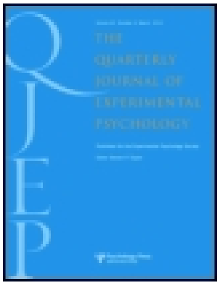
This article was downloaded by: [amy criss]
On: 14 March 2014, At: 13:41
Publisher: Routledge
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



The Quarterly Journal of Experimental Psychology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/pqje20

Age and response bias: Evidence from the strength-based mirror effect

Amy H. Criss[a], William Aue[a] & Aslı Kılıç[a]

[a] Department of Psychology, Syracuse University, Syracuse, NY, USA

Accepted author version posted online: 06 Jan 2014.Published online: 03 Mar 2014.

To cite this article: Amy H. Criss, William Aue & Aslı Kılıç (2014): Age and response bias: Evidence from the strength-based mirror effect, The Quarterly Journal of Experimental Psychology, DOI: 10.1080/17470218.2013.874037

To link to this article: http://dx.doi.org/10.1080/17470218.2013.874037

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the “Content”) contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at http://www.tandfonline.com/page/terms-and-conditions

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY, 2014
http://dx.doi.org/10.1080/17470218.2013.874037

Routledge
Taylor & Francis Group

Age and response bias: Evidence from the strength-based mirror effect

Amy H. Criss, William Aue, and Aslı Kılıç

Department of Psychology, Syracuse University, Syracuse, NY, USA

Performance in episodic memory is determined both by accurate retrieval from memory and by decision processes. A substantial body of literature suggests slightly poorer episodic memory accuracy for older than younger adults; however, age-related changes in the decision mechanisms in memory have received much less attention. Response bias, the willingness to endorse an item as remembered, is an important decision factor that contributes to episodic memory performance, and therefore understanding age-related changes in response bias is critical to theoretical development. We manipulate list strength in order to investigate two aspects of response bias. First, we evaluate whether criterion placement in episodic memory differs for older and younger adults. Second, we ask whether older adults have the same degree of flexibility to adjust the criterion in response to task demands as younger adults. Participants were tested on weakly and strongly encoded lists where word frequency (Experiment 1) or similarity between targets and foils (Experiment 2) was manipulated. Both older and younger adults had higher hit rates and lower false-alarm rates for strong lists than for weak lists (i.e., a strength-based mirror effect). Older adults were more conservative (less likely to endorse an item as studied) than younger adults, and we found no evidence that older and younger adults differ in their ability to flexibly adjust their criterion based on the demands of the task.

Keywords: Episodic memory; Memory models; Response bias; Mirror effects; Recognition memory.

Downloaded by [amy criss] at 13:41 14 March 2014

Recognition memory is the ability to identify information experienced in a specific prior episode. In the laboratory, recognition is commonly studied by presenting a participant with a series of words and later asking them to determine which words were presented in the experimental context (i.e., targets) and which were not (i.e., foils). A substantial body of research has evaluated age-related changes in episodic memory, showing both global and task-specific deficits (see Burke & Light, 1981; Light, 1991; Verhaeghen, Marcoen, & Goossens, 1993). Deficits in episodic memory, especially recall, are often associated with both healthy ageing and age-related disorders such as Alzheimer's. However, relatively little research has addressed age-related changes in decision processes as a mechanism for changes in memory.

Signal detection theory (SDT) is a framework for measuring factors underlying the decision process, and SDT can be applied to recognition

Correspondence should be addressed to Amy Criss, Department of Psychology, Syracuse University, Syracuse, NY, USA. E-mail: acriss@syr.edu

Current Address: Aslı Kılıç is now at the Department of Psychology, Koç University, Istanbul, Turkey.

We thank Matt Williams and Stephanie Cuevas for data collection and management.

This material is partially based upon work supported by the National Science Foundation [grant number 0951612]. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

© 2014 The Experimental Psychology Society

memory. SDT characterizes memory as consisting of two distributions, one for targets and one for foils, along an axis of memory strength (see Parks, 1966). Since the targets are studied, the mean memory strength of the target distribution is greater than that of the foil distribution. In order for an item to be endorsed as a target, the strength of the item must exceed a criterion. Items with memory strength above the criterion are endorsed as having been studied previously, while items falling below the threshold are judged as new. Trials on which an item is correctly endorsed as being old are called hits (HR) while trials on which an item is incorrectly endorsed as old are called false alarms (FAR), and these two measurements are the critical dependent variables in recognition.

Within the SDT framework, there are two factors that influence memory decisions and, by extension, memory performance: criterion placement and discrimination (Green & Swets, 1966). The ability to discriminate between studied and unstudied items during test is an indication of accuracy and is determined by the overlap between the target and foil distributions (e.g., d' of 0 indicates perfectly overlapping distributions). Discrimination, or d', can be manipulated experimentally—for example, by strengthening memory via repetition or levels of processing (e.g., Criss, 2009; Kılıç, 2013). Discrimination might also reflect properties of the stimulus or the participant. Criterion placement can be manipulated experimentally, for example, through altering expectations. For example, if 75% of the test items have been studied, participants tend to endorse more items (both targets and foils) as studied. Conversely, when only 25% of the test items are targets, participants tend to endorse fewer items as studied (Criss, 2010; Healy & Kubovy, 1978; Rotello, Macmillan, Hicks, & Hautus, 2006). Within SDT, this behaviour is reflected in a change in the placement of the criterion, measured by c. An optimal criterion (i.e., $c = 0$) indicates that the criterion is placed at the intersection of the target and foil distributions so as to maximize performance. A criterion can be conservative in that memory strength must be very high in order for an item to be endorsed as old or liberal such that even low memory strength results in an endorsement. In addition to experimental manipulations such as target expectancy, criterion placement may be affected by properties of individuals or groups of participants. Thus response bias and discriminability represent two aspects of memory that work in concert to produce the behavioural measurements of HR and FAR.

Older adults tend to have slightly lower discriminability than younger adults (e.g., Craik & McDowd, 1987), but little research has examined whether there exists age-related differences in criterion placement. Criterion placement is a crucial determinant of memory performance because it indicates the level of evidence required for a participant to claim that they remember an item. Understanding age-related changes in criterion placement is therefore critical to theoretical development (e.g., Treisman & Williams, 1984; Turner, Van Zandt, & Brown, 2011). In the current paper, we were interested in two questions about age-related changes in criterion placement. First, does criterion placement for a given task differ as a function of age? That is, do older adults set a more liberal or conservative criterion relative to younger adults? Second, when the experimental situation warrants a change in criterion placement, are older adults as flexible in changing their criterion as younger adults?

Criterion placement and ageing

Research examining age-related changes in criterion placement has produced mixed data (see McCormack, 1984, for a review). Some studies report no significant difference in criterion placement across age groups (e.g., Baron & Le Breck, 1987; Isingrini, Fontaine, Taconnat, & Duportal, 1995) while other studies find that older adults use a more liberal criterion. For example, Huh, Kramer, Gazzaley, and Delis (2006) observed a tendency for older adults to set a more liberal criterion with increasing age. Even when individual characteristics (e.g., gender, IQ, education) of the older adults were taken into account via a regression model, age still accounted for a significant proportion of the variance. Still others report that

Downloaded by [amy criss] at 13:41 14 March 2014

older adults adopt a more conservative response bias than younger adults. For example, Poon and Fozard (1980) compared response bias across three age groups in a continuous recognition memory task. They observed that older adults tended to set a more conservative criterion than the youngest group. Unfortunately, there appears to be no consensus on age-related changes in criterion placement.

Criterion flexibility and ageing

The literature on criterion flexibility is just as mixed as the literature on criterion placement. Baron and Surdy (1990) manipulated response bias through a pay-off matrix and reported that while older adults adjusted their criteria in response to the pay-off, they did so less than younger adults and less than required to reach the maximum pay-off. Other studies have reported that older adults are as skilled at flexibly adjusting their criterion placement as younger adults (e.g., Koutstaal, Schacter, Galluccio, & Stofer, 1999). For example, in Pendergrass, Olfman, Schmalstig, Seder, and Light (2012), participants studied a list of word pairs and then received both a difficult test and an easy test. The difficult test list was composed of target pairs and foil pairs where foil pairs were constructed by rearranging studied items. The easy test list was composed of target pairs and foil pairs, where foil pairs were constructed from non-studied items. Participants were not informed about the difference between the tests. When participants were given the easy test first, both younger and old participants tended to set a stricter criterion for the difficult test. However, when the difficult test was administered first, there was no shift in criterion across tests, instead participants maintained the strict criterion even for the easy test. The same pattern and magnitude of criterion flexibility was observed for older and younger adults, suggesting that older adults are not any less flexible than younger adults in changing their criterion based on task demands. The criterion considered in the studies described above comes from response bias in the SDT framework and concerns whether to endorse an item as remembered or not. An alternative view of criterion comes from considering how quickly the response is made.

In recognition memory experiments—and in fact most cognitive tasks—older adults are slower than younger adults. This general slowing is often attributed to a general cognitive decline associated with ageing (e.g., Cerella, 1985; Salthouse, 1996). However, diffusion model analyses provide an alternative explanation—that older adults tend to be slower due to cautiousness rather than an overall decline in ability. The diffusion model can be thought of as a dynamic version of SDT. Like SDT, parameters of the diffusion model have interpretable cognitive correlates; however, the mapping between SDT and the diffusion model is not one-to-one. The distributions of memory strength correspond to the drift rate parameters in the diffusion model and reflect the quality of the evidence. However, the diffusion model has multiple different types of criteria. One, boundary separation, has been the focus of studies evaluating age-related changes. Boundary separation reflects the amount of evidence collected before a decision is made. When the two decision boundaries (i.e., studied and not studied in a single item recognition experiment) are close together, little evidence is collected before a decision is made, which leads to fast, inaccurate responses. As the boundaries move farther apart, response time slows down, and accuracy improves. Thus, boundary separation is viewed as reflecting cautiousness. As described next, many studies have shown that older adults are not as skilled as younger adults at modifying their cautiousness to reflect the experimental situation.

Ratcliff, Thapar, and McKoon (2004) had participants engage in multiple study–test blocks, half under the instruction to respond quickly and the other half under the instruction to respond accurately. Response time data were interpreted within the diffusion model, and this analysis showed that memory strength did not differ for older and younger adults but reading and motor time was longer for older than younger adults. Under instructions to emphasize speed over accuracy, younger adults followed instructions to respond quickly and sacrificed accuracy but older adults did not—they remained cautious. Similar

Downloaded by [amy criss] at 13:41 14 March 2014

Downloaded by [amy criss] at 13:41 14 March 2014

data were reported by Starns and Ratcliff (2010)—namely, no differences in memory strength but more conservative responding for older adults. Starns and Ratcliff computed optimal values of cautiousness (defined as the value of the boundary separation parameter that maximized accuracy per unit of time). Younger adults were closer to optimal than older adults, and older adults were overly cautious such that they could have responded faster while maintaining the same level of accuracy. Even when specifically instructed to optimize boundary settings, older adults were not able to do so as well as younger adults (Starns & Ratcliff, 2012). These findings showing that older adults do not adjust to instructions, but instead remain overly cautious, suggest that older adults are not able to flexibly adjust their criterion according to task demands.

It is important to note that these two types of criteria—response bias favouring a response (remembered or not) and boundary separation reflecting cautiousness about when to respond—need not show the same age-related changes. The literature based on using response times (and accuracy) to measure cautiousness seems to show a consistent deviation from optimality for older adults. The literature based on using accuracy to measure response bias shows mixed results with regard to whether older adults are as flexible as younger adults in adapting the criterion to experimental demands. Clearly, there is a need to better understand the relationship between response bias, criterion flexibility, and ageing. The aim of the current study was to examine age-related differences in response bias and flexibility using the strength-based mirror effect paradigm.

The strength-based mirror effect

In a typical strength-based mirror effect paradigm, participants study a weak list and a strong list where strength is defined as longer study duration, more repetitions, or “deeper” processing. A test follows each study list so there are two different false-alarm rates, one measured in the context of strong targets and one measured in the context of weak targets. For simplicity, the foils are often referred to as strong or weak despite the fact that the foils themselves were not studied and are drawn from the same pool (i.e., the preexperimental memory strength of the foils should be equivalent on average). In this paradigm, HRs are higher, and FARs are lower for strong lists than weak lists, and this is referred to as the strength-based mirror effect (SBME). The SBME is a robust finding, having been replicated many times in many different laboratories (e.g., Cary & Reder, 2003; Criss, 2006, 2009, 2010; Hockley & Niewiadomski, 2007; Starns, White, & Ratcliff, 2010; Stretch & Wixted, 1998, among others).

In one theoretical framework, strengthening items shifts the target distribution along the memory strength axis, producing an increase in the HR. The memory strength distribution for foils remains constant because there are no differences between foils tested after a weakly versus a strongly encoded list other than the encoding conditions of the target items. To account for a strength-based mirror effect within this framework, one must assume that the criterion changes as a function of list strength (cf. Criss, 2006, 2009, 2010; McClelland & Chappell, 1998; Stretch & Wixted, 1998). The idea is that participants become aware that accuracy for a strong list is very high, from either their experience during encoding or the initial test trials (Hirshman, 1995; Stretch & Wixted, 1998), and adopt a strict criterion based on this knowledge. The vast majority of memory models, both dual and single process, have adopted the criterion shift explanation for the strength-based mirror effect (e.g., Cary & Reder, 2003; Starns, White, & Ratcliff, 2010; Stretch & Wixted, 1998). Thus the strength-based mirror effect paradigm is an ideal situation in which to investigate differences in criterion placement and flexibility in older adults.

This manuscript includes two experiments, both of which make use of the strength-based mirror effect paradigm. The same group of people participated in both experiments. Participants were not aware that they were participating in two different experiments. From their perspective, they completed four study–test blocks. This maximized our ability to find any age-related changes in the

Downloaded by [amy criss] at 13:41 14 March 2014

flexibility of criterion placement because the criterion was, presumably, adjusted for each of the four study–test blocks. In Experiment 1, normative word frequency was manipulated in addition to list strength. Word frequency and list strength mirror effects are similar in magnitude. In the event that a SBME was not observed for older adults, we wanted to be sure to demonstrate that it was not simply a lack of experimental power, but rather a meaningful finding. In Experiment 2, we manipulated similarity between foils and studied items to place additional demands on memory. We used similarity to tax memory because older adults are more vulnerable to falsely claiming to remember foils that are similar to studied targets (e.g., Balota et al., 1999; Benjamin, 2001; Jacoby, 1999).

EXPERIMENT 1

Experiment 1 makes use of the basic strength-based mirror effect manipulation of a strong study list followed by a test list and a weak study list followed by a test list. In addition, normative word frequency was manipulated. Word frequency is also known to produce a mirror effect (e.g., Glanzer & Adams, 1985) though in this case the stimuli differ in preexperimental frequency. Stretch and Wixted (1998) distinguish between stimulus-based mirror effects (e.g., via a word frequency manipulation), presumably resulting from different memory strength distributions, and strength-based mirror effects (e.g., via a study repetition manipulation), presumably resulting from criterion shifts. In other words, the word frequency mirror effect and strength-based mirror effect are separate findings that result from different empirical manipulations and different cognitive operations. Despite the different origins, word frequency and strength-based mirror effects are similar in magnitude (e.g., Criss, 2010; Stretch & Wixted, 1998), and including a word frequency manipulation serves to demonstrate sufficient experimental power to detect an effect. Because the word frequency mirror effect is the result of different memory strength distributions rather than a criterion shift (e.g., Glanzer & Bowles, 1976), age-related changes in the word frequency effect are not critical for the present purposes (see Balota, Burgess, Cortese, & Adams, 2002; Ratcliff et al., 2004, for a discussion of word frequency effects and ageing). If older adults are not able to flexibly control their criterion placement in response to task demands, then the SBME should be distorted. If older and younger adults have similar flexibility in criterion placement, then the magnitude of the strength-based mirror effect should not differ with age. The literature is mixed with respect to criterion placement—some studies show that older adults are more liberal, and others show that older adults are more conservative. This experimental design is optimal to help clarify whether older adults are generally more conservative or more liberal in a recognition memory paradigm and to evaluate the ability of older adults to flexibly adjust their criterion.

Method

Participants

A total of 36 younger adults and 21 (57% female) older adults from the Syracuse University community received $10 per hour or course credit for participation in the experiment. All participants reported good health and scored fewer than eight errors on the memory concentration portion of the Blessed exam, indicating they were free of dementia (Blessed, Tomlinson, & Roth, 1968). All participants were free of memory impairments and in good health based on self-report. Table 1 presents the demographics and background information about the participants. Three additional

Table 1. *Participant demographics*

Measure	*Younger adults*	*Older adults*
Age (years)	18.46 (0.76)	63.43 (7.51)
College	1 (0)	.83 (.38)
Vocabulary	29.49 (3.86)	34.52 (3.59)
Reasoning	15.92 (2.30)	14.95 (4.15)

Note: College indicates the proportion of participants who completed some college. Vocabulary (maximum score 40) and reasoning (maximum score 20) measured using the Shipley Institute of Living Scale (Shipley, 1940).

younger adults that participated but performed at or below chance were excluded.

Stimulus materials

The word pool consisted of 800 low-frequency (LF) and 800 high-frequency (HF) words between 4 and 7 letters in length. HF words ranged between 9 and 13 log frequency ($M = 10.46$) in the Hyperspace Analog to Language corpus (HAL; Balota et al., 2007), and LF words ranged between 3.5 and 6 log frequency ($M = 5.22$).

Design

The experiment was a 2 (word frequency: LF and HF) × 2 (list strength: strong and weak) × 2 (test type: target and foil) × 2 (age: younger and older) mixed design with age serving as the between-subjects factor. Each participant received a weak and a strong block with block order randomly chosen for each individual. Participants advanced to the next block at their own pace. The weak block consisted of 40 words, half LF and half HF, each studied for 2.5 s with a 500-ms interstimulus interval (ISI). Participants were simply instructed to study the list with no specific encoding task provided. A 45-s arithmetic distractor task separated study and test. The test was self-paced yes/no single item recognition with 80 words equally divided among the conditions. The strong block was identical with one exception: Each of the 40 items was studied four times each. To equate study–test lag across strength conditions, the entire set of 40 items was presented before any item repeated for each of the four study presentations. Within each study–test list, words were randomly assigned to conditions (strength and test type), and the order of words was randomly assigned for each participant. We present this as Experiment 1 but participants performed these two study–test blocks after completing Experiment 2.

Results and discussion

A 2 (word frequency: LF and HF) × 2 (list strength: strong and weak) × 2 (age: older and younger) mixed analysis of variance (ANOVA) was conducted for each dependent variable (FAR and HR) and each composite measure (discriminability, d', and response bias, c).[1] Figure 1 shows mean values of all measures.

Consistent with the literature, we observed better discrimination accuracy for LF than HF items and for strong than weak lists. The ANOVA confirmed higher d' for LF than for HF words, $F(1, 55) = 44.52$, $p < .001$, $\eta^2 = .447$, and for the strong than for the weak list, $F(1, 55) = 61.52$, $p < .001$, $\eta^2 = .528$. Both the effect of repetition and the effect of word frequency appear as a mirror pattern. Strong HRs exceeded weak HRs, $F(1, 55) = 54.68$, $p < .001$, $\eta^2 = .498$, and strong FARs were lower than weak FARs, $F(1, 55) = 7.42$, $p = .009$, $\eta^2 = .119$. LF targets were better remembered, $F(1, 55) = 12.76$, $p = .001$, $\eta^2 = .188$, and LF foils were less likely to be endorsed than their HF counterparts, $F(1, 55) = 31.24$, $p < .001$, $\eta^2 = .362$. Older adults had lower FARs overall than younger adults, $F(1, 55) = 5.34$, $p = .025$, $\eta^2 = .088$. No other main effects or interactions were significant.

The primary purpose of including the word frequency manipulation was to demonstrate that we had sufficient power to detect a strength-based mirror effect; however, given our data, this is no longer a concern. For discussion of ageing and word frequency effects see Balota et al. (2002), Ratcliff et al. (2004), and Buchler and Reder (2007).

[1]Using SDT requires assumptions, though in many cases the assumptions are not explicitly acknowledged. For example, SDT assumes normally distributed memory evidence, which is inconsistent with many process models of memory. In both experiments, we compute SDT measures from yes/no data, which assumes identical variance for target and foil distributions. Further, the values of d' and c are computed separately for the different conditions, which implicitly assumes equal variance across conditions, and that optimality is condition-specific. Some empirical studies have demonstrated that these may be invalid assumptions, at least for certain types of behavioural responses (e.g., Mickes, Wixted, & Wais, 2007; Ratcliff, Sheu, & Gronlund, 1992), which indicates a limitation in the interpretation of these measures (a limitation shared by the vast majority of publications in the field). We conducted additional analyses assuming a variance ratio of .8 for foils to targets and replicated every result reported in the paper.

Downloaded by [amy criss] at 13:41 14 March 2014

Downloaded by [amy criss] at 13:41 14 March 2014

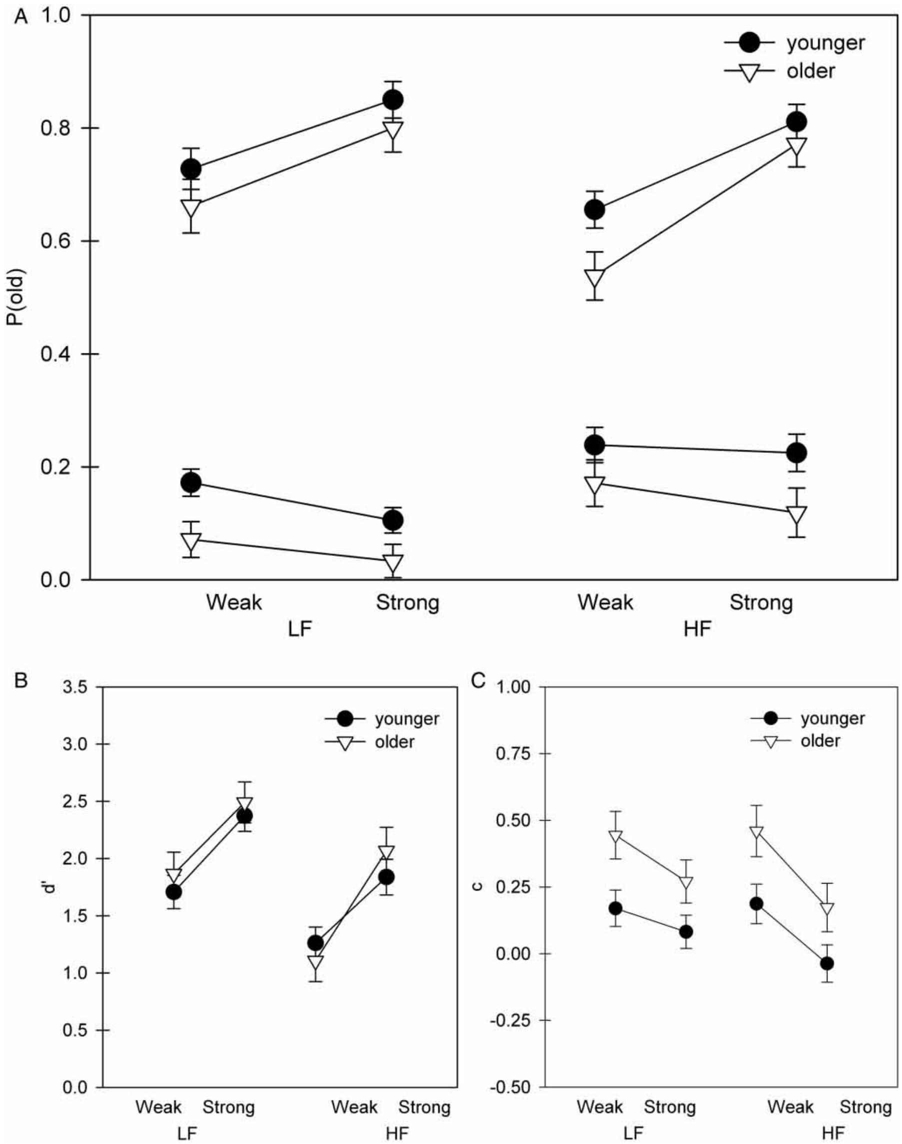


Figure 1. *Strength-based mirror effects were obtained for both for older and younger adults and for low frequency (LF) and high frequency (HF) words. Panel A: P(old) is the probability of endorsing an item as studied. Hits (top four lines) are higher, and false alarms (bottom four lines) are lower for strong than for weak lists. A word frequency mirror effect was observed with higher hits and lower false alarms for LF than HF words. Panel B: Discrimination (*d′*) was higher for LF words and strong lists. Panel C: Response bias (*c*) was more conservative and farther from optimal for older adults than for younger adults. Error bars are ±1 standard error.*

Analysis of response bias, *c*, showed an overall tendency to respond new. Note that *c* measures distance to the optimal criterion as measured by the point of intersection of the relevant target and foil distributions. A value of 0 is optimal, values greater than 0 indicate a conservative bias to respond new, and values less than 0 indicate a liberal bias to respond old. Overall the weak list elicited relatively more new responses than the strong list, $F(1, 55) = 20.20$, $p < .001$, $\eta^2 = .269$. Older adults were more cautious in their responding than younger adults, favouring the new response, $F(1, 55) = 6.92$, $p = .011$, $\eta^2 = .112$, but not differentially so across list strength. No other main effects or interactions were significant.

The data indicate that older adults are more cautious in the placement of their criteria, tending to reject more items than younger adults. However, the flexibility to adjust criterion placement in response to task demands is similar for older and younger adults as evidenced by strength-based mirror effects of similar magnitude and response bias values that change similarly across list strength (e.g., there were no interactions between age and any other variable).

EXPERIMENT 2

Experiment 2 is another strength-based mirror effect paradigm; however, word frequency was controlled, and word similarity was manipulated. The aim, however, was the same: to evaluate criterion placement and criterion flexibility in older adults. We include a similarity manipulation because older adults are typically more susceptible to false memory for foils that are similar to studied items (e.g., Balota et al., 1999; Benjamin, 2001; Jacoby, 1999). If older adults are at a disadvantage because they are impaired by similar foils, this may also manifest as a decline in their flexibility of criterion placement. In other words, we are specifically taxing older adults by including the similarity manipulation and expect this additional strain to reveal any deficits in criterion placement. The current experiment is similar to an experiment in Criss (2006) who reported higher FARs for similar than for unrelated foils and a lower FAR for strong than weak lists for both foil types. In Experiment 2, we expected younger adults to replicate the Criss (2006) results. If older adults do not have flexible control over their criterion placement, then they should show a distorted strength-based mirror effect. If older and younger adults have similar flexibility in criterion placement, then the magnitude of the strength-based mirror effect should not differ with age.

Method

Participants

The same participants completed both Experiment 1 and Experiment 2. However, only one younger adult was excluded from Experiment 2 for chance performance, leaving a total of 38 younger participants.

Stimulus materials

The word pool consisted of 100 pairs of rhyming words that differed by a single letter. The two words were of equal length and were approximately equal in normative word frequency. The average log HAL frequency for the first item of each pair was $M = 8.23$ ($SD = 2.39$) and for the second item was $M = 8.67$ ($SD = 2.02$), where the designation of first or second was arbitrary (Balota et al., 2007; Lund & Burgess, 1996). The difference between the log HAL frequency for each pair of items was computed, and the average difference across all pairs was $M = 0.79$. Whether the first set of items served as the target and the second served as the similar foil or vice versa was randomly selected for each participant.

Design

The experiment was a 2 (foils: similar and unrelated) × 2 (target list strength: strong and weak) × 2 (age: younger and older) mixed design with age serving as the between-subjects factor. Each participant received a weak and a strong block with block order randomly chosen for each individual. The weak study block consisted of 40 words studied with no specific encoding instructions for 2.5 s followed by a 500-ms ISI separating items. The strong study block was simply the weak list repeated for a total of four presentations with study–test lag equated in the same manner as in Experiment 1. A 45-s arithmetic distractor task separated each study list from the corresponding test, and the transition between the two study–test blocks was self-paced. Each test list consisted of 40 items, half targets and half foils. The nature of the foils was manipulated such that half were similar to a target, and half were not, where similarity was defined as rhyming with and being one

Downloaded by [amy criss] at 13:41 14 March 2014

letter different from a target word (e.g., COAT, BOAT). The unrelated foils were selected from the remaining items in the pool that were not assigned to the study list for the individual participant. Either a target or the corresponding similar foil was tested, but not both. The test was a self-paced yes/no single item recognition task.

Results and discussion

Data for all dependent measures are plotted in Figure 2. A 2 (foil type: unrelated and similar) × 2 (list strength: strong and weak) × 2 (age: older and younger) mixed ANOVA was conducted on d'. Discrimination was better for strong than for weak lists, $F(1, 57) = 98.79$, $p < .001$, $\eta^2 = .634$, and for unrelated than for similar foils, $F(1, 57) = 21.99$, $p < .001$, $\eta^2 = .278$, but there was no age-related difference. We used this design because older adults are more susceptible to false-memory effects, but we found no age-related decrement in discrimination for similar items. The similarity manipulation used here is related to the spelling and sound of the word rather than semantics, which is the more typical realm for measuring false memory. We tentatively suggest that the high susceptibility of false memory for older adults is restricted to meaning-based manipulations. There was a marginal triple interaction between foil type, list strength, and age, $F(1, 57) = 4.04$, $p = .049$, $\eta^2 = .066$, where there appears to be no (or a smaller) difference in d' for older and younger adults in the weak similar condition than in the other three conditions, which show a small benefit for older adults. Neither the effect size nor the p-value inspires confidence in this interaction. Even if the result is replicable, the nature of the interaction is such that it is removable with a transformation and therefore should be interpreted with extreme caution, if at all (see Loftus, 1978; Wagenmakers, Krypotos, Criss, & Iverson, 2012). Nevertheless, we speculate that the interaction could be attributed to older adults having more difficulty rejecting the weak-similar foils than the other three foil types. We initially hypothesized that older adults would more susceptible to false memory for similar foils. Additionally, it is likely that the influence of similarity is greater when information is poorly encoded in memory. As such the decrease in older adults' discriminability for similar foils may simply reflect their susceptibility to the similarity manipulation. There were no other interactions (Fs < 1.41, ps $> .24$)

As expected, repetition resulted in a strength-based mirror effect. Strong HRs exceeded weak HRs, $F(1, 57) = 74.01$, $p < .001$, $\eta^2 = .565$, and strong FARs were lower than weak FARs, $F(1, 57) = 15.96$, $p < .001$, $\eta^2 = .219$. FARs were higher for similar than for unrelated foils, $F(1, 57) = 21.15$, $p < .001$, $\eta^2 = .271$. No other main effects or interactions were significant. Note also that these results fully replicate Criss (2006) and extend those same findings to older adults.

We compute the value of c based on targets and unrelated foils. Older adults were marginally more cautious in their responding, $F(1, 57) = 3.26$, $p = .076$, $\eta^2 = .054$, and this did not interact with list strength, $F(1, 57) = 1.97$, $p = .17$, $\eta^2 = .03$.

Like Experiment 1, the data indicate that older adults are slightly more cautious in the placement of their criterion but adjust the criterion with equal flexibility. Manipulating similarity of the foils did not affect the flexibility with which older adults adjust their criterion in response to list strength.

GENERAL DISCUSSION

Age-related strength-based mirror effects were examined in two experiments. In the first experiment, both word frequency and strength were manipulated. We observed a word frequency mirror effect and a SBME. In the second experiment, similarity of the foils was manipulated, and we observed a strength-based mirror effect for both similar and unrelated foils. In both experiments, older adults set a more conservative criterion, preferring the new response, more than the younger adults as evidenced in both in FARs and in response bias. However, this did not interact with the strength-based mirror effect in either experiment. Older adults showed strength-based mirror effects comparable in size to those for younger adults, indicating that older adults can

Downloaded by [amy criss] at 13:41 14 March 2014

Downloaded by [amy criss] at 13:41 14 March 2014

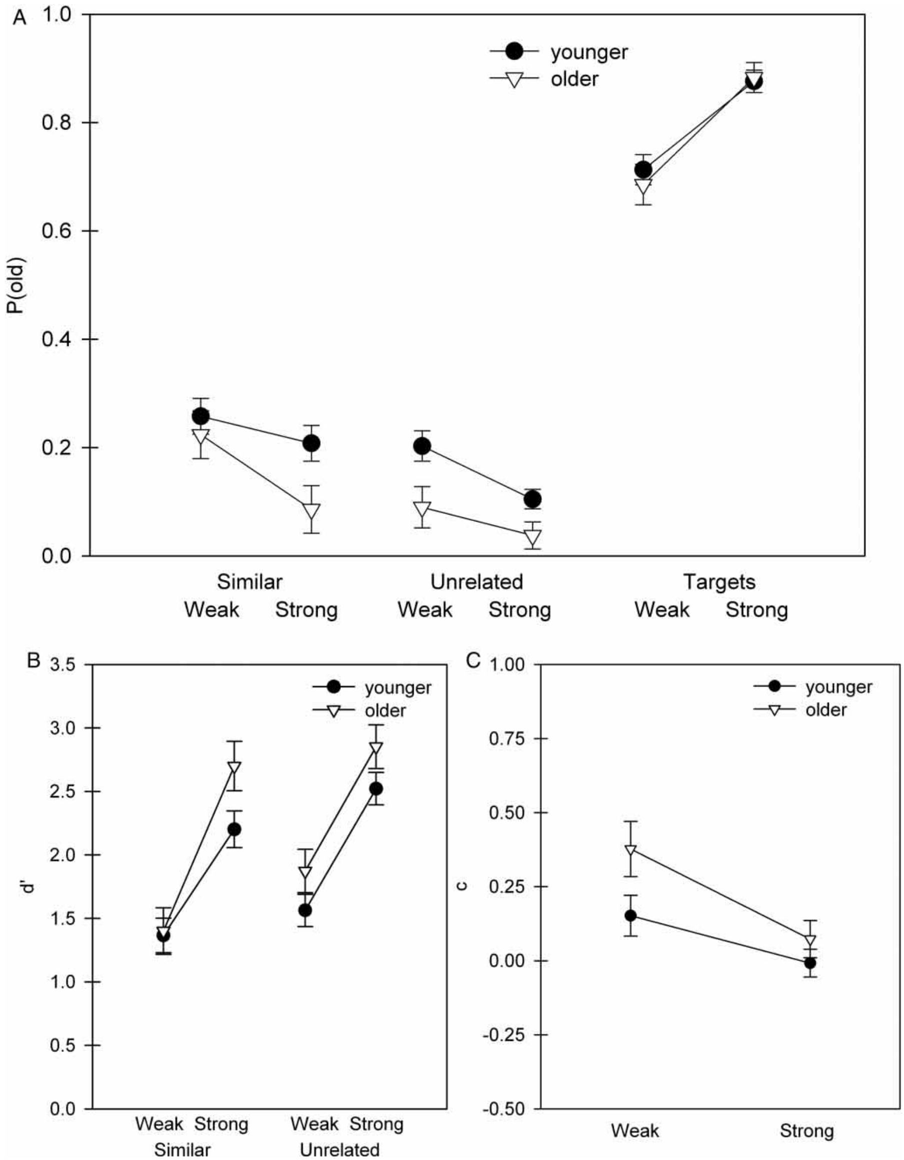


Figure 2. *Strength-based mirror effects were obtained for both for older and younger adults and for similar and unrelated words foils. Panel A: P(old) is the probability of endorsing an item as studied. Hits (top two lines) are higher, and false alarms (bottom four lines) are lower for strong lists than for weak lists. Panel B: Discrimination (d′) was higher for strong lists. Panel C: Response bias (c) was more conservative and farther from optimal for older adults than for younger adults. Error bars are ±1 standard error.*

flexibly place the criterion in response to the demands of the task.

We find evidence that older adults are more conservative in their criterion setting than younger adults but they are equally flexible in adjusting the criterion. As Burke and Light (1981) note, ageing effects in forced-choice recognition tasks make criterion differences less interesting with regard to explanations of memory and ageing. As such, theories of ageing tend to

Downloaded by [amy criss] at 13:41 14 March 2014

be indifferent or simply do not address changes in criterion in older adults. However, if we infer possible expectations from different theories of ageing both a conservative and liberal criterion seem plausible depending on the theoretical perspective. For example, an older adult might adopt a more liberal criterion if they are concerned that their encoding of the list was insufficient due to a general processing deficit (e.g., processing-speed theory; Salthouse, 1996) or if they are concerned about insufficient item-to-list context binding (e.g., associative deficit hypothesis; Naveh-Benjamin, 2000) and overcompensate by endorsing any item that evokes minimal memory strength (e.g., Touron & Hertzog, 2004). Conversely, if the nature of the ageing deficit involves less efficient retrieval (e.g., Burke & Light, 1981) then older adults might adopt a more cautious strategy to avoid false alarms. In the current experiments, there was no age-related deficit in discrimination. That, combined with the fact that judgements of learning improve with ageing (e.g., Daniels, Toth, & Hertzog, 2009; Hertzog, Sinclair, & Dunlosky, 2010), suggests that the above explanations are incomplete or unsatisfying. If older adults are aware that their recognition memory performance was not impaired, then they should not adopt a different criterion. A final possibility is simply that older adults have many more items in episodic memory and more variability in those items by virtue of having several additional years of life spanning a longer range. The larger number of memory traces combined with the longer lag associated with those memories results in a more variable distribution of episodic memory traces than younger adults. In order to make a decision in a recognition memory task, a criterion must be set that discriminates between a specific episode (e.g., the list) and all of the remaining episodes of one's life. If the foil distribution is more variable for older adults, then adopting a more conservative criterion is warranted. The hypothesis that the foil distribution is more variable for older adults is an excellent avenue for future research that would advance theories of memory and ageing.

We used SDT to measure criterion placement and discrimination. As we discussed, SDT brings a host of assumptions that are sometimes treated as claims about the processes underlying memory. Our strong preference is to treat SDT as a measurement device, similar to using an ANOVA or percentage correct and not as a model of memory. There exist a number of successful models of memory that do explain, in great detail, encoding and retrieval mechanisms. However, we offer two possibilities for how to resolve our hypothesis that older adults have higher variability in the foil distribution than younger adults with the assumptions of SDT. If one assumes equal variance in the target and foil distributions (as we did in our statistical analysis), then SDT must also assume that the target distribution for older adults is also more variable (due to the similar level of discrimination across groups). This seems like a plausible assumption, especially if variability is characterized by additional noise in the encoding and retrieval processes or if variability is simply the result of a general slowing of the cognitive system (e.g., Salthouse, 1996). If, on the other hand, one adopts unequal variance between the target and foil distributions as the preferred model (as noted in Footnote 1, we conducted this analysis as well and replicate all findings reported in this manuscript), then the foil distribution may be more variable for older adults than for younger adults without requiring a more variable target distribution. To our knowledge, there are no data that speak to the age-related differences in the variability of foil or target distributions, so the model that best describes ageing and human memory awaits further research.

We discussed two types of criterion flexibility—one that determines whether to respond "studied" or "not studied" and the other that determines when to give a response. The latter has been studied primarily within the context of the diffusion model, and the consistent finding is that older adults are more cautious than optimal—that is, they gather more evidence, thereby taking longer to respond than is necessary. Our data speak only to the former and suggest that older adults are equally flexible at adjusting their response

Downloaded by [amy criss] at 13:41 14 March 2014

threshold to accommodate the demands of the experimental situation, consistent with Pendergrass et al. (2012).

Our data indicate an intact strength-based mirror effect for older adults. Benjamin (2001) found a more complicated SBME pattern for older adults when similarity was manipulated by using two levels of semantic similarity. Benjamin had younger and older adults study items from semantically related word lists (i.e., lists from Deese, 1959; Roediger & McDermott, 1995). Half of the lists were weakly encoded, and half were strongly encoded. At test, participants were given two types of foils: critical lures (e.g., the category label) and categorical distractors (unstudied words from the category list). The results showed an expected increase in HRs for strengthened words in both age groups. The FARs for the critical lures decreased with list strength for younger adults, demonstrating a strength-based mirror effect. However, the opposite pattern was observed for older participants; FARs to critical lures increased with list strength. The FARs for the categorical distractors did not differ as a function of strength for either younger or older adults. The different patterns of data reported here and by Benjamin (2001) are not surprising when considering the differences in methodology. Critically, Benjamin presented all study lists (weak and strong) prior to presenting any test lists—a sort of hybrid between a typical SBME paradigm and a mixed-strength design. A SBME is only predicted for lists of pure strength, not mixed strength, on the assumption that the criterion for a given test list is fixed and does not change trial to trial (i.e., Stretch & Wixted, 1998). In addition, the similarity manipulation was quite different for the two studies. Our lists were structured such that a similar foil shared orthography and phonology with a single studied item whereas Benjamin used foils that were semantically similar to every item on the study list. We propose that the similarity manipulation falls on a continuum (unrelated foils at one extreme and critical lures toward the other) and suggest that a comprehensive study of list strength effects across levels of similarity would demonstrate both patterns of data. Specifically, Criss (2006) presented simulations demonstrating an interaction between strength and similarity. Foils that are less than approximately 50% similar to targets showed a typical SBME pattern: the FAR decreased with the strength of encoding of the target list. Foils that were more than approximately 50% similar to a target showed the opposite pattern: FARs increased with increased encoding of the study list. The exact predictions are determined by the mechanisms of the model, but the general pattern can be understood by considering what it means to be a target. Memory is imperfect—there are errors made during encoding and retrieval. The memory system takes this imperfection into account and does not expect a perfect match between the contents of episodic memory and the cue presented for a recognition task. For these reasons, an item simply needs to match "well enough" to be judged as a target item. Foils that share over 50% of their features with a target item will behave as if they are (poorly encoded) targets. FARs to these items will be high, and the more similar the foil, the higher the FAR. In addition, manipulations that affect targets will similarly affect the FAR for highly similar foils.

Differentiation

Models of the SBME described earlier present a metacognitive account where the criterion placement is determined by the strength of the study list. An alternative account based on encoding and retrieval processes is presented by differentiation models (Criss & McClelland 2006; McClelland & Chappell 1998; Shiffrin, Ratcliff, & Clark, 1990; Shiffrin & Steyvers 1997). Differentiation models propose that the SBME results from changes in the memory strength distributions rather than changes in the criterion. Specifically, the target distribution increases, and the foil distribution decreases for a strong study list compared to a weak study list. Differentiation models assume that better encoding of target items results in more accurate memory traces for those targets. The more accurate a given memory trace, the more likely that it will match its corresponding target (increasing the hit rate). In addition, the more accurate a given memory trace, the less likely it will match a foil (decreasing the false-alarm rate). In other words, the more that is known about a set of items (targets in this case), the less confusable they are with *other* items. Together these assumptions cause the distribution of subjective memory strength to increase for targets and simultaneously decrease for foils following a strong list compared to a weak list (see Criss, 2006, 2009, 2010). Differentiation models, like all memory models, must assume a criterion in order to determine which items are called old and which are called new; however, the criterion does not need to change with list strength in these models to predict a SBME. The interpretation of differential criterion placement for older and younger adults holds under a differentiation account. The difference is in the interpretation of criterion flexibility. If differentiation provides the causal explanation for the SBME then the data presented here should not be interpreted as a change in criterion flexibility. A differentiation account of this data suggests that the fundamental processes underlying encoding and retrieval do not differ for older and younger adults, though the exact parameters contributing to those processes (e.g., accuracy of encoding or location of decision criterion) may differ with age.

SUMMARY

Our goal was to evaluate response bias in older adults using a strength-based mirror effect paradigm. Older adults were more conservative than younger adults, but not differentially so across word frequency, similarity, or list strength. We found no evidence that older adults are less flexible in their criterion placement than younger adults. Change in response bias across the lifespan is underinvestigated. Here we provide evidence that changes in response bias are not caused by a decrease in the ability to flexibly adjust the criterion in response to task demands.

Original manuscript received 9 April 2013
Accepted revision received 10 October 2013
First published online 4 March 2014

REFERENCES

Balota, D. A., Burgess, G. C., Cortese, M. J., & Adams, D. R. (2002). The word-frequency mirror effect in young, old, and early-stage Alzheimer's disease: Evidence for two processes in episodic recognition performance. *Journal of Memory and Language*, *46*(1), 199–226.

Balota, D. A., Cortese, M. J., Duchek, J. M., Adams, D., Roediger, H. L., III, McDermott, K. B., & Yerys, B. E. (1999). Veridical and false memories in healthy older adults and in dementia of the Alzheimer's type. *Cognitive Neuropsychology*, *16*, 361–384.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., … Treiman, R. (2007). The English Lexicon project. *Behavior Research Methods*, *39*(3), 445–459.

Baron, A., & Le Breck, D. B. (1987). Are older adults generally more conservative? Some negative evidence from signal detection analyses of recognition memory and sensory performance. *Experimental Aging Research*, *13*(3), 163–165.

Baron, A., & Surdy, T. H. (1990). Recognition memory in older adults: Adjustment to changing contingencies. *Journal of the Experimental Analysis of Behavior*, *54*(3), 201–212.

Benjamin, A. S. (2001). On the dual effects of repetition on false recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(4), 941–947.

Blessed, G., Tomlinson, B. E., & Roth, M. (1968). The association between quantitative measures of dementia and of senile change in the cerebral grey matter of elderly subjects. *The British Journal of Psychiatry*, *114*, 797–811.

Buchler, N. E. G., & Reder, L. M. (2007). Modeling age-related memory deficits: A two-parameter solution. *Psychology and Aging*, *22*(1), 104–121.

Burke, D. M., & Light, L. L. (1981). Memory and aging: The role of retrieval processes. *Psychological Bulletin*, *90*(3), 513–546.

Cary, M., & Reder, L. M. (2003). A dual-process account of the list-length and strength-based mirror effects in recognition. *Journal of Memory and Language*, *49*(2), 231–248.

Downloaded by [amy criss] at 13:41 14 March 2014

Downloaded by [amy criss] at 13:41 14 March 2014

Cerella, J. (1985). Information processing rates in the elderly. *Psychological Bulletin*, *98*(1), 67–83.

Craik, F. I. M., & McDowd, J. M. (1987). Age differences in recall and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 474–479. doi:10.1037/0278-7393.13.3.474.

Criss, A. H. (2006). The consequences of differentiation in episodic memory: Similarity and the strength based mirror effect. *Journal of Memory and Language*, *55*(4), 461–478.

Criss, A. H. (2009). The distribution of subjective memory strength: List strength and response bias. *Cognitive Psychology*, *59*(4), 297–319.

Criss, A. H. (2010). Differentiation and response bias in episodic memory: Evidence from reaction time distributions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(2), 484–499.

Criss, A., & McClelland, J. L. (2006). Differentiating the differentiation models: A comparison of the retrieving effectively from memory model (REM) and the subjective likelihood model (SLiM). *Journal of Memory and Language*, *55*, 447–460.

Daniels, K. A., Toth, J. P., & Hertzog, C. (2009). Aging and recollection in the accuracy of judgments of learning. *Psychology And Aging*, *24*(2), 494–500.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, *58*, 17–22.

Glanzer, M., & Adams, J. K. (1985). The mirror effect in recognition memory. *Memory & Cognition*, *13*, 8–20.

Glanzer, M., & Bowles, N. (1976). Analysis of the word-frequency effect in recognition memory. *Journal of Experimental Psychology: Human Learning And Memory*, *2*(1), 21–31. doi:10.1037/0278-7393.2.1.21.

Green, D. M., & Swets, J. A. (1966). *Signal Detection Theory and Psychophysics.* New York, NY: Wiley.

Healy, A. F., & Kubovy, M. (1978). The effects of payoffs and prior probabilities on cutoff location in recognition memory. *Memory & Cognition*, *6*, 544–553.

Hertzog, C., Sinclair, S. M., & Dunlosky, J. (2010). Age differences in the monitoring of learning: Cross-sectional evidence of spared resolution across the adult life span. *Developmental Psychology*, *46*(4), 939–948.

Hirshman, E. (1995). Decision processes in recognition memory: Criterion shifts and the list-strength paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(2), 302–313.

Hockley, W. E., & Niewiadomski, M. W. (2007). Strength-based mirror effects in item and associative recognition: Evidence for within-list criterion changes. *Memory & Cognition*, *35*(4), 679–688.

Huh, T. J., Kramer, J. H., Gazzaley, A., & Delis, D. C. (2006). Response bias and aging on a recognition memory task. *Journal of the International Neuropsychological Society*, *12*(1), 1–7.

Isingrini, M., Fontaine, R., Taconnat, L., & Duportal, A. (1995). Aging and encoding in memory: False alarms and decision criteria in a word-pair recognition task. *The International Journal of Aging & Human Development*, *41*(1), 79–88.

Jacoby, L. L.1 (1999). Deceiving the elderly: Effects of accessibility bias in cued-recall performance. *Cognitive Neuropsychology*, *16*, 417–436.

Kılıç, A. (2013). Output interference and strength based mirror effect in recognition memory. *Dissertation Abstracts International*, *74*, 2-B(E).

Koutstaal, W., Schacter, D. L., Galluccio, L., & Stofer, K. A. (1999). Reducing gist-based false recognition in older adults: Encoding and retrieval manipulations. *Psychology and Aging*, *14*(2), 220–237.

Light, L. L. (1991). Memory and aging: Four hypothesis in search of data. *Annual Review of Psychology*, *42*, 333–376.

Loftus, G. R. (1978). On interpretation of interactions. *Memory and Cognition*, *6*, 312–319.

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments & Computers*, *28*(2), 203–208.

McClelland, J. L., & Chappell, M. (1998). Familiarity breeds differentiation: A subjective likelihood approach to the effects of experience in recognition memory. *Psychological Review*, *105*(4), 724–760.

McCormack, P. D. (1984). Aging and recognition memory: Methodological and interpretive problems. *Experimental Aging Research*, *10*(4), 215–219.

Mickes, L., Wixted, J. T., Wais, P. E. (2007). A direct test of the unequal-variance signal detection model of recognition memory. *Psychonomic Bulletin & Review*, *14*, 858–865.

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.

Parks, T. E. (1966). Signal-detectability theory of recognition-memory performance. *Psychological Review*, *73*, 44–58.

Pendergrass, R., Olfman, D., Schmalstig, M., Seder, K., & Light, L. (2012). Age, criterion flexibility, and associative recognition. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, *67*, 36–42. doi:10.1093/geronb/gbr071.

Poon, L. W., & Fozard, J. L. (1980). Age and word frequency effects in continuous recognition memory. *Journal of Gerontology*, *35*(1), 77–86.

Ratcliff, R., Sheu, C-F, & Gronlund, S. D. (1992). Testing global memory models using ROC curves. *Psychological Review*, *99*, 518–535.

Ratcliff, R., Thapar, A., & McKoon, G. (2004). A diffusion model analysis of the effects of aging on recognition memory. *Journal of Memory and Language*, *50* (4), 408–424.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(4), 803–814.

Rotello, C. M., Macmillan, N. A., Hicks, J. L., & Hautus, M. (2006). Interpreting the effects of response bias on remember-know judgments using signal-detection and threshold models. *Memory & Cognition*, *34*, 1598–1614.

Salthouse, T. (1996). The processing speed theory of adult age differences in cognition. *Psychological Review*, *103*(3), 403–428.

Shiffrin, R. M., Ratcliff, R. & Clark, S. (1990). The list-strength effect: II. Theoretical mechanisms. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 179–195.

Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM — retrieving effectively from memory. *Psychonomic Bulletin and Review*, *4* (2), 145–166.

Shipley, W. (1940). A self-administering scale for measuring intellectual impairment and deterioration. *Journal of Psychology*, *9*, 371–377.

Starns, J. J., & Ratcliff, R. (2010). The effects of aging on the speed–accuracy compromise: Boundary optimality in the diffusion model. *Psychology and Aging*, *25*(2), 377–390.

Starns, J. J., & Ratcliff, R. (2012). Age-related differences in diffusion model boundary optimality with both trial-limited and time-limited tasks. *Psychonomic Bulletin and Review*, *19*, 139–145.

Starns, J. J., White, C. N., & Ratcliff, R. (2010). A direct test of the differentiation mechanism: REM, BCDMEM, and the strength-based mirror effect in recognition memory. *Journal of Memory and Language*, *63*(1), 18–34.

Stretch, V., & Wixted, J. T. (1998). On the difference between strength-based and frequency-based mirror effects in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(6), 1379–1396.

Touron, D. R., & Hertzog, C. (2004). Distinguishing age differences in knowledge, strategy use, and confidence during strategic skill acquisition. *Psychology and Aging*, *19*, 452–466.

Treisman, M., & Williams, T. C. (1984). A theory of criterion setting, with an application to sequential dependencies. *Psychological, Review*, *91*, 68–111.

Turner, B. M., Van Zandt, T., & Brown, S. (2011). A dynamic stimulus-driven model of signal detection. *Psychological Review*, *118*(4), 583–613. doi:10.1037/a0025191.

Verhaeghen, P., Marcoen, A., & Goossens, L. (1993). Facts and fiction about memory aging: A quantitative integration of research findings. *Journals of Gerontology: Psychological Sciences*, *48*, 157–171.

Wagenmakers, E., Krypotos, A., Criss, A. H., & Iverson, G. (2012). On the interpretation of uninterpretable interactions: A survey of the field 32 years after loftus. *Memory & Cognition*, *40*, 145–160.

Downloaded by [amy criss] at 13:41 14 March 2014