Journals of Gerontology: Psychological Sciences
cite as: *J Gerontol B Psychol Sci Soc Sci*, 2015, 1–8
doi:10.1093/geronb/gbv039
Advance Access publication July 29, 2015

Original Research Report

Age Effects in Adaptive Criterion Learning

Brittany S. Cassidy[1,2] and Angela H. Gutchess[2]

[1]Department of Psychological and Brain Sciences, Indiana University, Bloomington. [2]Department of Psychology, Brandeis University, Waltham, Massachusetts.

Correspondence should be addressed to Brittany S. Cassidy, PhD, Department of Psychological and Brain Sciences, Indiana University, 1101 E 10th St., Bloomington, IN 47405. E-mail: bscassid@indiana.edu

Received September 26 2014; Accepted February 24 2015.

Decision Editor: Nicole Anderson, PhD

Abstract

Objective. Although prior work has examined age-related changes to criterion placement and flexibility, no study tested these constructs through a paradigm that employs adaptive feedback to encourage specific criterion changes. The goal of this study was to assess age differences in how young and older adults adapt and shift criteria in recognition memory decisions based on trial-by-trial feedback.

Method. Young and older adults completed an adaptive criterion learning paradigm. Over 3 study/test cycles, a biased feedback technique at test encouraged more liberal or strict responding by false-positive feedback toward false alarms or misses.

Results. Older adults were more conservative than young, even when feedback first encouraged a liberal response bias, and older adults adaptively placed criteria in response to biased feedback, much like young adults. After first being encouraged to respond conservatively, older adults shifted criteria less than young when feedback encouraged more lenient responding.

Discussion. These findings evidence labile adaptive criteria placement and criteria shifting with age. However, age-related tendencies toward conservative response biases may limit the extent to which criteria can be shifted in a lenient direction.

Key words: Response bias—Adaptive criterion learning—Aging—Memory

Receiving benefits in many situations, from remembering to whom you have introduced yourself at a conference leading to networking success, to knowing you have already seen a movie while perusing Netflix, would be impossible without recognition memory. The myriad situations utilizing recognition memory (i.e., distinguishing old from new information) have elicited much research, including work illustrating age-related memory deficits (for reviews, see McDaniel, Einstein, & Jacoby, 2008; Zacks, Hasher, & Li, 2000). Recognition has two components: (a) retrieval, in which people generate information from memory based on cues, and (b) decision-making, in which people choose how to act on information (Yonelinas, 2002; Yonelinas & Parks, 2007). Although much aging work has evaluated accuracy, less has considered decision-making.

Signal detection theory (SDT) provides a framework to study recognition. Decisions arise based on where memory strength lies along two distributions representing targets and lures relative to a decision criterion. If strength exceeds a criterion, items will be called "old." If not, items will be called "new." SDT helps assess discriminability (i.e., distinguishing targets from lures; accuracy) and response bias (i.e., the tendency to say "old" or "new;" criterion placement). Evidence of age-related changes to criterion placement has been inconsistent, showing age-invariance (Deason, Hussey, Ally, & Budson, 2012; Isingrini, Fontaine, Taconnat, & Duportal, 1995), and both a tendency to become more liberal (Bastin & Van der Linden, 2003; Flicker, Ferris, Crook, & Bartus, 1990; Kapucu, Rotello, Ready, & Seidl, 2008), especially when more lures than targets are present (Huh, Kramer, Gazzaley, & Delis,

2006) and conservative (Criss, Aue, & Kilic, in 2014; Ferris, Crook, Clark, McCarthy, & Rae, 1980; Vakil, Mosak, & Ashkenazi, 2003).

Beyond general placement, other work has evaluated how *flexibly* older adults shift criteria based on experience (i.e., the ability to move criterion given different situations). For instance, when a difficult test on word pairs follows an easy one, both young and older adults shift criteria to become more conservative (Pendergrass, Olfman, Schmalstig, Seder, & Light, 2012). Related work reports similar findings of intact flexibility, although older adults tend to be more conservative (Criss et al., 2014). Other work suggests relative inflexibility, such that older adults adjust criteria to a lesser extent than young to maximize payoffs (Baron & Surdy, 1990). Work showing cautious responding among older adults despite no age differences in memory strength also suggests inflexibility in that older adults were unable to adjust criteria based on task instructions (Starns & Ratcliff, 2010). Critically, these studies investigated placement and flexibility through reaction to different task demands. No study has assessed age effects in criterion placement and flexibility based on *trial-to-trial* feedback. This is important, as many situations involve adapting criterion based on feedback to optimize the likelihood of success. To this end, we explored age differences in criterion placement and flexibility using an adaptive criterion learning paradigm.

We used a biased feedback technique effective in inducing criterion learning (Han & Dobbins, 2008). People become more lax given false-positive feedback for false alarms and stricter given false-positive feedback for misses (Han, 2009; Han & Dobbins, 2008, 2009). Limiting false-positive feedback to errors leaves participants relatively unaware of the manipulation (see Experiment 3; Han & Dobbins, 2008). We anticipated that there would be age differences in criteria setting (Hypothesis 1), based on related findings in the literature, as no studies have assessed how age-related changes in placement might change after *feedback*. One possibility is that potential gains (positive feedback) and losses (negative feedback) might cause older adults to respond in a conservative manner, given that aging is associated with less risk tolerance (Deakin, Aitken, Robbins, & Sahakian, 2004). However, because reduced memory accuracy with age broadly results in more false alarms (e.g., Rosa & Gutchess, 2013), older adults might respond more liberally than young despite feedback. We anticipated modified criterion placement (Hypothesis 2) and flexible shift (Hypothesis 3) across age groups. However, we also believed that differences in criterion placement with age might limit the extent to which criteria could be shifted. Given that many studies of aging and recognition memory do not analyze decision criteria, systematically assessing criteria through an adaptive learning paradigm may inform our theoretical understanding of how older adults adopt decision-making strategies.

We also expected that trial-to-trial feedback could produce age differences in criterion shifting. Providing false-positive feedback means participants would view more positive than negative feedback overall if they shifted criteria to optimize the presence of positive feedback when making errors (i.e., when they weren't sure whether an item was new or old). Similar to reinforcement learning (White & Wixted, 1999), this may be particularly salient for older adults, who may feel increased self-efficacy with more positive feedback (Karl, O'Leary-Kelly, & Martocchio, 1993; Maurer, 2001). Older adults also could have an increased criterion shift to optimize positive feedback because they are more motivated to avoid negative outcomes than young (Frank & Kong, 2008). Under optimal feedback conditions, older adults have higher motivation and goal commitment than young (West, Bagwell, & Dark-Freudman, 2005). Thus, we anticipated older adults might shift criteria *more* than young (Hypothesis 3b).

Although not our primary aim, we also explored a potential relationship between working memory and decision criterion. Individual variability underlies criterion placement (Aminoff et al., 2012; Han, 2009; Kantner & Linsday, 2012). Healthy aging is related to decreases in risk tolerance (Deakin et al., 2004). Although some work shows similarly cautious behavior with age (Peters, Hess, Vastfjall, & Auman, 2007), older adults are particularly risk averse when considering potential gains (Lauriola & Levin, 2001; Weller, Levin, & Denburg, 2011) (e.g., the reward from receiving positive feedback). Given that lower cognitive ability is broadly associated with greater risk aversion (Boyle, Yu, Buchman, Laibson, & Bennett, 2011; Dohmen, Falk, Huffman, & Sunde, 2009) older adults with lower working memory capacity could exhibit more conservative responding. However, limited executive function (i.e., in Alzheimer's patients) is also associated with increased liberal responding (e.g., Snodgrass & Corwin, 1988), and liberal response biases have been tied to the frontal aging hypothesis (Huh et al., 2006). This alternatively predicts an association between working memory and lenient responding. We used the letter-number sequencing task (Wechsler, 1997), a working memory task with an executive function component (Crowe, 2000), to explore a potential relationship between working memory capacity and more conservative or lenient responding (Hypothesis 4).

Method

Participants

Twenty-four older ($M = 78.42$, $SD = 5.99$) and 24 young adults (Note that these young adults also comprised the nonsocial source group in related work on adaptive criterion learning (Cassidy, Dube, & Gutchess, 2014).) ($M = 18.46$, $SD = 0.83$) from Brandeis University and the community participated. Older adults had Mini-Mental State Examination (Folstein, Folstein, & McHugh, 1975) scores > 26 ($M = 29.29$, $SD = 1.16$). Older adults had more years of education ($M_{young} = 12.40$, $SD = 0.86$; $M_{older} = 16.75$, $SD = 2.27$; $t(46) = 8.79$, $p < .001$) and higher vocabulary (Shipley, 1986) scores ($M_{young} = 32.41$, $SD = 3.60$; $M_{older} = 37.50$, $SD = 2.23$; $t(46) = 5.89$, $p < .001$) than young. Young adults had faster processing speed ($M = 83.25$, $SD = 11.10$) than older adults ($M = 58.00$, $SD = 10.40$), $t(46) = 8.13$, $p < .001$, as measured by digit-comparison (Hedden et al., 2002), and higher letter-number sequencing scores ($M_{young} = 11.54$, $SD = 2.04$; $M_{older} = 8.75$, $SD = 2.51$; $t(46) = 4.23$, $p < .001$).

Materials

Three lists of 200 nouns each (100 assigned as targets and 100 as lures) were drawn from the English Lexicon Project (Balota et al., 2007) for three study/test cycles. Two versions counterbalanced targets and lures. No significant effects emerged when comparing the number of letters, syllables, and phonemes, and Kucera–Francis frequency (Kucera & Francis, 1967) of the words in five 3 (Cycles: 1, 2, 3) × 2 (Assignment: Target, Lure) ANOVAs, $ps > .22$. On average, nouns had 7.11 letters ($SD = 1.68$), 6.02 phonemes ($SD = 1.64$), and 2.40 syllables ($SD = 0.81$), with a Kucera–Francis frequency of 12.26 ($SD = 7.00$).

Procedure

We replicated the study/test procedure of Han and Dobbins (2008), which has been previously employed in young adults, but never in comparison to older adults. Stimuli were presented via E-Prime (Psychology Software Tools, Pittsburgh, PA). At study, words appeared one at a time for 2000 ms. Below each word, the question, "How many syllables?" appeared, with choices of "1," "2," "3," and "4+" listed below. Participants pressed 1, 2, 3, or 4 to indicate how many syllables were in each word. All participants completed 10 study task practice trials, repeating them if necessary, and were aware of the memory tests. After practice, the first of three study/test cycles began. One-hundred randomly presented words were studied, with a blank screen for 500 ms between each trial.

Immediately following, participants began the first self-paced memory test. In all tests, participants viewed 200 randomly presented words (100 targets/100 lures) one at a time. Responding "1" indicated "old" and "2" indicated "new." After old/new decisions, the question "How confident are you?" appeared on the screen with a 3-point scale (1 = unsure; 3 = certain). Note that in the scale, "2" did not have a label, consistent with a prior iteration of this task (Han & Dobbins, 2008). However, the "2" response was defined to participants at the introduction of the memory test as not 100% certain, but not an unsure guess. After each rating, the word "Correct" or "Incorrect" at the center of the screen for 1000 ms indicated old/new accuracy (this was not always accurate feedback, as described below), followed by a blank screen for 250 ms. Because piloting revealed older adults had difficulty utilizing the confidence scale, older adults practiced 10 memory trials. After completion of the experiment (i.e., at the end of the three cycles), participants noted their belief in the feedback (1 = not at all; 7 = very much so).

Biased Feedback Manipulation

Biased feedback induced criterion shifting. Participants received biased feedback in the first and third study/test cycles, and fully correct feedback in the second. The nature of feedback was reversed across two groups (Lax–Neutral–Strict [LNS], Strict–Neutral–Lax [SNL]). LNS participants received positive feedback for false alarm responses in the first test (i.e., receiving "Correct" as feedback for "new" words called "old"), encouraging lax (L) criteria. Other responses were correctly identified. Feedback in the second test was fully correct (N; neutral). LNS participants received positive feedback to miss responses in the third test, (i.e., receiving "Correct" as feedback for "old" words as "new"), encouraging strict (S) criteria. SNL participants saw positive feedback for misses in the first test, fully correct feedback in the second, and positive feedback for false alarms in the third. Note that participants received feedback on every trial during the experiment, and always received "Correct" for fully correct responses. It was only the nature of biased feedback to incorrect responses that varied across the experiment.

Results

Manipulation Check

Young (M = 4.79, SD = 1.86) and older (M = 5.75, SD = 1.48) adults reported moderate belief in our manipulation, with older adults marginally higher, $t(46) = 1.97$, $p = .06$ (Because of above moderate feedback belief, on average, we re-ran our analyses covarying for the extent to which participants believed feedback. None of the reported results were impacted. The success of this manipulation replicates prior work showing that people tend not to suspect biased feedback because it occurs for only errors (Han & Dobbins, 2008). Future work could assess how young and older adults adapt criteria based on whether false feedback is given across all responses of a certain type (e.g., all misses) versus whether it is probabilistically manipulated (e.g., 75% of misses beget "Correct" as feedback) to understand how much biased feedback begets criterion shift with age.).

Signal Detection Measures

To replicate Han and Dobbins (2008), we computed A_z, a measure of accuracy, and c_a, a measure of criterion placement, for our analyses (Macmillan & Creelman, 1991). A_z and c_a were calculated for each participant on each test and entered into ANOVAs, as described below. A_z and c_a were calculated after obtaining the zROC slope estimates from a linear regression. These estimates are appropriate to use (e.g., see Sauvage, Beer, & Eichenbaum, 2010; Trippas, Handley, & Verde, 2014) as they are comparable to $1/\sigma^2_{\text{Target}}$ as obtained from model fitting techniques (Ratcliff, McKoon, & Tindall, 1994; Ratcliff, Sheu, & Gronlund, 1992). It is important to note that using measures derived from the unequal-variance SDT model provide *independent* measures of memory accuracy (A_z) and response bias (c_a) (Macmillan & Creelman, 1991), as response bias is theoretically independent of discriminability (Miller, Handy, Cutler, Inati, & Wolford, 2001; Snodgrass & Corwin, 1988). Accuracy and bias may change across conditions, but our measures are unlikely to be influenced by this or result in artifact (Dube & Rotello, 2012; Green & Swets, 1966; Pazzaglia, Dube, & Rotello, 2013).

Decision Criteria (c_a)

Placement

To assess differences in criterion *placement,* we used c_a as the dependent variable in a 2 (Age Group: Young, Older) × 2 (Order: LNS, SNL) × 3 (Test: 1, 2, 3) ANOVA. See Table 1 for complete descriptives. A main effect of Test, $F(2, 88) = 3.40$, $p = .04$, $\eta_p^2 = 0.07$, indicated that participants were stricter in Test 1 (M = .23, SD = 0.32) than 2 (M = 0.11, SD = 0.26), $F(1, 44) = 10.64$, $p = .002$, $\eta_p^2 = 0.20$. There were no differences between Tests 1 or 2 with 3 (Test 3: M = 0.15, SD = 0.27; Test 1: $F(1, 44) = 1.96$, $p = 0.17$); Test 2: $F < 1$, $p = 0.39$).

Supporting Hypothesis 1, a main effect of Age Group emerged such that older adults (M = 0.24, SD = 0.21) were stricter than young

Table 1. Mean (Standard deviation) Decision Criteria and Accuracy Estimates Across Age Group, Order, and Test

	Lax–Neutral–Strict			Strict–Neutral–Lax		
	Test 1	Test 2	Test 3	Test 1	Test 2	Test 3
Young adults						
c_a	−0.08 (0.31)	−0.14 (0.22)	0.18 (0.28)	0.39 (0.35)	0.27 (0.27)	−0.08 (0.19)
A_z	0.86 (0.07)	0.86 (0.06)	0.76 (0.13)	0.76 (0.14)	0.79 (0.10)	0.78 (0.11)
Hit	0.81 (0.09)	0.83 (0.08)	0.65 (0.19)	0.57 (0.16)	0.62 (0.20)	0.74 (0.12)
False alarm	0.24 (0.15)	0.26 (0.17)	0.20 (0.07)	0.11 (0.10)	0.19 (0.08)	0.31 (0.11)
Older adults						
c_a	0.26 (0.34)	0.19 (0.36)	0.41 (0.36)	0.36 (0.28)	0.13 (0.11)	0.07 (0.21)
A_z	0.73 (0.07)	0.78 (0.10)	0.70 (0.11)	0.73 (0.07)	0.73 (0.09)	0.69 (0.11)
Hit	0.59 (0.14)	0.62 (0.15)	0.50 (0.17)	0.55 (0.16)	0.63 (0.09)	0.62 (0.11)
False alarm	0.20 (0.12)	0.23 (0.14)	0.18 (0.09)	0.17 (0.09)	0.27 (0.08)	0.32 (0.11)

(M = 0.09, SD = 0.21), $F(1, 44) = 5.71$, $p = 0.02$, $\eta_p^2 = 0.12$. An interaction between Age Group and Order qualified the effect of Age Group, $F(1, 44) = 5.45$, $p = 0.02$, $\eta_p^2 = 0.13$. Older LNS participants (M = 0.29, SD = 0.21) were stricter than young LNS participants (M = −0.01, SD = 0.21), $F(1, 44) = 12.15$, $p = .001$, $\eta_p^2 = 0.22$. No differences in criterion existed for the SNL participants ($M_{YA} = 0.20$, SD = 0.21; $M_{OA} = 0.19$, SD = 0.21; $F < 1$, $p = .91$).

Replicating Han and Dobbins' (2008) findings across age groups (Hypothesis 2), an Order × Test interaction emerged, $F(2, 88) = 22.55$, $p < .001$, $\eta_p^2 = 0.34$. SNL participants were stricter than LNS in tests 1 ($M_{SNL} = 0.37$, $SD_{SNL} = 0.32$; $M_{LNS} = 0.09$, $SD_{LNS} = 0.32$; $F(1, 44) = 9.33$, $p = .004$, $\eta_p^2 = 0.18$) and 2 ($M_{SNL} = 0.20$, $SD_{SNL} = 0.25$; $M_{LNS} = 0.02$, $SD_{LNS} = 0.25$; $F(1, 44) = 6.06$, $p = .02$, $\eta_p^2 = 0.12$). LNS participants (M = 0.30, SD = 0.27) were stricter than SNL (M = −0.002, SD = 0.27) in Test 3, $F(1, 44) = 15.60$, $p < .001$, $\eta_p^2 = 0.26$. No other effects were significant (A separate question regards whether developing a criterion level was the same for the LNS and SNL groups when manipulated to be lax and strict, given that these manipulations occurred in different orders. A 2 (Age Group: young adults, older adults) × 2 (Order: LNS, SNL) × 3 (Manipulation: Lax, Neutral, Strict) ANOVA using response bias (c_a) as the dependent variable tested this possibility. We report findings that were not redundant with our original analyses. A main effect of Manipulation emerged, $F(2, 88) = 21.59$, $p < .001$, $\eta_p^2 = 0.33$. Criterion when manipulated to be lax (M = 0.05, SD = 0.30) was more lenient than criterion given fully correct feedback (M = 0.11, SD = 0.29), $F(1, 44) = 4.79$, $p = .03$, $\eta_p^2 = 0.10$, and when manipulated to be strict strict (M = 0.34, SD = 0.32), $F(1, 44) = 24.99$, $p < .001$, $\eta_p^2 = 0.36$. Criterion when manipulated to be more conservative was also stricter than criterion given fully correct, $F(1, 44) = 23.98$, $p < .001$, $\eta_p^2 = 0.35$. An interaction between Manipulation and Order also emerged, $F(2, 88) = 4.36$, $p = .02$, $\eta_p^2 = 0.09$ (See Table 1 for descriptive statistics). When manipulated to be lax and when manipulated to be strict, the LNS and SNL groups exhibited similar criterion, $ps > 0.24$. Redundant with our original analyses, given fully correct feedback, the LNS group was more lax than the SNL group, $F(1, 44) = 6.06$, $p = .02$, $\eta_p^2 = 0.12$. No other effects were significant.).

Relative Shift

Our first analysis assessed age differences in criterion placement and if biased feedback encouraged placement changes with age, but did not directly address the extent of flexible criterion *shift*. To assess potential changes, we calculated shift indexes for each participant by subtracting criterion (c_a) from Tests 2 and 3 from the criterion of the preceding test. We entered these indexes in a 2 (Age Group: Young, Older) × 2 (Order: LNS, SNL) × 2 (Shift Index: Test 1–2, Test 2–3) ANOVA. Given directional differences in index based on becoming stricter or liberal, it is most useful to interpret differences from this ANOVA *within* versus *across* Order.

Supporting Hypothesis 3, an interaction between Age Group, Order, and Shift Index emerged, $F(1, 44) = 5.43$, $p = .02$, $\eta_p^2 = 0.11$ (Figure 1), that suggested age-related difficulty in shifting toward lenient criteria. For SNL participants, young (M = 0.35, SD = 0.22) shifted criteria more than older adults (M = 0.06, SD = 0.24) from the second to third test, $F(1, 44) = 6.06$, $p = .02$, $\eta_p^2 = 0.12$, with no difference from the first to second, $F < 1$, $p = .33$. LNS participants had no differences, $ps > 0.41$.

Additionally, a main effect of Order showed that LNS participant shifted criterion in a stricter direction than the SNL participants, $F(1, 44) = 24.99$, $p < .001$, $\eta_p^2 = 0.36$ A main effect of Shift Index showed that overall criteria shifted in a more lenient direction from the first to second test (M = 0.12, SD = 0.25) than the second to third (M = −0.04, SD = 0.29), $F(1, 44) = 8.78$, $p = .005$, $\eta_p^2 = 0.17$. An interaction between Shift Index and Order also emerged, $F(1, 44) = 13.39$, $p = .001$, $\eta_p^2 = 0.23$. LNS participants shifted criteria to be stricter from the second to third test (M = −0.28, SD = 0.29) versus first to second (M = 0.07, SD = 0.25), $F(1, 44) = 21.95$, $p < .001$, $\eta_p^2 = 0.33$. No difference emerged in SNL participants, $F < 1$ $p = .63$.

Working Memory and c_a

We correlated the average decision criterion across tests with letter-number sequencing scores to assess a relationship between executive processing and response bias. Lower working memory capacity was linked to stricter criteria in older adults, $r = -0.51$, $p = .01$ (Hypothesis 4; Figure 2). The association persisted ($r = -0.48$, $p = .03$) when controlling for age and education. No association existed in young, $r = 0.01$, $p = .98$, and no associations between criterion *shift* and working memory emerged for either age group, $ps > 0.14$.

Accuracy (A_z)

Although not our primary aim, we assessed accuracy (A_z) in a 2 (Age Group: Young, Older) × 2 (Order: LNS, SNL) × 3 (Test: 1, 2, 3) ANOVA to provide a more comprehensive picture of performance with age. See Table 1 for complete descriptives. A main effect of Age Group showed that older adults (M = 0.72, SD = 0.08) were less accurate than young (M = 0.80, SD = 0.08), $F(1, 44) = 11.95$, $p = .001$, $\eta_p^2 = 0.21$ (Although our use of measures derived from unequal-variance SDT models means A_z and c_a are independent of each other, decreased accuracy among older relative to young adults suggests that older adults might encounter more instances for biased feedback relative to young. This suggests that differences in

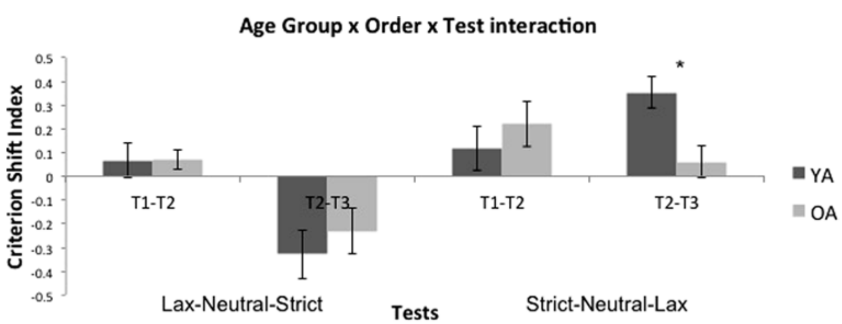

Figure 1. Older adults (OA) in the SNL group shifted criterion less than young adult (YA) from the second to third test (T2–T3). No other age differences emerged. Error bars represent standard error.

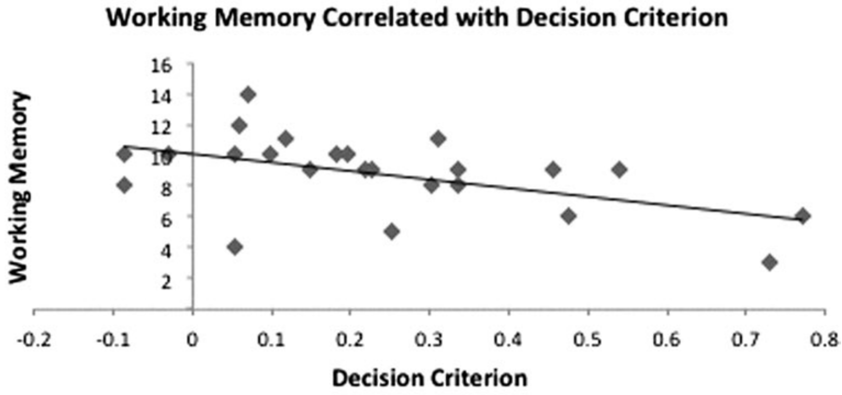

Figure 2. Older adults with lower working memory capacity had stricter decision criteria.

discrimination might in part impact the extent of criterion placement. We tested this possibility by using average accuracy (A_z) across the three tests as a covariate in the discussed 2 (Age Group: Young, Older) × 2 (Order: LNS, SNL) × 3 (Test: 1, 2, 3) ANOVA where c_a was entered as the dependent variable. Covarying for accuracy eliminated the main effect of Age Group ($F(1, 43) = 1.60$, $p = .21$), suggesting that decision criteria were similar with age. Covarying for accuracy did not eliminate the demonstrated interactions between Age Group and Order, and between Order and Test, suggesting that differences in accuracy cannot account for all of the effects demonstrated in the task.). A main effect of Test showed that accuracy decreased over time, $F(2, 88) = 5.62$, $p = .01$, $\eta_p^2 = 0.11$, with Test 3 ($M = 0.73$, $SD = 0.11$) lower than Test 1 ($M = 0.77$, $SD = 0.10$; $F(1, 44) = 5.10$, $p = .03$, $\eta_p^2 = 0.10$) and Test 2 ($M = 0.78$, $SD = 0.08$; $F(1, 44) = 10.80$, $p = .002$, $\eta_p^2 = 0.20$). Tests 1 and 2 did not differ, $p = .36$.

Age Group, Order, and Test marginally interacted, $F(1, 44) = 2.77$, $p = .07$, $\eta_p^2 = 0.06$. For young LNS participants, there was a main effect of Test, $F(2, 22) = 12.52$, $p < .001$, $\eta_p^2 = 0.53$. Test 3 accuracy ($M = 0.76$, $SD = 0.13$) was lower than tests 1 ($M = 0.86$, $SD = 0.07$; $F(1, 11) = 23.25$, $p < .001$, $\eta_p^2 = 0.68$) and 2 ($M = 0.86$, $SD = 0.06$; $F(1, 11) = 11.63$, $p = .01$, $\eta_p^2 = 0.51$). No difference existed between tests 1 and 2, $p = .92$. No order effects emerged for older LNS participants or for younger and older SNL participants, *ps* > .12. No other effects were significant.

Discussion

We assessed age differences in criterion placement and flexibility through a paradigm in which trial-to-trial feedback motivated criterion placement. Older adults were more conservative than young. Consistent with our predictions, both age groups adaptively placed criteria in response to feedback, suggesting an intact ability to implicitly adapt criterion with age. However, older adults displayed less criterion *shift* than young when encouraged to become lenient after first being encouraged to be strict. Finally, our exploratory analysis showed an association between lower working memory capacity and stricter criteria among older adults.

Our findings complement work showing that both young and older adults place decision criteria based on task demands. For instance, both age groups respond conservatively if a difficult memory test follows an easy one (Pendergrass et al., 2012). Likewise, both age groups show strength-based mirror effects (i.e. higher hit rates and lower false alarm rates for strongly vs. weakly encoded items) (Criss et al., 2014). We expand the literature by showing this ability through trial-to-trial adaptation. Adapting and maintaining criteria based on feedback might be implicit and inherently different from responding to more overt placement cues. For instance, explicit instructions to respond leniently may evoke deliberate strategy changes. Unintentional changes may occur if cues to maximize success are not overt. Given that species incapable of explicit strategy use show response learning (White & Wixted, 1999), adaptive criterion learning could represent an interaction between implicit-procedural and explicit memory systems (Han & Dobbins, 2008). This is an intriguing possibility, as it suggests that older adults retain the ability to adaptively change criteria even when unaware of evidence that they should do so.

These findings may prove fruitful for future work because they occurred despite age-related cognitive losses. An interesting direction would be to extend this paradigm to a cognitively impaired group. People with Alzheimer's disease (AD) change criterion placement given explicit instructions (Waring, Chong, Wolk, & Budson, 2008). We might expect similar adaptive placement among patients using the present paradigm, as AD-affected individuals show evidence of implicit learning (Knopman & Nissen, 1987). Such studies could test the limits of how and when different patient populations are able to learn and retain information.

Despite similarly adapting criteria, older adults were stricter than young, complementing past work (Criss et al., in press; Poon & Fozard, 1980). When faced with a memory test, older adults may be particularly aware of age differences, impacting their response style. Indeed, older adults are often less confident than young in their ability to perform memory tests, regardless of their actual accuracy (Touron & Hertzog, 2004). Despite being anchored in an overall conservative response style, the data suggest our manipulation nevertheless caused adaptive criterion placement among older adults. However, not all studies have found older adults to be more conservative than young; in fact, one recent study reports older adults to actually be more liberal. Critically, Huh et al. (2006) compared response bias in education-matched cohorts of young and older adults, whereas our older cohort had more years of education than our young. Because prior work has found stricter response styles limited to highly educated older adults (Marquié & Baracat, 2000), future work using a broader range of education levels among both young and older adults could help to resolve this inconsistency.

Intriguingly, the tendency for older adults to be more conservative than young existed only among those in the LNS group.

Young and older adults in the SNL group were similarly strict overall. This suggests the tendency for older adults to be conservative existed even when we sought to anchor their bias as more liberal. Dovetailing this, while both age groups had adaptive criterion *placement*, older adults had less *shift* than young when feedback cued adaptively lenient, versus strict, responding. Even though younger and older adults adaptively place criterion, anchoring in or shifting toward a liberal response style may be difficult with age as it combats a natural tendency to be stricter. These findings suggest age differences in how an overall relatively strict response style impacts both adaptive criterion *placement* and *shift* in recognition decisions. Although older adults may have labile decision criteria, the extent of age-related lability may be limited by the direction of the desired response pattern.

Finally, our exploratory analyses found lower working memory capacity was associated with *stricter* responding in older adults. This could potentially reflect risk aversion given less cognitive ability (Boyle et al., 2011). Because older adults often more accurately judge their learning than young (Hertzog, Sinclair, & Dunlosky, 2010), negative aging- and memory-related stereotypes could alternatively reduce performance and increase cautious responding (Levy, 2003). Speculatively, those with less executive ability and potentially less prefrontal function, as indicated by lower working memory scores, might be most susceptible to memory-related stereotypes and threat within an evaluation context, and most likely respond cautiously. Indeed, older adults evaluate their recognition memory as poorer than young, and adjust their decision criteria accordingly (Humphries, Flowe, Hall, Williams, & Ryder, 2015). Conditions maximizing threat also lower performance in threatened older adults versus nonthreatened peers and young (Hess, Auman, Colcombe, & Rahhal, 2003) and also induce stricter biases (Barber & Mather, 2013). Future work could test this idea by comparing criteria in this paradigm between high and low functioning older adults receiving primes meant to induce or ameliorate stereotype threat.

A tendency for older adults to respond more conservatively than young and a relationship between conservative responding with lower working memory ability might seem counterintuitive at first blush. Indeed, some extant work suggests that healthy older adults (Huh et al., 2006) in addition to AD patients (Bartok et al., 1997; Deason et al., 2012; Snodgrass & Corwin, 1988), and individuals with frontal (Curran, Schacter, Norman, & Galluccio, 1997) and parietal (Drowos, Berryhill, Andre, & Olson, 2010) lobe damage tend to respond more *liberally* than controls. Much of this research involves tasks that simply assess old-new decisions. Our work critically differs by providing trial-to-trial *feedback*. The evaluative aspect of our memory test could potentially explain the divergence of the present results from this literature. Being evaluated could induce risk aversion in memory decisions among older adults that leads to conservative responding. Liberal responding with age is consistent with the frontal aging hypothesis (Huh et al., 2006), in that differences in frontal lobe function influence relative response bias (Alexander, Stuss, & Fansabedian, 2003). However, the evaluative aspect of our test could negate and potentially *reverse* this effect due to memory-related stereotype threat (Barber & Mather, 2013) or risk aversion (Boyle et al., 2011). To test this possibility, it would be interesting for future work to examine how receiving versus not receiving feedback influences response bias with age.

Compared to studies of discriminability, investigations of age-related changes to decision-making in recognition memory are few. Our results extend the literature on this topic in several ways. They evidence preservation in the adaptive placement of criterion in response to trial-by-trial feedback, but also change in the age-related tendency to adopt a conservative response style and the extent to which older adults shift toward more lax criteria. Our exploratory analysis suggests working memory capacity may be related to stricter responding with age. These findings broaden our understanding of how older adults might learn and respond in countless recognition-based situations.

Funding

This project was funded by the Brandeis University and was approved by the Brandeis University Institutional Review Board.

Acknowledgments

We thank Jenny Crawford, Naomi Weinblatt, Kyra Borenstein, and Tammy Hsieh for research assistance. We also thank Eric Leshikar for his help programming the computer task. The authors declare no conflict of interest.

References

Alexander, M., Stuss, D., & Fansabedian, N. (2003). California verbal learning test: performance by patients with focal frontal and non-frontal lesions. *Brain*, 126, 1493–1503. doi:10.1093/brain/awg128

Aminoff, E., Clewett, D., Freeman, S., Frithsen, A., Tipper, C., Johnson, A., … Miller, M. (2012). Individual differences in shifting decision criterion: A recognition memory study. *Memory & Cognition*, 40, 1016–1030. doi:10.3758/s13421-012-0204-6

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, 445–459. doi:10.3758/BF03193014

Barber, S., & Mather, M. (2013). Stereotype threat can reduce older adults memory errors. *The Quarterly Journal of Experimental Psychology*, 66, 1888–1895. doi:10.1080/17470218.2013.840656

Baron, A., & Surdy, T. (1990). Recognition memory in older adults: adjustment to changing contingencies. *Journal of the Experimental Analysis of Behavior*, 54, 201–212. doi:10.1901/jeab.1990.54-201

Bartok, J., Wilson, C., Giordani, B., Keys, B., Persad, C., Foster, N., & Berent, S. (1997). Varying patterns of verbal recall, recognition, and response bias with progression of Alzheimer's disease. *Aging, Neuropsychology, and Cognition*, 4, 266–272. doi:10.1080/13825589708256651

Bastin, C., & Van der Linden, M. (2003). The contribution of recollection and familiarity to recognition memory: a study of the effects of test format and aging. *Neuropsychology*, 17, 14–24. doi:10.1037/0894-4105.17.1.14

Boyle, P., Yu, L., Buchman, A., Laibson, D., & Bennett, D. (2011). Cognitive function is associated with risk aversion in community-based older persons. *BMC Geriatrics*, 11, 53. doi:10.1186/1471-2318-11-53

Cassidy, B., Dube, C., & Gutchess, A. (2014). Social influences on adaptive criterion learning. *Memory & Cognition*.

Criss, A., Aue, W., & Kilic, A. (2014). Age and response bias: evidence from the strength based mirror effect. *The Quarterly Journal of Experimental Psychology*, 67, 1910–1924. doi:10.1080/17470218.2013.874037

Crowe, S. F. (2000). Does the letter number sequencing task measure anything more than digit span? *Assessment*, 7, 113–117. doi:10.1177/107319110000700202

Curran, T., Schacter, D. L., Norman, K. A., & Galluccio, L. (1997). False recognition after a right frontal lobe infarction: memory for general and specific information. *Neuropsychologia*, 35, 1035–1049. doi:10.1016/S0028-3932(97)00029-8

Deakin, J., Aitken, M., Robbins, T., & Sahakian, B. (2004). Risk taking during decision-making in normal volunteers changes with age. *Journal of the International Neuropsychological Society*, 10, 590–598. doi:10.1017/S1355617704104104

Deason, R., Hussey, E., Ally, B., & Budson, A. (2012). Changes in response bias with different study-test delays: evidence from young adults, older adults, and patients with Alzheimer's disease. *Neuropsychology*, 26, 119–126. doi:10.1037/a0026330

Dohmen, T., Falk, A., Huffman, D., & Sunde, U. (2009). *Are risk aversion and impatience related to cognitive ability?* Bonn, Germany: Institute for the Study of Labor.

Drowos, D., Berryhill, M., Andre, J., & Olson, I. (2010). True memory, false memory, and subjective recollection deficits after focal parietal lobe lesions. *Neuropsychology*, 24, 465–475. doi:10.1037/a0018902.

Dube, C., & Rotello, C. (2012). Binary ROCs in perception and recognition memory are curved. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38, 130–151. doi:10.1037/a0024957

Ferris, S. H., Crook, T., Clark, E., McCarthy, M., & Rae, D. (1980). Facial recognition memory deficits in normal aging and senile dementia. *Journal of Gerontology*, 35, 707–714. doi:10.1093/geronj/35.5.707

Flicker, C., Ferris, S. H., Crook, T., & Bartus, R. T. (1990). Impaired facial recognition memory in aging and dementia. *Alzheimer Disease and Associated Disorders*, 4, 43–54.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12, 189–198. doi:10.1016/0022-3956(75)90026-6

Frank, M., & Kong, L. (2008). Learning to avoid in older age. *Psychology and Aging*, 23, 392–398. doi:10.1037/0882-7974.23.2.392

Green, D., & Swets, J. (1966). *Signal detection theory and psychophysics*. Oxford, UK: Wiley.

Han, S. (2009). Trait individual difference of reinforcement-based learning during episodic recognition judgments. *Korean Journal of Cognitive Science*, 20, 357–381

Han, S., & Dobbins, I. G. (2008). Examining recognition criterion rigidity during testing using a biased-feedback technique: evidence for adaptive criterion learning. *Memory & Cognition*, 36, 703–715. doi:10.3758/MC.36.4.703

Han, S., & Dobbins, I. (2009). Regulating recognition decisions through incremental reinforcement learning. *Psychonomic Bulletin & Review*, 16, 469–474. doi:10.3758/PBR.16.3.469

Hedden, T., Park, D. C., Nisbett, R., Ji, L. J., Jing, Q., & Jiao, S. (2002). Cultural variation in verbal versus spatial neuropsychological function across the life span. *Neuropsychology*, 16, 65–73. doi:10.1037/0894-4105.16.1.65

Hertzog, C., Sinclair, S., & Dunlosky, J. (2010). Age differences in the monitoring of learning: Cross-sectional evidence of spared resolution across the adult lifespan. *Developmental Psychology*, 46, 939–948. doi:10.1037/a0019812

Hess, T. M., Auman, C., Colcombe, S. J., & Rahhal, T. A. (2003). The impact of stereotype threat on age differences in memory performance. *The Journals of Gerontology. Series B: Psychological Sciences and Social Sciences*, 58, P3–P11. doi:10.1093/geronb/58.1.P3

Huh, T., Kramer, J. H., Gazzaley, A., & Delis, D. (2006). Response bias and aging on a recognition memory task. *Journal of the International Neuropsychological Society*, 12, 1–7. doi:10.1017/S1355617706060024

Humphries, J., Flowe, H., Hall, L., Williams, L., & Ryder, H. (2015). The impact of beliefs about face recognition ability on memory retrieval processes in young and older adults. *Memory*.

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. *International Journal of Aging & Human Development*, 41, 79–88. doi:10.2190/ERDG-LHA8-EBYN-L9LX

Kantner, J., & Linsday, D. (2012). Response bias in recognition memory as a cognitive trait. *Memory & Cognition*, 40, 1163–1177. doi:10.3758/s13421-012-0226-0

Kapucu, A., Rotello, C., Ready, R., & Seidl, K. (2008). Response bias in "remembering" emotional stimuli: A new perspective on age differences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 703–711. doi:10.1037/0278-7393.34.3.703

Karl, K., O'Leary-Kelly, A., & Martocchio, J. (1993). The impact of feedback and self-efficacy on performance in training. *Journal of Organizational Behavior*, 14, 379–394. doi:10.1002/job.4030140409

Knopman, D., & Nissen, M. (1987). Implicit learning in patients with probably Alzheimer's disease. *Neurology*, 37, 784. doi:10.1212/WNL.37.5.784

Kucera, H., & Francis, W. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

Lauriola, M., & Levin, I. (2001). Personality and risky decision-making in a controlled experimental task: an exploratory study. *Personality and Individual Differences*, 31, 215–226. doi:10.1016/S0191-8869(00)00130-6

Levy, B. R. (2003). Mind matters: cognitive and physical effects of aging self-stereotypes. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 58, P203–P211. doi:10.1093/geronb/58.4.P203

Macmillan, N., & Creelman, C. (1991). *Detection theory: A user's guide*. New York: Cambridge University Press.

Marquié, J. C., & Baracat, B. (2000). Effects of age, education, and sex on response bias in a recognition task. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 55, P266–P272. doi:10.1093/geronb/55.5.P266

Maurer, T. (2001). Career-relevant learning and development, worker age, and beliefs about self-efficacy for development. *Journal of Management*, 27, 123–140. doi:10.1177/014920630102700201

McDaniel, M. Einstein, G., & Jacoby, L. (2008). New considerations in aging and memory. In: F.I.M. Craik, T.A. Salthouse (Eds.), *The handbook of aging and cognition* (3rd ed., pp. 251–310). New York: Psychology Press.

Miller, M., Handy, T., Cutler, J., Inati, S., & Wolford, G. (2001). Brain activations associated with shifts in response criterion on a recognition test. *Canadian Journal of Experimental Psychology*, 55, 162–173. doi:10.1037/h0087363

Pazzaglia, A., Dube, C., & Rotello, C. (2013). A critical comparison of discrete-state and continuous models of reocgnition memory: Implications for recognition and beyond. *Psychological Bulletin*, 139, 1173–1203. doi:10.1037/a0033044

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

Peters, E., Hess, T. M., Vastfjall, D., & Auman, C. (2007). Adult age differences in dual information processes: Implications for the role of affective and deliberative processes in older adults' decision making. *Perspectives on Psychological Science*, 2, 1–23. doi:10.1111/j.1745-6916.2007.00025.x

Poon, L. W., & Fozard, J. L. (1980). Age and word frequency effects in continuous recognition memory. *Journal of Gerontology*, 35, 77–86. doi:10.1093/geronj/35.1.77

Ratcliff, R., McKoon, G., & Tindall, M. (1994). Empirical generality of data from recognition memory receiver-operating characteristic functions and implications for the global memory models. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 20, 763–785. doi:10.1037/0278-7393.20.4.763

Ratcliff, R., Sheu, C., & Gronlund, S. (1992). Testing global memory models using ROC curves. *Psychological Review*, 3, 518–535. doi:10.1037/0033-295X.99.3.518

Rosa, N., & Gutchess, A. (2013). False memory in aging resulting from self-referential processing. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 68, 882–892. doi: 10.1093/geronb/gbt018

Sauvage, M., Beer, Z., & Eichenbaum, H. (2010). Recognition memory: Adding a response deadline eliminates recollection but spares familiarity. *Learning and Memory*, 17, 104–108. doi:10.1101/lm.1647710

Shipley, W. (1986). *Shipley Institute of Living Scale*. Los Angeles, CA: Western Psychological Services.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: applications to dementia and amnesia. *Journal of Experimental Psychology. General*, 117, 34–50. doi:10.1037/0096-3445.117.1.34

Starns, J., & Ratcliff, R. (2010). The effects of aging on the speed-accuracy compromise: Boundary optimality in the diffusion model. *Psychology and Aging*, 25, 377–390. doi:10.1037/a0018022

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

Trippas, D., Handley, S., & Verde, M. (2014). Fluency and belief bias in deductive reasoning: new indices for old effects. *Frontiers in Psychology*, 5, 631. doi:10.3389/fpsyg.2014.00631

Vakil, E., Mosak, C., & Ashkenazi, M. (2003). The effect of aging on script memory for typical and atypical actions. *Applied Neuropsychology*, 10, 239–245. doi:10.1207/s15324826an1004_6

Waring, J., Chong, H., Wolk, D., & Budson, A. (2008). Preserved metamemorial ability in patients with mild Alzheimer's disease: Shifting response bias. *Brain and Cognition*, 66, 32–39. doi:10.1016/j.bandc.2007.05.002

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale* (3rd edn). San Antonion, TX: Harcourt Assessment.

Weller, J., Levin, I., & Denburg, N. (2011). Trajectory of risky decision making for potential gains and losses from ages 5 to 85. *Journal of Behavioral Decision Making*, 24, 331–344. doi:10.1002/bdm.690

West, R., Bagwell, D., & Dark-Freudman, A. (2005). Memory and goal setting: the response of older and younger adults to positive and objective feedback. *Psychology and Aging*, 20, 195–201. doi:10.1037/0882-7974.20.2.195

White, K., & Wixted, J. (1999). Psychophysics of remembering. *Journal of the Experimental Analysis of Behavior*, 71, 91–113. doi:10.1901/jeab.1999.71-91

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:10.1006/jmla.2002.2864

Yonelinas, A., & Parks, C. (2007). Receiver operating characteristics (ROCs) in recognition memory: A review. *Psychological Bulletin*, 133, 800–832. doi:10.1037/0033-2909.133.5.800

Zacks, R., Hasher, L., & Li, K. (2000). Human memory. In F. I. Craik & T. Salthouse (Eds.), *The handbook of aging and cognition* (2nd edn, pp. 293–357). Mahwah, NJ: Erlbaum.