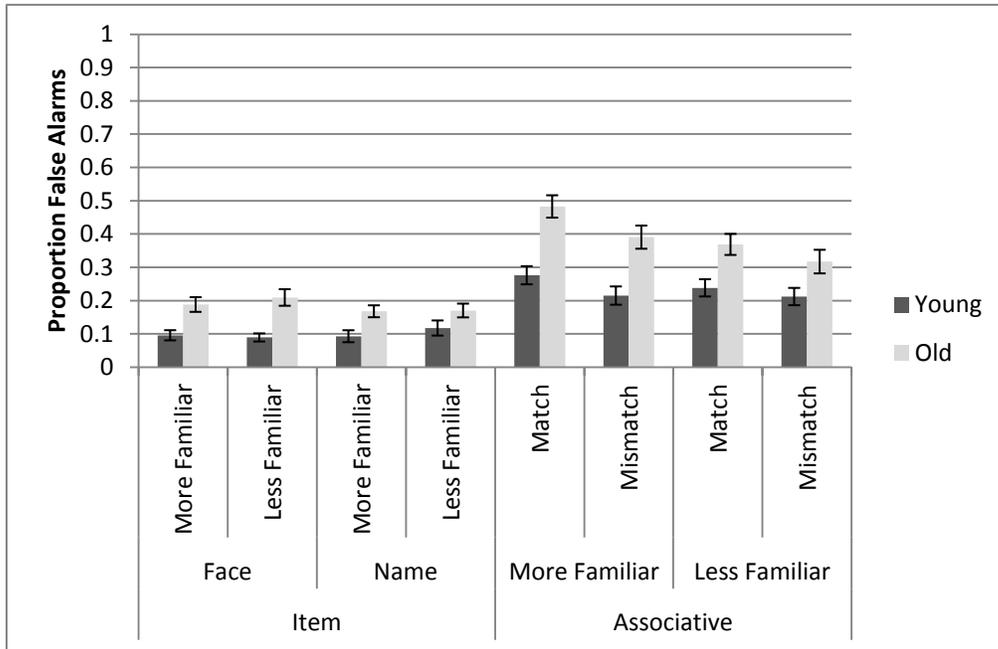
more false alarms in item ($M = .18, SD = .11$) and associative ($M = .39, SD = .18$) tests relative to younger adults and especially so in the associative test ($M = .24, SD = .14; M = .10, SD = .10$ in the item test). To compare Experiment 1 to Experiment 2, we also analyzed the false alarms using the standard conditions analysis (for the match condition only) and found a significant age and test interaction ($F(1,81) = 8.97, p = .004, \eta_p^2 = .10$). Follow-up analyses confirmed that older adults ($M = .43, SD = .19$) were making more false alarms than younger adults ($M = .26, SD = .16$) in the associative test ($t(81) = 4.42, p < .001$) and the item test ($t(81) = 3.69, p < .001; M = .18, SD = .11$ for older adults and $M = .10, SD = .10$ for younger adults). In addition, a significant interaction of age and familiarity ($F(1, 81) = 5.29, p = .02, \eta_p^2 = .06$) was found. Further tests indicate that older adults made significantly more false alarms in the more familiar condition ($M = .31, SD = .14$) than in the less familiar condition ($M = .27, SD = .15; t(39) = 3.624, p < .001$); whereas there were no significant differences between familiarity conditions for younger adults ($M = .17, SD = .12$ and $M = .16, SD = .11$ for more and less familiar, respectively; $t(42) = .538, p = .59$). Test and item familiarity also significantly interacted with each other ($F(1, 81) = 20.67, p < .001, \eta_p^2 = .20$). Follow-up t-tests revealed item familiarity differences in the associative test ($t(82) = 3.90, p < .001$), and not the item test ($t(82) = 1.70, p = .10; M = .13, SD = .11$, and $M = .15, SD = .12$, for more familiar and less familiar, respectively), with more false alarms made in the more familiar condition ($M = .34, SD = .20$) than the less familiar condition ($M = .28, SD = .18$) in the associative test. This age by familiarity interaction can also be interpreted in the context of the significant three-way interaction of age, test, and familiarity, ($F(1, 81) = 6.57, p = .02, \eta_p^2 = .08$). Follow-up ANOVAs for each test, showed no age by familiarity interaction in the item

test ($F(1, 81) = .05, p = .83, \eta_p^2 = .001$). However, the age by familiarity interaction was significant in the associative test ($F(1,81) = 7.06, p = .01, \eta_p^2 = .08$), indicating significant familiarity effects in older adults ($t(39) = 4.16, p < .001$, reflecting higher false alarms in the more familiar than the less familiar condition; $M = .44, SD = .19$ for more familiar items, and $M = .34, SD = .20$ for less familiar items), with a lack of familiarity effects in younger adults ($t(42) = 1.24, p = .22, M = .25, SD = .16$ and $M = .23, SD = .14$, for more and less familiar conditions, respectively).

Another ANOVA was conducted for the associative test only and analyzed the effects of the change of the schematic relationship and item familiarity for false alarms. A main effect of age was found ($F(1,81) = 18.81, p < .001, \eta_p^2 = .19$), with older adults making more false alarms on the associative test ($M = .39, SD = .18$) than younger adults ($M = .24, SD = .14$). There was a main effect of change of schematic relationship ($F(1, 81) = 16.92, p < .001, \eta_p^2 = .17$), with a higher proportion of false alarms in the match condition ($M = .34, SD = .19$), than in the mismatch condition ($M = .28, SD = .19$). There was also a main effect of familiarity ($F(1, 81) = 17.06, p < .001, \eta_p^2 = .17$), with a higher proportion of false alarms being made when items were more familiar ($M = .34, SD = .20$) than when they were less familiar ($M = .28, SD = .18$). Importantly, there was a significant two-way interaction of age and familiarity ($F(1, 81) = 7.06, p = .01, \eta_p^2 = .08$). A follow-up paired sample t-test indicated that the interaction was driven by differences in familiarity for older adults ($t(39) = 4.16, p < .001$) with older adults making more false alarms when items were more familiar ($M = .44, SD = .19$) than less familiar ($M = .34, SD = .20$). A t-test that examined the effects of familiarity on younger adults was not significant ($t(42) = 1.24, p = .22; M = .25, SD = .16$, and $M = .23, SD =$

.14, for more familiar and less familiar, respectively). No other interactions were significant.

Figure 13. Mean Proportion False Alarms in Experiment 2 Error Bars Indicating Standard Errors



Discussion

First, in this experiment, older adults have shown an associative deficit, as their performance in the associative test relative to the item test was poorer than that of younger adults. This deficit was mostly due to the higher false alarm rate older adults exhibited in the associative test. Furthermore, we hypothesized that the combination of low item familiarity and high recall-to-reject would result in the highest memory performance for older adults and the lowest associative deficit. The overall results of Experiment 2, as those of Experiment 1, support this hypothesis. Looking specifically at

the results for each of the different manipulations, Experiment 2 showed the hypothesized effect of the match-mismatch manipulation, with better memory performance in the mismatch condition. Interestingly, in this experiment, both older and younger adults improved their memory performance in the mismatch relative to the match condition, and this was due to a lower false alarm rate in the associative test in the mismatch condition. Furthermore, as hypothesized, Experiment 2 did show a significant age by familiarity interaction for overall memory performance, reflecting that older adults show a lower false alarm rate when there was a decrease in item familiarity.

One key difference between the two experiments is in how recall-to-reject was defined. In Experiment 1, the change of schematic relationship was between the two item-components, the product and the price. In order to perform well, the participant had to remember whether a given product was paired with either an underestimated, market value, or overestimated price and how that had changed from study to test. In this case, the price had some ecologically valid meaning when paired with a given product. In Experiment 2, the change of schematic support (age of face) was only in the individual face component, as the name that was paired with each face did not really have any schematic meaning. As it turned out, despite the differences in the operational definition of the match-mismatch manipulation in both experiments, the results showed in both cases that older adults can improve their associative memory performance in the mismatch condition. A follow-up analysis was conducted in order to equate the two experiments with respect to the meaning of the match-mismatch manipulation. Post-test questionnaires completed by the participants in Experiment 2 indicated that one strategy used by many of them involved the assessment of whether the person's age as reflected

by the face fit the name that appeared with it. Specifically, many of the participants reported that some of the names sounded like an older person's name (i.e. Velma Berry) while others sounded like a younger person's name (i.e. Conner Mullins). To assess the potential effect of the age of the name, a name rating scale was given to an independent group of younger ($n = 9$) and older ($n = 10$) adults. Participants were asked to rate a name on a scale of 1 to 10, with 1 being "definitely sounds like a younger person's name" and 10 being "definitely sounds like an older person's name." Forty of the more younger-sounding names ($M = 3.95$, $SD = .52$) and 40 of the more older-sounding names ($M = 7.40$, $SD = .68$) were used for further analysis. In order to equate Experiment 2 with Experiment 1, we analyzed how the younger-sounding names and the older-sounding names affected the already defined age-match or age-mismatch (see Figure 14a for more details). This analysis equates Experiment 2 to Experiment 1 by having the match-mismatch manipulation based on an association between the two item-components (younger- or older-sounding name with the age of the face) rather than being a manipulation that includes only an individual item-component (schematic change of age of face).

A three-way ANOVA was conducted with age, test, and item familiarity as the independent variables and overall memory accuracy (proportion of hits minus false alarms) as the dependent variable (see Figure 14a). It is important to note that this time the item test just consisted of the name test. There was a significant main effect of age ($F(1, 81) = 4.15$, $p = .045$, $\eta_p^2 = .05$), with younger adults ($M = .54$, $SD = .18$) performing better than older adults ($M = .46$, $SD = .17$). There was a significant main effect of test ($F(1, 81) = 29.13$, $p < .001$, $\eta_p^2 = .27$), with higher memory performance for

the item test ($M = .60, SD = .17$) than the associative test ($M = .45, SD = .24$). Finally there was a main effect of familiarity ($F(1, 81) = 17.50, p < .001, \eta_p^2 = .18$) with more familiar items ($M = .57, SD = .20$) being better remembered than less familiar items ($M = .48, SD = .17$). The only significant interaction was of test and familiarity ($F(1, 81) = 13.37, p < .01, \eta_p^2 = .14$) indicating that there was no difference in the associative test between more familiar ($M = .45, SD = .30$) and less familiar pairs ($M = .44, SD = .27; t(82) = .34, p = .74$), and a larger difference in the item test ($t(82) = 5.68, p < .001$) between more familiar ($M = .69, SD = .22$) and less familiar items ($M = .52, SD = .20$). This interaction, although not a key interaction with age, was significant in all three analyses indicating that item familiarity greatly improves performance on the item test.

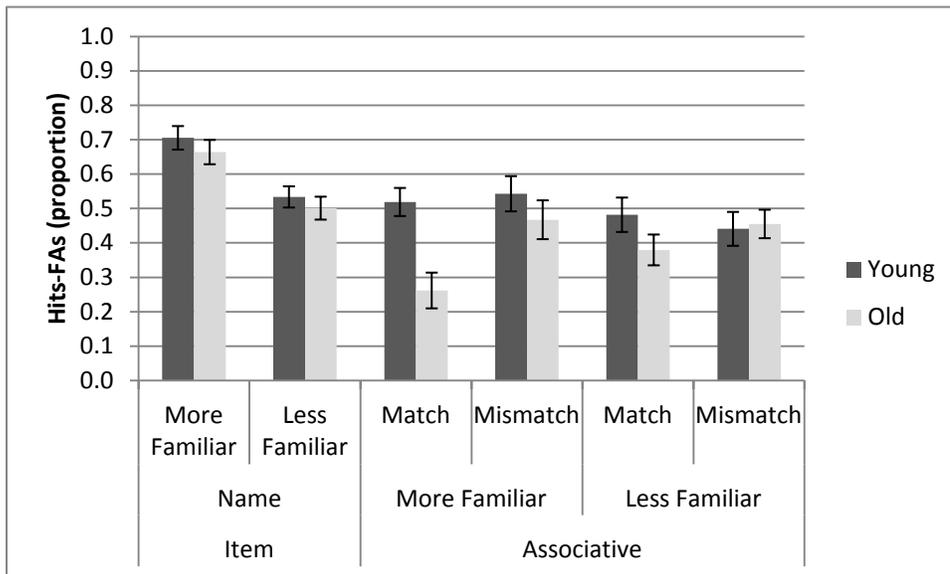
We ran a follow-up ANOVA using the standard condition (item test and match condition only). This analysis did reveal a significant interaction of age and test ($F(1, 81) = 5.53, p = .02, \eta_p^2 = .06$). Follow-up t-tests indicated that this interaction was driven by age-related differences ($M = .50, SD = .25$ and $M = .32, SD = .27$, for younger and older adults, respectively) in the associative test ($t(81) = 3.16, p = .002$) but not in the item test ($t(81) = 1.00, p = .32; M = .62, SD = .16$ and $M = .58, SD = .17$), younger and older adults, respectively). This result suggests that by equating the match-mismatch manipulation between the two experiments, basing the association between two item-components rather than the individual item-component, we are getting results similar to that of Experiment 1 for this analysis.

We also performed a three-way ANOVA examining the effects of age, item familiarity, and schematic relationship on the associative test. There was main effect of age ($F(1,81) = 3.98, p = .049, \eta_p^2 = .05$), with younger adults ($M = .50, SD = .24$) having

better memory performance than older adults ($M = .39, SD = .24$) and a main effect of the match-mismatch manipulation ($F(1, 81) = 8.29, p = .005, \eta_p^2 = .09$), with better memory performance in the mismatch ($M = .48, SD = .26$) than in the match condition ($M = .41, SD = .27$). Again these were the only main effects found in Experiment 1. Like Experiment 1 but unlike the original results of Experiment 2, there was a significant interaction of schematic relationship and age ($F(1,81) = 10.57, p = .002, \eta_p^2 = .12$), suggesting that age differences were significant in the match condition ($t(81) = 3.16, p = .002$) but not the mismatch condition ($t(81) = .52, p = .60$). Older adults ($M = .32, SD = .27$) showed a deficit in the match condition relative to younger adults, ($M = .50, SD = .25$) but not in the mismatch condition ($M = .49, SD = .25$ and $M = .46, SD = .28$, for younger and older adults, respectively). There was also an age by familiarity interaction ($F(1, 81) = 3.96, p = .05, \eta_p^2 = .05$) which was shown in the original analysis of Experiment 2 but not in Experiment 1. When conducting follow-up t-tests on this interaction, results revealed that there were significant differences between younger ($M = .53, SD = .28$) and older adults ($M = .36, SD = .30$) in the more familiar condition ($t(81) = 2.62, p = .01$). However, there were no significant differences between younger ($M = .46, SD = .29$) and older adults ($M = .42, SD = .24$) in the less familiar condition ($t(81) = .75, p = .45$).

Figure 14a. Mean Proportion Hits minus Proportion False Alarms in Experiment 2
 Follow-up Analyses with Error Bars Indicating Standard Errors

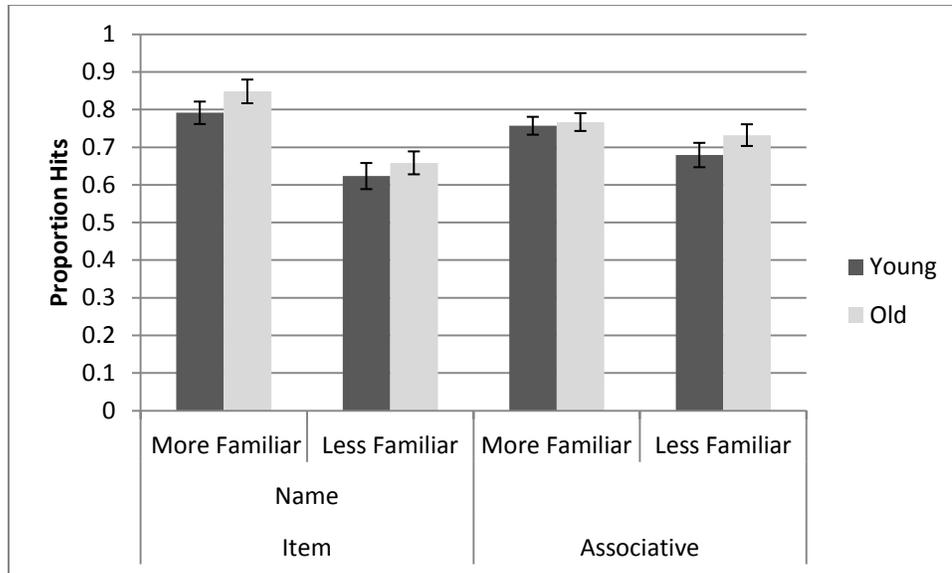
A match pair is a younger or older adult face-name pair recombined with a different younger or older adult face-name pair within the same face age category. A mismatch pair is a younger or older adult face-name pair recombined with a different older or younger face-name pair between face-age categories.



In order to determine, whether it was the hits or the false alarms driving the above-presented results, similar ANOVAs were conducted to measure item familiarity and change of schematic relationship separately for hits and false alarms (see Figures 14b and 14c, for hits and false alarms, respectively). For the analysis of hits the only main effect was familiarity ($F(1, 81) = 42.95, p < .001, \eta_p^2 = .35$), with more hits obtained when items were made more familiar through the use of the pre-exposure phase ($M = .79, SD = .13$) than when they were less familiar and shown only in the study phase ($M = .67, SD = .15$). Neither test nor familiarity interacted with age. This suggests that younger

and older adults are performing equivalently for hits on the item and associative regardless of whether items are more or less familiar.

Figure 14b. Mean Proportion Hits in Experiment 2 Follow-up Analyses with Error Bars Indicating Standard Errors



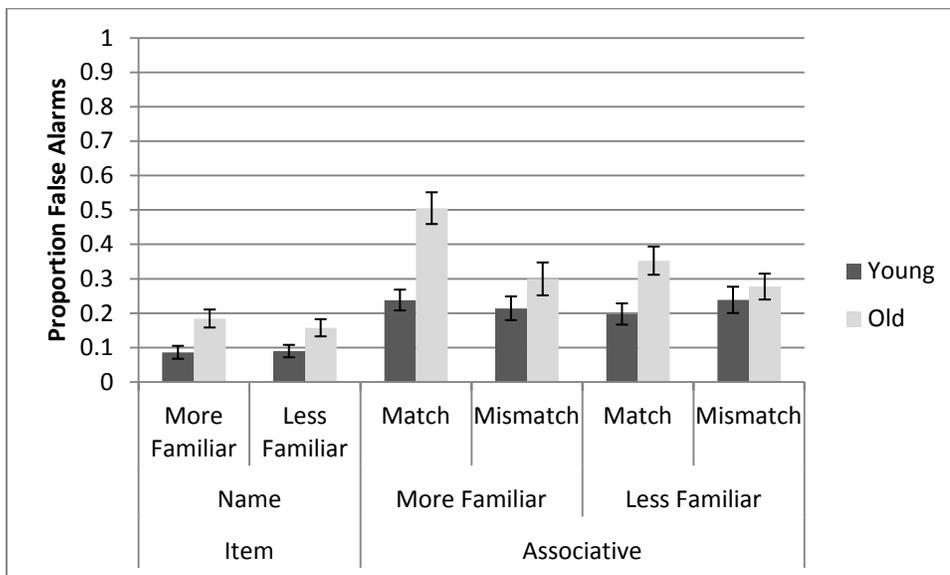
The effects of item familiarity on false alarms indicated main effects of age ($F(1, 81) = 16.35, p < .001, \eta_p^2 = .17$), test ($F(1, 81) = 64.71, p < .001, \eta_p^2 = .44$) and familiarity ($F(1, 81) = 5.40, p = .02, \eta_p^2 = .06$). Specifically, older adults ($M = .30, SD = .16$) made significantly more false alarms than younger adults ($M = .18, SD = .11$). Participants made more false alarms on the associative test ($M = .29, SD = .19$) than the item ($M = .13, SD = .13$). Finally, more false alarms were made when items were more familiar ($M = .25, SD = .17$) than less familiar ($M = .22, SD = .16$). As in the previous analysis, examining overall performance, there was no significant age and test interaction. This was also shown in Experiment 1. When conducting the follow-up

ANOVA in the standard condition only, there was a significant age and test interaction ($F(1,81) = 8.42, p = .005, \eta_p^2 = .09$). Follow-up t-tests revealed that there were age-related differences in the item test ($M = .17, SD = .14$ and $M = .09, SD = .10$, for old and young, respectively, $t(70.94) = 3.00, p = .004$) and larger ones in the associative test ($M = .43, SD = .25$ and $M = .22, SD = .16$, for old and young, respectively, $t(64.12) = 4.56, p < .001$). Like in the original analysis of Experiment 2, there was an age and familiarity interaction ($F(1,81) = 4.61, p = .04, \eta_p^2 = .05$). Follow-up tests indicated that there were no significant differences in the false alarm rate for younger adults ($t(42) = .19, p = .85$) between the more familiar ($M = .18, SD = .13$) and the less familiar conditions ($M = .18, SD = .14$). However, older adults committed significantly fewer false alarms ($t(39) = 3.56, p = .001$) when items were made less familiar ($M = .26, SD = .17$) relative to when they were made more familiar ($M = .33, SD = .17$).

The final ANOVA conducted examined the change of schematic relationship and item familiarity for false alarms in the associative test. A main effect of age was found ($F(1, 81) = 12.21, p = .001, \eta_p^2 = .13$), with older adults making more false alarms on the associative test ($M = .36, SD = .21$) than younger adults ($M = .22, SD = .13$). There was a main effect of change of schematic relationship ($F(1, 81) = 8.29, p = .005, \eta_p^2 = .09$), with a higher proportion of false alarms in the match ($M = .32, SD = .23$) than in the mismatch condition ($M = .26, SD = .21$). Finally, a main effect of familiarity was found ($F(1,81) = 5.05, p = .03, \eta_p^2 = .06$), with a higher proportion of false alarms made when items were more familiar ($M = .31, SD = .23$) than when they were less familiar ($M = .27, SD = .20$). In Experiment 1, age interacted with change of schematic relationship, but not with familiarity. In the original analysis of Experiment 2, age interacted with familiarity,

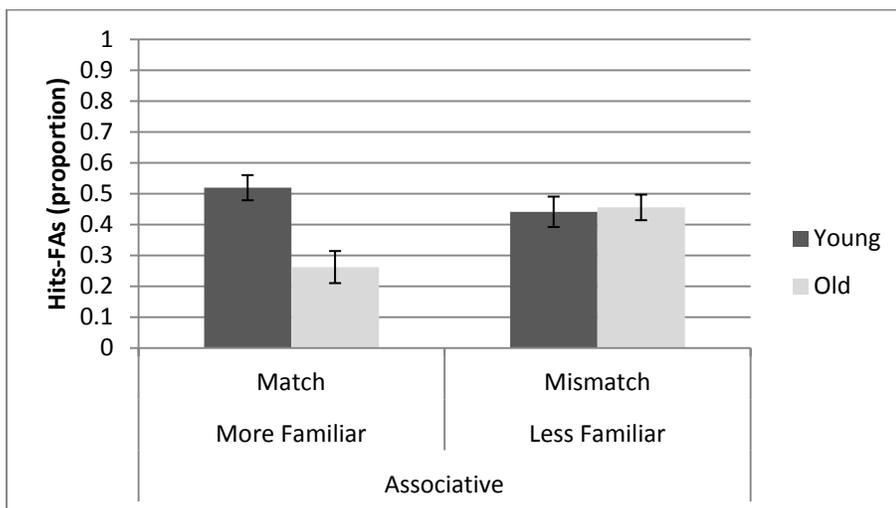
but not with change of schematic relationship. In the current analysis, age interacted with change of relationship ($F(1, 81) = 10.65, p = .002, \eta_p^2 = .12$) and marginally with familiarity ($F(1, 81) = 3.48, p = .07, \eta_p^2 = .04$). Follow-up t-tests indicate that there were significant age-related differences in the match condition ($t(64.12) = 4.56, p < .001$), with older adults ($M = .42, SD = .25$) making a higher proportion of false alarm errors than younger adults ($M = .22, SD = .16$), whereas younger ($M = .23, SD = .17$) and older adults ($M = .29, SD = .24$) performed similarly in the mismatch condition ($t(68.40) = 1.36, p = .18$). Younger adults performed similarly in the more familiar ($M = .23, SD = .17$) and less familiar ($M = .22, SD = .17$) conditions, whereas older adults made more false alarms in the associative test when items were more familiar ($M = .40, SD = .25$) than less familiar ($M = .32, SD = .21$).

Figure 14c. Mean Proportion False Alarms in Experiment 2 Follow-up Analyses with Error Bars Indicating Standard Errors



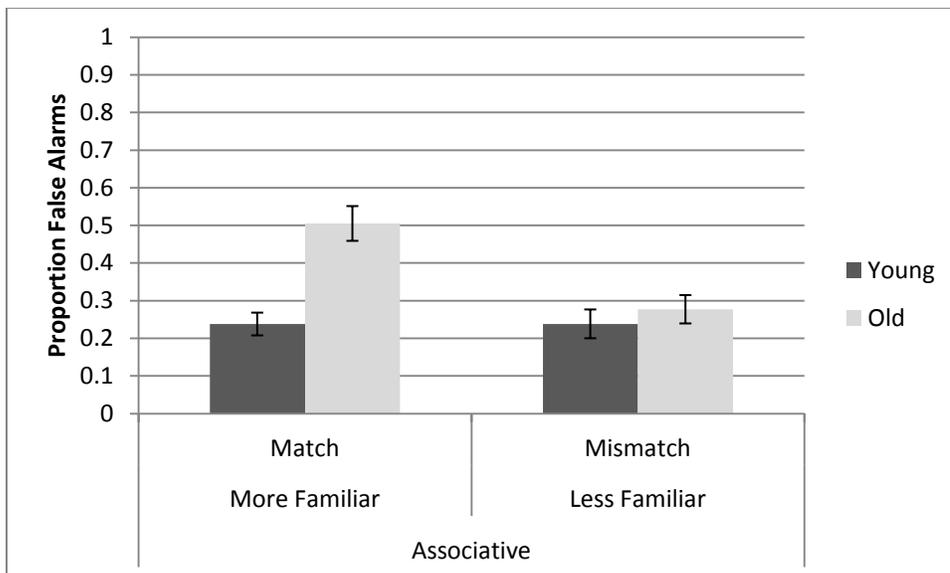
Finally, a set of statistical analyses was run on the follow-up results of Experiment 2 to analyze the contrast between the conditions in which we expected older adults to have the best associative memory performance and the worst associative memory performance, similar to the final analysis presented in Experiment 1. The results indicated a significant interaction between age and the combination of less familiar-mismatch vs. more familiar-match conditions ($F(1,81) = 12.91, p = .001, \eta_p^2 = .14$) in overall associative memory performance (proportion of hits minus false alarms; see Figure 15). Follow-up t-tests indicated that there were no significant differences between younger ($M = .44, SD = .32$) and older adults ($M = .45, SD = .26$) in the less familiar mismatch condition ($t(81) = .22, p = .83$); however there were significant differences between the two age groups in the more familiar match condition ($t(81) = 3.93, p = .000; M = .52, SD = .27$ and $M = .26, SD = .33$, for younger and older adults, respectively).

Figure 15. Mean Proportion Hits minus Proportion False Alarms in the Two Extreme Conditions for Experiment 2 Follow-up Analyses with Error Bars Indicating Standard Errors



Similarly, an ANOVA run on false alarms (see Figure 16) found that age interacted with the combination of less familiar-mismatch vs. more familiar-match conditions ($F(1, 81) = 12.38, p = .001, \eta_p^2 = .13$). Follow-up analysis indicated significant differences between younger ($M = .24, SD = .19$) and older adults ($M = .50, SD = .29$) in the high familiar-match condition ($t(68.04) = 4.84, p < .001$). There were no significant differences in the less familiar-mismatch condition ($t(81) = .72, p = .47; M = .24, SD = .25$ and $M = .28, SD = .24$, for younger and older adults, respectively).

Figure 16. Mean Proportion False Alarms in the Two Extreme Conditions for Experiment 2 Follow-up Analyses with Error Bars Indicating Standard Errors



To summarize, the results of the follow-up analyses, when equating the operational definition of the match-mismatch manipulation in Experiment 2 to the one used in Experiment 1, reveal some interesting findings. First, when conducting the analysis with the standard condition, the interaction between age and test was significant,

revealing age-related differences in the associative test but not the item test. The poor performance of the older adults in the associative test appears to be driven by their high false alarm rate as there were no age-related differences in the hit rate. Also, the results indicated age-related interactions with item familiarity and change of schematic relationship in the associative test for overall memory performance and for false alarms. Whereas in Experiment 1 we only saw an interaction of age and change of schematic relationship and in the original analysis of Experiment 2 we only saw an interaction of age and familiarity, the follow-up analysis of Experiment 2 indicated that both interactions were significant.

GENERAL DISCUSSION

According to the associative deficit hypothesis, older adults' have trouble binding item-components and later retrieving these bound components, while their ability to remember individual item-components stays relatively intact with age (Naveh-Benjamin, 2000). This is a robust finding and has been found with many different modalities and stimuli (see Old & Naveh-Benjamin, 2008a). Previous research has found that a high false alarm rate is an important factor in the associative deficit (Castel & Craik, 2003; Cohn, Emrich, & Moscovitch, 2008; Kilb & Naveh-Benjamin, 2011; Naveh-Benjamin, Shing, Kilb, Werkle-Bergner, Lindenberger & Li, 2009). High item familiarity and low-recall-to-reject processing have been implicated as potential mediators of the high false alarm rate (Cohn, Emrich, Moscovitch, 2008; Light, Chung, Pendergrass, & Van Ocker, 2006; Light, Patterson, Chung, & Healy, 2004).

The purpose of the current set of experiments was an attempt to lower the false alarm rate in older adults' associative memory and subsequently decrease the associative

memory deficit. In Experiment 1, younger and older adults were shown product-price pairings at study while item familiarity (more familiar vs. less familiar) and change of schematic relationship (match-mismatch) were manipulated. Results indicated that when performance in the standard condition used in previous research is evaluated (the match condition in the current experiments) older adults show an associative memory deficit (poorer performance than the young in the associative but not in item memory). These results seem to be mostly driven by high false alarm rates in older adults' associative memory performance.

When assessing the effects of item familiarity and match-mismatch manipulations, specifically on the associative test in Experiment 1, there were differential effects of age on change of schematic relationship (match vs. mismatch). In particular, as hypothesized, older adults took a greater advantage of the price category mismatch condition in improving their associative memory compared to the baseline match condition. This suggests that older adults are able to use schematic support and the potential use of the recall-to-reject strategy to decrease their false alarm rate, and as a result increase their associative memory accuracy (proportion of hits minus false alarms).

As for the effect of item familiarity, in Experiment 1, the item familiarity manipulation did not affect older adults' associative memory performance, as hypothesized. One reason for the lack of item familiarity effect could be the weak item familiarity manipulation (potentially due to the low performance in the price item test, the pre-exposure phase appearing before each list, or the lists being short), and in Experiment 2 we have changed several features of the experiment in order to strengthen this manipulation.

The purpose of Experiment 2 was to extend the results of Experiment 1 to different stimuli and to strengthen these results by using a modified design. In particular, when the pre-exposure phase happened before the beginning of the four experimental study-test blocks (rather than before each list, as done in Experiment 1), the results of Experiment 2 did reveal an age by item familiarity interaction which was absent in Experiment 1, indicating that whereas younger adults took advantage of an increase in item familiarity in the associative test, older adults did not. Furthermore, the results of Experiment 2 replicated the results of Experiment 1, showing an age-related associative deficit for overall memory performance, which was mostly driven by a high false alarm rate in the associative test. Interestingly, although the results of Experiment 2 did replicate those of Experiment 1 showing an improvement in older adults' associative memory (via a decline in false alarms) in the mismatch over the match condition, they also showed a similar improvement in younger adults. These results can be compared to the results McGillivray and Castel (2010) that found that when using the age of the face to aid in associative memory, age-related differences in performance were not eliminated to the same degree as when participants used product-price schematic associations at test (Castel, 2005). The improvement of associative memory in both age-groups in the mismatch over the match condition in Experiment 2, in comparison to the improvement only in older adults reported in Experiment 1 could potentially be explained by how recall-to-reject was defined in each experiment. In Experiment 1, the change of schematic relationship was defined as the change of the price category (price estimate) between the product and the price which required a stronger association between the two item-components compared to Experiment 2 in which the change of schematic category

(age) was associated with the face-component and not as much with the association (face-name). A follow-up analysis was conducted to try to equate the two experiments (see Experiment 2 Discussion for details), which found that when a stronger association is made between the two item-components when manipulating recall-to-reject, older adults take a larger advantage of it relative to younger adults. Finally, Experiment 2 showed an interaction between age and familiarity, with the follow-up analysis indicating that increasing item familiarity helped younger adults in the associative test, while it decreased older adults' performance in this test, in line with the hypothesis.

We would like to note several limitations to the reported experiments. First, the price-products pairings in Experiment 1 can be somewhat subjective, especially for larger products which was why we used picture-price and not just label-price pairings. However, it could still be somewhat difficult to judge how much a piece of technology or appliance costs without knowing any features (i.e. brand, size) of the product. Second, in Experiment 1 prices were repeated between lists which could have caused interference for the price test. Prices had to be repeated particularly for small grocery items because even rounding to the next closest dollar amount may have made that product change price category. In Experiment 2, we tried to eliminate this problem by using each of the stimuli (face and name) only once.

Another potential methodological limitation could be the assessment of item familiarity and recollection. We manipulated item familiarity by having a pre-exposure phase before each list in Experiment 1. Several of the participants realized that the items that were shown in the pre-exposure phase were also shown in the study phase, even when explicitly telling them that they would only be tested on stimuli presented in the

study phase. We tried to eliminate this problem in Experiment 2 by manipulating item familiarity within list rather than between lists, presenting participants with the pre-exposure phase for all of the lists before the beginning of the presentation of the four lists. Still, several of the participants noticed the repetition. Future research could give participants a judgment task such as judging whether the presented item was a product or a price when the items were inter-mixed in the pre-exposure phase in Experiment 1, or judging whether a face or name is of a male or a female in Experiment 2.

As mentioned before, in both experiments older adults improved their associative memory performance in the mismatch (over the match) condition, with this effect potentially mediated by an increase use of the recall-to-reject strategy. For example, all of the younger adults and all but one of the older adults in Experiment 1 noticed that product-price pairings were either market value, underestimated, or overestimated. Furthermore, approximately 84% of the younger adults and 65% of the older adults said that having the product-price category helped them in the associative test. A similar question was also given on the post-test questionnaire for Experiment 2. All of the younger and older adults noticed that the faces were either a young face or an old face. In addition, approximately 74% of younger adults and 68% of older adults said that the age of the face did help them in the associative test, when asked about it. Although not all these participants reported using the category manipulation in helping them perform well in the associative test, these results suggest that even using schematic support does seem to be related to improved memory performance. Future research could assess the effects of other manipulations intended to improve recall-to-reject processes in older adults, for example by informing participants about the recall-to-reject retrieval strategy

(e.g. Gallo, Bell, Beir, & Schacter, 2006). This can be especially helpful for the older adults in case they are not spontaneously using such a strategy (e.g., the production deficit hypothesis, Craik & Byrd, 1982; Naveh-Benjamin, Brav, & Levy, 2007). Furthermore, future research could assess the potential effects of familiarity and recollection manipulations by asking participants directly about the underlying sources of their judgments (familiarity and recollection) using a remember/know task (e.g. Kilb & Naveh-Benjamin, 2011; Rotello & Heit, 2000).

Future work could also investigate other types of retrieval monitoring, such as recall-to-accept (as described in the introduction) or the distinctiveness heuristic. Whereas recall-to-reject is a type of retrieval processing in which participants reject a recombined pair by successfully remembering the original pair presented at study, the distinctiveness heuristic involves recollecting that you did not see the lure at study, particularly a pictorial nonstudied item, because pictures are more distinctive. In addition, Fandakova, Shing, and Lindenberger (2013a) suggest that not only is associative memory impaired by aging but also memory monitoring, which could also affect the high false alarm rate is involved. For example, older adults made significantly more false alarms for lure pairs in a continuous recognition task in which participants had to discriminate if that pair was already presented on that particular run (repeated) or was seen in another run (lure). Fandakova et al.'s results show that false alarms for lures increased across runs for all age groups, but were particularly pronounced in older adults. This could be due to the high monitoring demands of this task, in which older adults are severely impaired relative to children and young adults. It should be easier to correctly reject a lure in the first run compared to later runs due to less interference caused by

familiarity. However, as the experiment continues and more runs are presented the monitoring task becomes more demanding, making it more difficult for older adults to use recollection. When presented with rearranged pairs, older adults significantly made more false alarms than children and younger adults with false alarms declining for all three age groups across run, suggesting a potential use of some type of recall-to-reject mechanism.

Unlike many other studies that suggested that older adults cannot use recall-to-reject, Gallo, Cotel, Moore, and Schacter (2007) found that older adults can effectively use recall-to-reject when they study items that were mutually exclusive at test; that is, when items were presented either as a word or a picture, but not as both at study. Furthermore, participants in the Gallo et al. study were told that if they remember seeing that item as a picture at study then they did not see that item as word, subsequently allowing them the use of a recall-to-reject strategy. Performance in this condition was compared to a nonexclusive condition in which stimuli could be presented as both a red word and a picture at study. Adding the mutually exclusive condition resulted in a greater reduction of false alarms for studied pictures in the red word test for younger and older adults suggesting that both groups were using the recall-to-reject strategy. When investigating the distinctiveness heuristic, Gallo et al. found that false alarms decreased more on the picture test than the word test confirming the use of this retrieval strategy. Moreover, Gallo et al.'s study suggests that even though older adults show a decline in their ability to use recall-to-reject, they are able to use such a strategy when given distinct stimuli and a mutually exclusive rule to decrease their false alarms.

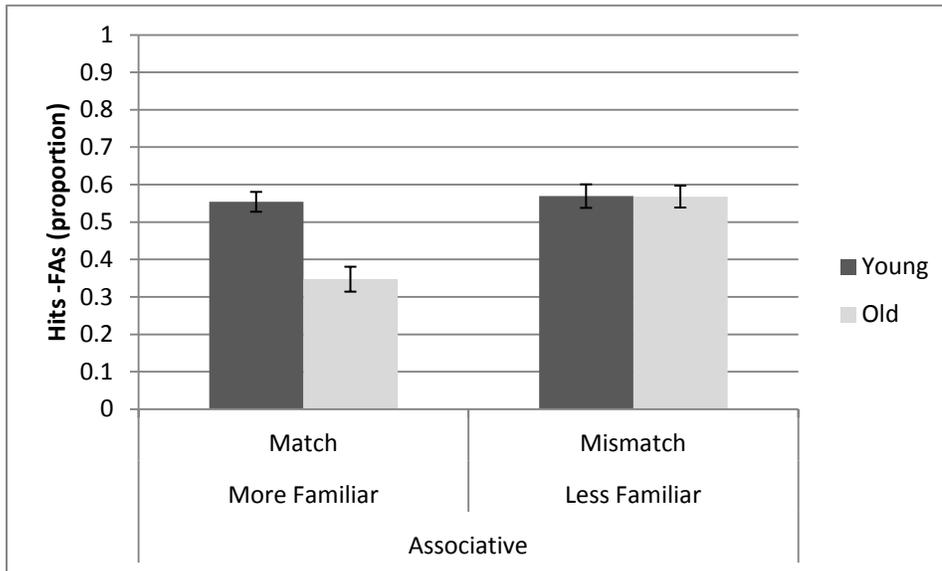
The results of our studies support Gallo, Cotel, Moore, and Schacter (2007) results using a direct associative memory recognition task instead of a criterial recognition task. Furthermore, whereas Gallo et al.'s study examined false recognition based on recollection expectations and the usage of the recall-to-reject and distinctiveness heuristic in younger and older adults with familiarity differences minimized between red words and pictures in an attempt to equate the distinctiveness of pictures and words, we examined simultaneously a manipulation of a recall-to-reject retrieval strategy (mismatch vs. match) and of familiarity (by repetition) to assess how item familiarity and recall-to-reject interacted with age.

There is another study that found that older adults are able to use a recall-to-reject strategy. Patterson, Light, Van Ocker, and Olfman (2009) examined if semantic relatedness aided in the utilization of the recall-to-reject strategy for younger and older adults when analyzing associative memory using an associative recognition task. Patterson et al. manipulated recall-to-reject by taking two pairs which were semantically related at study (e.g. chair-table and envy-jealousy) and rearranging them at test (e.g. chair-jealousy). If participants could remember that "chair" was presented with a semantically related word at study, they could use recall-to-reject at test. False alarms for recombined pairs declined for younger and older adults when the original pair was semantically related at study and when given enough time to respond. Moreover, semantic relatedness minimizes the demands at retrieval by facilitating the binding at encoding. Although Patterson et al.'s study examined false alarm rates, they did not assess whether the overall associative deficit of older adults was smaller due to changes only in false alarms or also in hits. In addition, they also did not employ an item

recognition test to confirm that there were no memory deficiencies for the individual item-components.

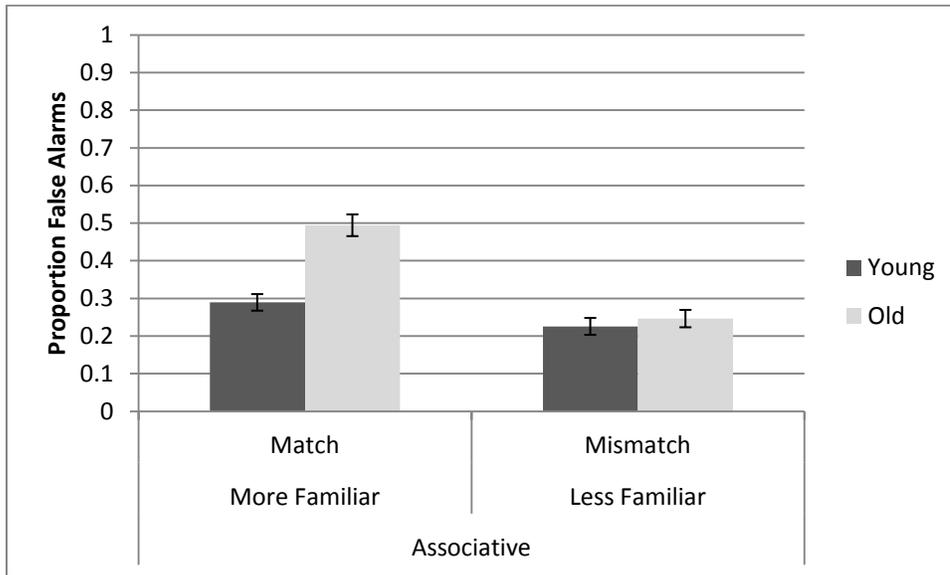
The results of both experiments suggest that older adults' overall associative memory performance improves when lowering their associative false alarm rates by utilizing the recall-to-reject strategy (in these experiments in the mismatch condition) to overcome feelings of item familiarity. This was indicated in a final set of statistical analyses assessing our hypothesis regarding the predictions for the combined effects of the two variables over the results of both experiments. An ANOVA examining overall memory performance (see Figure 17) showed a significant interaction between the type of condition (more familiar-match condition vs. the less familiar-mismatch) and age of the participant (younger vs. older adults; $F(1, 164) = 21.80, p < .001, \eta^2 = .12$). Follow-up t-tests suggested there were significant differences between younger ($M = .55, SD = .25$) and older adults ($M = .35, SD = .30$) in the more familiar-match condition ($t(164) = 4.88, p < .001$). However, there were no significant differences in the less familiar-mismatch condition ($t(164) = .03, p = .98$) between younger ($M = .57, SD = .29$) and older adults ($M = .57, SD = .26$).

Figure 17. Mean Proportion Hits minus Proportion False Alarms Hypothesis Predictions (Experiment 1 and Experiment 2 Follow-up Analyses) with Error Bars Indicating Standard Errors



A similar ANOVA was conducted on false alarms (see Figure 18). There was a significant interaction between age of the participant and the type of associative test condition ($F(1, 164) = 23.14, p < .001, \eta^2 = .12$). Similar to the results of overall memory performance, older adults ($M = .49, SD = .26$) showed significantly higher proportion of false alarms committed in the more familiar-match condition ($t(150.22) = 5.63, p < .001$), relative to younger adults ($M = .29, SD = .20$), but there were no differences between the two age groups in the less familiar-mismatch condition ($t(164) = .64, p = .52; M = .23, SD = .21$, and $M = .25, SD = .21$, for younger and older adults, respectively).

Figure 18. Mean Proportion False Alarms Hypothesis Predictions (Experiments 1 and Experiment 2 Follow-up Analyses) with Error Bars Indicating Standard Errors



In conclusion, the current set of experiments provides support for the notion that as people age there is a decline in their overall associative memory performance (Naveh-Benjamin, 2000). The results suggest that the deficit is mediated to a greater degree by an increase in false alarms rather than a decline in the hit rate in the associative recognition test. Furthermore, the current results assess several mechanisms underlying this deficit and show the manipulations to augment the operation of these mechanisms to improve older adults' associative memory performance by reducing their high false alarm rate. Specifically, a decline in item familiarity was shown to decrease older adults' associative false alarms (Experiment 2) and an increase in schematic support also decreased associative false alarms (Experiments 1 and 2) possibly by allowing older adults to use a type of recall-to-reject strategy. By identifying that an increase in false

alarms is mediating older adults' associative deficit, researchers can continue to investigate other encoding and retrieval monitoring strategies that older adults could use to further decrease their false alarms and therefore improve their associative memory.

REFERENCES

- Brainerd, C. J., & Reyna, V. F. (1998). Fuzzy-trace theory and children's false memories. *Journal of Experimental Child Psychology, 71*, 81-129. doi:10.1006/jecp.1998.2464
- Brainerd, C.J., Reyna, V.F., & Kneer, R. (1995). False recognition reversal: When similarity is distinctive. *Journal of Memory and Language, 34*, 157-185. doi:10.1006/jmla.1995.1008
- Castel, A. D. (2005). Memory for grocery prices in younger and older adults: the role of schematic support. *Psychology and Aging, 20*(4), 718–721. doi:10.1037/0882-7974.20.4.718
- 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*(4), 873–885. doi:10.1037/0882-7974.18.4.873
- Chalfonte, B. L., & Johnson, M. K. (1996). Feature memory and binding in young and older adults. *Memory & Cognition, 24*(4), 403–416. doi:10.3758/BF03200930
- Cohn, M., Emrich, S. M., & Moscovitch, M. (2008). Age-related deficits in associative memory: The influence of impaired strategic retrieval. *Psychology and Aging, 23*(1), 93–103. doi:10.1037/0882-7974.23.1.93
- Cohn, M., & Moscovitch, M. (2007). Dissociating measures of associative memory: Evidence and theoretical implications. *Journal of Memory and Language, 57*(3), 437–454. doi:10.1016/j.jml.2007.06.006
- Craik, F. I. M., & Byrd, M. (1982). Aging and cognitive deficits: The role of attentional resources. In F. I. M. Craik & S. E. Trehub (Eds.), *Aging and cognitive processes* (pp.191-211). New York: Plenum Press.
- Ebner, N. C., Riediger, M., & Lindenberger, U. (2010). FACES—A database of facial expressions in young, middle-aged, and older women and men: Development and validation. *Behavior Research Methods, 42*, 351-362. doi:10.3758/BRM.42.1.351
- Fandakova, Y., Shing, Y. L., & Lindenberger, U. (2013a). Differences in binding and monitoring mechanisms contribute to lifespan age differences in false memory. *Developmental Psychology, 49*(10), 1822-1832. doi:10.1037/a003161
- Fandakova, Y., Shing, Y. L., & Lindenberger, U. (2013b). High-confidence memory errors in old age: The roles of monitoring and binding processes. *Memory, 21*(5), 732-750. doi:10.1080/09658211.2012.756038

- Gallo, D. A. (2004). Using recall to reduce false recognition: Diagnostic and disqualifying monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(1), 120–128. doi:10.1037/0278-7393.30.1.120
- Gallo, D. A., Bell, D. M., Beier, J. S., & Schacter, D. L. (2006). Two types of recollection-based monitoring in younger and older adults: Recall-to-reject and the distinctiveness heuristic. *Memory*, *14*(6), 730–741. doi:10.1080/09658210600648506
- Gallo, D. A., Cotel, S. C., Moore, C. D., & Schacter, D. L. (2007). Aging can spare recollection-based retrieval monitoring: The importance of event distinctiveness. *Psychology and Aging*, *22*(1), 209–213. doi:10.1037/0882-7974.22.1.209
- Gardiner, J. M., Ramponi, C., & Richardson-Klavehn, A. (1998). Experiences of remembering, knowing, and guessing. *Consciousness and Cognition: An International Journal*, *7*(1), 1-26. doi:10.1006/ccog.1997.0321
- Healy, M. R., Light, L. L., & Chung, C. (2005). Dual-process models of associative recognition in young and older adults: evidence from receiver operating characteristics. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *31*(4), 768–788. doi:10.1037/0278-7393.31.4.768
- Howard, M. W., Bessette-Symons, B., Zhang, Y., & Hoyer, W. J. (2006). Aging selectively impairs recollection in recognition memory for pictures: Evidence from modeling and receiver operating characteristic curves. *Psychology and Aging*, *21*, 96-106. doi:10.1037/0882-7974.21.1.96
- Jacoby, L. L. (1999). Ironic effects of repetition: measuring age-related differences in memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *25*(1), 3–22. doi:10.1037/0278-7393.25.1.3
- Jennings, J. M., & Jacoby, L. L. (1997). An opposition procedure for detecting age-related deficits in recollection: Telling effects of repetition. *Psychology and Aging*, *12*(2), 352–361. doi:10.1037/0882-7974.12.2.352
- Jones, T. C., & Jacoby, L. L. (2001). Feature and conjunction errors in recognition memory: Evidence for dual-process theory. *Journal of Memory and Language*, *45*(1), 82–102. doi:10.1006/jmla.2000.2761
- Kausler, D. H., & Puckett, J. M. (1980). Frequency judgments and correlated cognitive abilities in young and elderly adults. *Journal of Gerontology*, *35*, 376-382. doi:10.1093/geronj/35.3.375
- Kilb, A., & Naveh-Benjamin, M. (2011). The effects of pure pair repetition on younger and older adults' associative memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(3), 706–719. doi:10.1037/a0022525

- Leding, J. K., & Lampinen, J. M. (2009). Memory conjunction errors: The effects of presentation duration and study repetition. *Memory, 17*(5), 597–607. doi:10.1080/09658210902984518
- Light, L. L., Chung, C., Pendergrass, R., & Van Ocker, J. C. (2006). Effects of repetition and response deadline on item recognition in young and older adults. *Memory & Cognition, 34*(2), 335–343. doi:10.3758/BF03193411
- 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
- Light, L. L., Prull, M. W., La Voie, D. J., & Healy, M. R. (2000). Dual-process theories of memories in old age. In T.J. Perfect & E.A. Maylor (Eds.), *Models of cognitive aging* (pp. 238-300). New York: Oxford University Press, Inc.
- McGillivray, S. & Castel, A. D. (2010). Memory for age-face associations in younger and older adults: The role of generation and schematic support. *Psychology and Aging, 25*, 822-832. doi:10.1037/a0021044
- Minear, M., & Park, D.C. (2004). A life-span database of adult facial features. *Behavior Research Methods, Instruments, & Computers, 36* (4), 630-633. doi:10.3758/BF03206543
- 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
- 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., Guez, J., Kilb, A., & Reedy, S. (2004). The associative memory deficit of older adults: further support using face-name associations. *Psychology and Aging, 19*(3), 541–546. doi:10.1037/0882-7974.19.3.541
- Naveh-Benjamin, M., Hussain, Z., Guez, J., & Bar-On, M. (2003). Adult age differences in episodic memory: Further support for an associative-deficit hypothesis. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 29*(5), 826–837. doi:10.1037/0278-7393.29.5.826
- Naveh-Benjamin, M., Shing, Y. L., Kilb, A., Werkle-Bergner, M., Lindenberger, U., & Li, S. (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

- Odegard, T. N., Lampinen, J. M., & Toglia, M. P. (2005). Meaning's moderating effect on recollection rejection. *Journal of Memory and Language*, *53*(3), 416–429. doi:10.1016/j.jml.2005.04.004
- Old, S. R., & Naveh-Benjamin, M. (2008a). 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
- Old, S. R., & Naveh-Benjamin, M. (2008b). Memory for people and their actions: Further evidence for an age-related associative deficit. *Psychology and Aging*, *23*, 467-472. doi:10.1037/0882-7974.23.2.467
- 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
- Rhodes, M. G. (2009). Age estimation of faces: A review. *Applied Cognitive Psychology*, *23*, 1-12. doi:10.1002/acp.1442
- Rotello, C. M., & Heit, E. (2000). Associative recognition: A case of recall-to-reject processing. *Memory & Cognition*, *28*(6), 907–922. doi:10.3758/BF03209339
- Schmid, J., Herholz, S. C., Brandt, M., & Buchner, A. (2010). Recall-to-reject: The effect of category cues on false recognition. *Memory*, *18*(8), 863–882. doi:10.1080/09658211.2010.517756
- Shing, Y. L., Werkle-Bergner, M., Li, S.-C., & Lindenberger, U. (2008). Associative and strategic components of episodic memory: A life-span dissociation. *Journal of Experimental Psychology: General*, *137*(3), 495–513. doi:10.1037/0096-3445.137.3.495
- Tulving, E. (1985). Memory and consciousness. *Canadian Psychology*, *26*, 1-12. doi:10.1037/h0080017
- Wilson, M. D. (1988). The MRC Psycholinguistic Database: Machine Readable Dictionary, Version 2. *Behavioural Research Methods, Instruments and Computers*, *20*(1), 6-11.
- Yonelinas, A. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*, 441-517. doi:10.1037/h0080017
- Zacks, R. T., Hasher, L., & Li, K. Z. H. (2000). Human memory. In T. A. Salthouse & F. I. M. Craik (Eds.), *Handbook of Aging and Cognition*, 2nd edition (pp.293-357). Mahwah, NJ: Lawrence Erlbaum.

APPENDIX

Appendix A

Table A1. Mean Proportion of Hits minus Proportion False Alarms with Standard Deviations for Experiment 1

| | | Item | | | | Associative | | | |
|--------------|-------------|---------------|---------------|---------------|---------------|---------------|-------------|---------------|-------------|
| | | Product | | Price | | More Familiar | | Less Familiar | |
| | | More Familiar | Less Familiar | More Familiar | Less Familiar | Match | Mismatch | Match | Mismatch |
| Young | Mean | 0.97 | 0.95 | 0.51 | 0.36 | 0.59 | 0.71 | 0.53 | 0.70 |
| | SD | 0.06 | 0.09 | 0.27 | 0.21 | 0.23 | 0.16 | 0.25 | 0.19 |
| Old | Mean | 0.97 | 0.93 | 0.5 | 0.27 | 0.43 | 0.69 | 0.41 | 0.68 |
| | SD | 0.05 | 0.11 | 0.23 | 0.23 | 0.24 | 0.23 | 0.19 | 0.21 |

Table A2. Mean Proportion Hits with Standard Deviations for Experiment 1

| | | Item | | | | Associative | |
|--------------|-------------|---------------|---------------|---------------|---------------|---------------|---------------|
| | | Product | | Price | | More Familiar | Less Familiar |
| | | More Familiar | Less Familiar | More Familiar | Less Familiar | More Familiar | Less Familiar |
| Young | Mean | 0.98 | 0.97 | 0.76 | 0.6 | 0.93 | 0.91 |
| | SD | 0.05 | 0.07 | 0.16 | 0.19 | 0.07 | 0.07 |
| Old | Mean | 0.98 | 0.94 | 0.76 | 0.56 | 0.92 | 0.9 |
| | SD | 0.05 | 0.09 | 0.2 | 0.24 | 0.07 | 0.08 |

Table A3. Mean Proportion False Alarms with Standard Deviations for Experiment 1

| | | Item | | | | Associative | | | |
|--------------|-------------|---------------|---------------|---------------|---------------|---------------|-------------|---------------|-------------|
| | | Product | | Price | | More Familiar | | Less Familiar | |
| | | More Familiar | Less Familiar | More Familiar | Less Familiar | Match | Mismatch | Match | Mismatch |
| Young | Mean | 0.004 | 0.02 | 0.25 | 0.24 | 0.34 | 0.22 | 0.38 | 0.21 |
| | SD | 0.02 | 0.04 | 0.21 | 0.17 | 0.20 | 0.14 | 0.21 | 0.16 |
| Old | Mean | 0.01 | 0.02 | 0.26 | 0.29 | 0.48 | 0.23 | 0.49 | 0.22 |
| | SD | 0.03 | 0.05 | 0.19 | 0.17 | 0.22 | 0.22 | 0.18 | 0.17 |

Table A4. Mean Proportion Hits minus Proportion False Alarms with Standard Deviations for Experiment 2

| | | Item | | | | Associative | | | |
|--------------|-------------|---------------|---------------|---------------|---------------|---------------|-------------|---------------|-------------|
| | | Face | | Name | | More Familiar | | Less Familiar | |
| | | More Familiar | Less Familiar | More Familiar | Less Familiar | Match | Mismatch | Match | Mismatch |
| Young | Mean | 0.72 | 0.56 | 0.70 | 0.51 | 0.49 | 0.55 | 0.43 | 0.45 |
| | SD | 0.20 | 0.17 | 0.19 | 0.16 | 0.25 | 0.26 | 0.24 | 0.25 |
| Old | Mean | 0.72 | 0.58 | 0.68 | 0.54 | 0.30 | 0.34 | 0.33 | 0.38 |
| | SD | 0.16 | 0.18 | 0.15 | 0.15 | 0.25 | 0.24 | 0.25 | 0.24 |

Table A5. Mean Proportion of Hits with Standard Deviations for Experiment 2

| | | Item | | | | Associative | |
|--------------|-------------|------------------|------------------|------------------|------------------|------------------|------------------|
| | | Face | | Name | | | |
| | | More Familiar | Less Familiar | More Familiar | Less Familiar | More Familiar | Less Familiar |
| Young | Mean | 0.81 | 0.65 | 0.79 | 0.63 | 0.77 | 0.67 |
| | SD | 0.14 | 0.14 | 0.15 | 0.16 | 0.14 | 0.14 |
| Old | Mean | 0.91 | 0.79 | 0.85 | 0.71 | 0.78 | 0.70 |
| | SD | 0.08 | 0.12 | 0.12 | 0.15 | 0.11 | 0.13 |

Table A6. Mean Proportion False Alarms with Standard Deviations for Experiment 2

| | | Item | | | | Associative | | | |
|--------------|-------------|------------------|------------------|------------------|------------------|------------------|-------------|------------------|-------------|
| | | Face | | Name | | More Familiar | | Less Familiar | |
| | | More Familiar | Less Familiar | More Familiar | Less Familiar | Match | Mismatch | Match | Mismatch |
| Young | Mean | 0.10 | 0.09 | 0.09 | 0.12 | 0.28 | 0.22 | 0.24 | 0.21 |
| | SD | 0.10 | 0.08 | 0.12 | 0.15 | 0.18 | 0.18 | 0.17 | 0.17 |
| Old | Mean | 0.19 | 0.21 | 0.17 | 0.17 | 0.48 | 0.34 | 0.37 | 0.32 |
| | SD | 0.14 | 0.16 | 0.11 | 0.13 | 0.21 | 0.22 | 0.20 | 0.22 |

Appendix B

Figure B1. Mean Proportion Hits minus Proportion False Alarms from Experiment 1 with error bars indicating standard error broken down into market-value (MV), over- (over), and underestimated (under) prices. A match is a market-value, over-, or underestimated product-price pair at study that is recombined with another product or price at test within the same price category. A mismatch is a market-value, over-, or underestimated price that appeared at study that was later recombined with a product-price pair at test outside of its category (i.e. a market value mismatch is product-price pair that appeared at market value at study and then was recombined with either an over- or underestimated product-price pair at test).

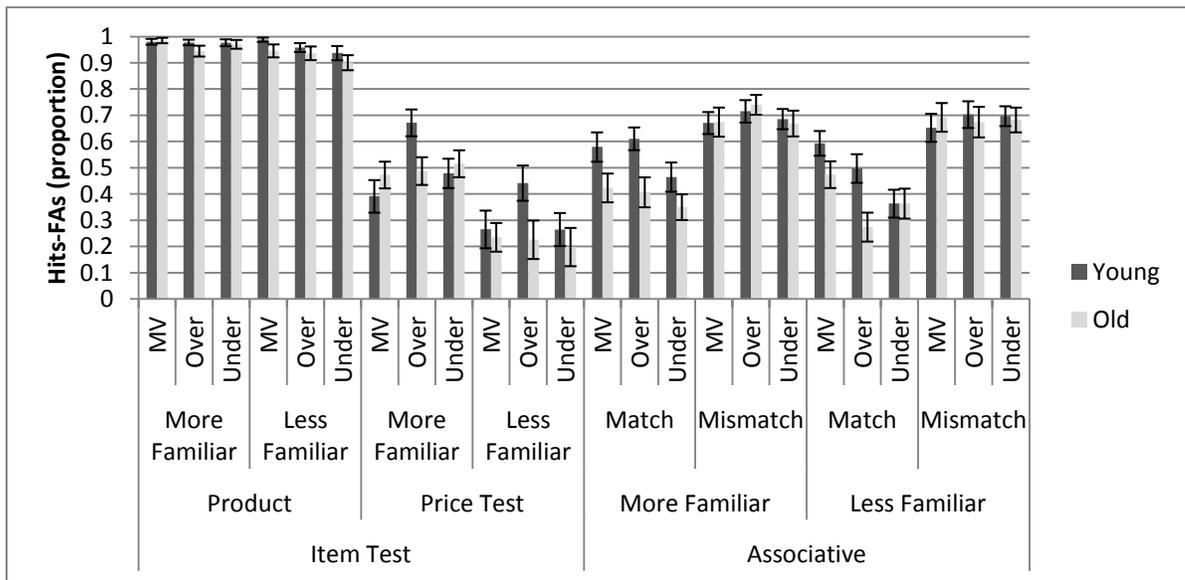


Figure B2. Mean Proportion Hits from Experiment 1 with error bars indicating standard error broken down into market-value (MV), over- (over), and underestimated prices

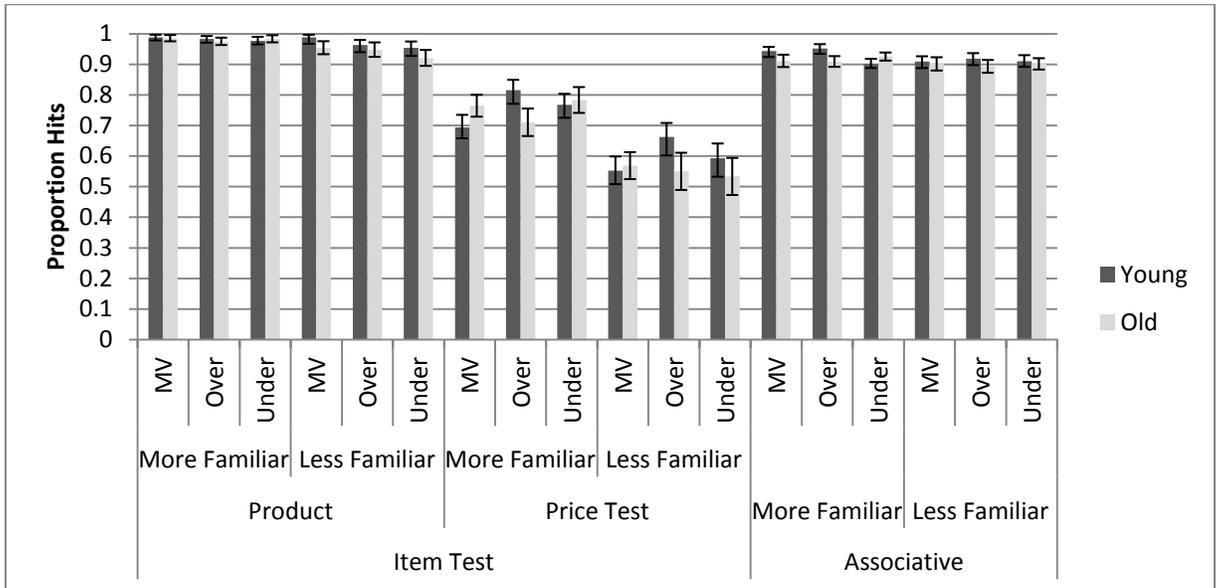


Figure B3. Mean Proportion False Alarms with error bars indicating standard error broken down into market-value (MV), over- (over), and underestimated prices

