

## Aging and the Role of Attention in Associative Learning

Sharon A. Mutter, Jared M. Holder, Cody A. Mashburn, and Catherine M. Luna  
Western Kentucky University

In this study, we investigated whether age-related deficits in cue–outcome associative learning (e.g., Mutter, Atchley, & Plumlee, 2012; Mutter, DeCaro, & Plumlee, 2009; Mutter, Haggblom, Plumlee, & Schirmer, 2006; Mutter & Williams, 2004) might be due to a decline in older adults' ability to modulate attention to relevant and irrelevant cues. In the first 2 experiments, we used standard blocking and highlighting tasks to indirectly measure the ability to shift attention away from irrelevant stimuli toward relevant, predictive cues (e.g., Kruschke, Kappenman, & Hetrick, 2005). Although there were age differences in prediction accuracy, like young adults, older adults learned to shift attention toward predictive stimuli and ignore irrelevant or less predictive stimuli. This attentional effect was unrelated to either working memory or executive function suggesting that it did not involve voluntary control processes. The third experiment provided further support for this idea. We alternated a category learning task with a dot probe task to more directly assess the development of automatic attentional biases. There were again age differences in category prediction, but young and older adults alike responded more rapidly to the location of a dot probe cued by a stimulus experienced as predictive during the learning task than one cued by a stimulus experienced as nonpredictive. These findings provide converging evidence that even though cue–outcome prediction declines with age, the ability to modulate attention based on the predictive relevance of cues during associative learning remains intact.

*Keywords:* attention, associative learning, predictive learning, aging

*Supplemental materials:* <http://dx.doi.org/10.1037/pag0000277.supp>

Attention plays a prominent role in many theories of associative learning (e.g., Kruschke, 2003; Mackintosh, 1975; Pearce & Hall, 1980; Pearce & Mackintosh, 2010). In these theories, learning processes driven by error reduction form associations among the representations of environmental stimuli (i.e., cues) that signal or cause the occurrence of outcome events (e.g., Rescorla & Wagner, 1972). However, efficient acquisition of these associations requires discovering and focusing on the most relevant stimuli while ignoring distracting or irrelevant stimuli. Research has shown that young adults readily learn to shift attention to the most relevant

stimuli to reduce errors during associative learning (e.g., Kruschke et al., 2005; Le Pelley, Beesley, & Griffiths, 2011). Older adults are less efficient at acquiring cue–outcome associations than young adults (Mutter et al., 2006, 2009, 2012; Mutter & Williams, 2004), but there have been few studies exploring whether this might be due to age-related differences in attention. The goal of this research was to determine whether older adults differ from young adults in their ability to modulate attention to relevant and irrelevant stimuli during associative learning.

“Blocking” was one of the first demonstrations of attention in associative learning. In a blocking task, participants first learn that cue A reliably predicts outcome X ( $A \rightarrow X$ ). They then learn that cue A presented in compound with a second cue B reliably predicts the outcome ( $A.B \rightarrow X$ ). Although cue B is paired with outcome X, participants do not learn this association. The initial explanation for this finding was that the existing  $A \rightarrow X$  association “blocked” acquisition of the association between B and the outcome (e.g., Kamin, 1969; Rescorla & Wagner, 1972). However, this associative account could not explain evidence that the ability to learn completely new associations involving the blocked cue is greatly attenuated (Beesley & Le Pelley, 2011; Mackintosh & Turner, 1971). It appears instead that participants learn to shift their attention away from the redundant cue, which reduces its associability (i.e., how readily it can be associated with the outcome; Mackintosh, 1975). Recent studies of eye gaze behavior offer further support for learned inattention to blocked cues. Gaze duration is reduced for blocked cues (Kruschke et al., 2005) and this is directly related to the reduced rate of learning novel associations for these stimuli (Beesley & Le Pelley, 2011).

---

Sharon A. Mutter, Jared M. Holder, Cody A. Mashburn, and Catherine M. Luna, Department of Psychological Sciences, Western Kentucky University.

Jared M. Holder is now at Research and Evaluation, Centerstone Research Institute, Nashville, Tennessee. Cody A. Mashburn is now at the School of Psychology, Georgia Institute of Technology.

This research reported in this article was presented at the 28th Annual Meeting of the Psychonomic Society. Experiment 2 is based on research conducted by Jared M. Holder in partial fulfillment of the requirements for the M.A. degree in Psychology; Experiment 3 was supported by the National Science Foundation under Cooperative Agreement 1355438. We thank Melanie Asriel, Candice Tucker-Groves, and Keely Lawrence for their assistance in collecting the data for this project.

Correspondence concerning this article should be addressed to Sharon A. Mutter, Department of Psychological Sciences, Western Kentucky University, 1908 College Heights Boulevard #22030, Bowling Green, KY 42101. E-mail: [sharon.mutter@wku.edu](mailto:sharon.mutter@wku.edu)

Just as individuals learn to shift their attention away from redundant stimuli, they also learn to shift their attention toward the most predictive stimuli. Kruschke (2003, 2009; Kruschke et al., 2005) has described a phenomenon called “highlighting” that demonstrates the beneficial effect of this type of attentional modulation during associative learning. Participants first learn that an imperfect predictor (I) in compound with an initial (early) perfect predictor (PE) is associated with an initial (early) outcome (I.PE→E). They then learn that the same imperfect predictor in compound with a new (late) perfect predictor (PL) is associated with a new (late) outcome (I.PL→L). At test, participants show a strong preference to predict the early outcome for I and strong preference to predict the late outcome for PE.PL, a new compound cue formed from the perfect early and perfect late predictors (see Table 1 for further details on the highlighting design). These preferences are correlated with increased gaze duration for PL as compared to I and PE. Kruschke suggested that these effects occur because the I→E association learned first leads to prediction errors when cue I is paired with PL and the new outcome L. To prevent these errors, participants learn to shift their attention away from I toward PL. Because PL is attentionally “highlighted,” the association between PL and outcome L is stronger than the association between PE and outcome E.

Learning to shift attention reduces interference between competing cues, accelerates the learning of new cue–outcome associations, and protects previously learned associations (Kruschke et al., 2005). There is evidence for a decline in both the rate and strength of older adults’ associative learning (e.g., Mutter et al., 2006, 2009, 2012; Mutter & Williams, 2004), which could be exacerbated by age-related differences in attention. Older adults

have difficulty attending to relevant stimuli while suppressing attention to irrelevant stimuli in perceptual and working memory (WM) tasks (e.g., Gazzaley, Cooney, Rissman, & D’Esposito, 2005; Schmitz, Cheng, & De Rosa, 2010), but there has been little research on whether there might be an age-related deficit in attentional modulation for relevant and irrelevant stimuli during associative learning.

Hannah, Allan, and Young (2012) examined blocking effects in young and older adults’ learning using a “streamed trial” adaptation of the classic learning task. Their results initially suggested that there was an age-related deficit in blocking, but this disappeared when irrelevant stimuli were perceptually distinct from the relevant cues and outcomes. Increased perceptual load apparently disrupted older adults’ attentional selection (cf. Schmitz et al., 2010), but when irrelevant stimuli were perceptually distinctive, older adults learned to ignore these stimuli. There have been no studies examining age differences in the highlighting effect. However, research suggests that older adults may have difficulty inhibiting attention to preexisting associations during learning. For example, Mutter, Strain, and Plumlee (2007; see also Mutter & Poliske, 1994) found that older adults’ contingency estimates were less accurate than young adults’ estimates, especially when new evidence regarding the contingency relationship contradicted their prior expectations. In line with studies showing that preexisting knowledge affects which features are attended during associative learning (e.g., Kim & Rehder, 2011), this finding suggests that older adults selectively attended to events that were nonpredictive, but consistent with their preexisting associative knowledge. However, it is not clear whether this age difference was due to poorer attentional modulation during learning or to weaker memory for

Table 1  
*Blocking and Highlighting Task Designs*

Phase	Blocks	Experiment 1a: Blocking				Blocks	Experiment 1b: Highlighting	
		Trial type					Trial type	
Early	18	A1→X1 A2→X2	F1→Y1 F2→Y2	— —	— —	12	I1.PE→E1 I2.PE2→E2 <i>3 times</i>	PE1.I1→E1 PE2.I2→E2 <i>1 time</i>
Late	9	A1.B1→X1 B1.A1→X1	A2.B2→X2 B2.A2→X2	C1.D1→Y1 D1.C1→Y1	C2.D2→Y2 D2.C2→Y2	6	I1.PE1→E1 PE1.I1→E1 I2.PE2→E2 PE2.I2→E2 <i>1 time</i>	I1.PL1→L1 PL1.I1→L1 I2.PL2→L2 PL2.I2→L2 <i>3 times</i>
						18	I1.PE1→E1 PE1.I1→E1 I2.PE2→E2 PE2.I2→E2 <i>1 time</i>	I1.PL1→L1 PL1.I1→L1 I2.PL2→L2 PL2.I2→L2 <i>3 times</i>
Test	1	A1.B1→X1 B1.A1→X1  B1.D1→? D1.B1→? B1.C1→? C1.B1→?	A2.B2→X2 B2.A2→X2  B1.D2→? D2.B1→? B1.C2→? C2.B1→?	C1.D1→Y1 D1.C1→Y1  B.D: B2.D1→? D1.B2→? B2.C1→? C1.B2→?	C2.D2→Y2 D2.C2→Y2  A.C: A2.D1→? D1.A2→? A2.C1→? C1.A2→?	1	I1.PE1→? PE1.I1→? I2.PE2→? PE2.I2→? I1→? I2→? I1→? I2→?	I1.PL1→? PL1.I1→? I2.PL2→? PL2.I2→? PE1.PL1→? PE2.PL2→? PE1.PL1→? PE2.PL2→?

Note. Each block consisted of two trials of a given trial type.

new contingency evidence relative to preexisting associations during contingency estimation.

The findings from the few studies addressing whether older adults learn to shift attention to relevant and irrelevant stimuli during associative learning provide little clarity on the issue. Moreover, these studies do not offer a particularly strong test of error-driven attentional modulation because the learning tasks used involved passive viewing of event–state pairs. Therefore, we used predictive learning tasks to examine age differences in the ability to flexibly shift attention among competing stimuli. Cue presentation was followed by outcome prediction and feedback, so there was more opportunity for trial-by-trial prediction error to drive changes in learning and attention. In addition, cue–outcome pairs were clearly segregated, so there was little reason to expect perceptual load to influence older adults' performance. In Experiments 1 and 2, we investigated age differences in blocking and highlighting and in Experiment 3, we alternated learning trials with a dot probe task to more directly assess age differences in the development of attentional biases to predictive stimuli.

### Experiment 1 (Blocking) and Experiment 2 (Highlighting)

Any difficulty that older adults might have in learning to shift attention away from preexisting, irrelevant cue–outcome associations during learning may affect their performance differently in blocking and highlighting tasks. In a blocking task, the original predictive cue is still relevant during later training and should induce both young and older adults to ignore the new, redundant cue. Thus, both groups should show less learning for the blocked cue and there should be no age difference in the magnitude of this blocking effect. In contrast, failure to shift attention away from preexisting associations could prevent older adults from showing a highlighting effect. During later training in the highlighting task, participants must learn to shift attention away from the imperfect predictor (I) toward the late perfect predictor (PL). If older adults

are less able to do this they should show a smaller highlighting effect than young adults; that is, they should show little or no preference for the late outcome when the early and late perfect predictors are presented in compound (PE.PL).

Alternatively, Kruschke (2009) has suggested that the same attentional shifting process is used in both blocking and highlighting and that individuals who are capable of learning to suppress attention to irrelevant stimuli should show both effects. From this viewpoint, if older adults do not learn to shift their attention away from irrelevant stimuli (i.e., the redundant cue in blocking and the imperfect predictor in highlighting), they should show weaker blocking and highlighting effects than young adults, but if they do learn to shift attention away from these stimuli, no age differences would be expected in either effect.

### Method

**Participants.** Forty-eight young adults (YA) were recruited from psychology courses at Western Kentucky University and received class credit for participating; 48 older adults (OA) were recruited from the community and were paid a small stipend. All participants were screened for the use of medications and health problems that could have an impact on cognitive functioning and none were excluded. Older adults were also screened for dementia via the Telephone Mini Mental State Examination (TMMSE) and all who participated met a passing criterion. Participants were randomly assigned to either the blocking task or the highlighting task. Demographic and cognitive characteristics for each group are shown in Table 2.

**Stimuli and task design.** The blocking and highlighting tasks were adapted from Kruschke et al. (2005) and Kruschke (2009). Table 1 shows the design for these tasks. Each letter and number combination represents a unique cue or outcome. To create the cues, a pool of 22 five-letter nouns were chosen from the MRC Psycholinguistic Database ([http://www.psy.uwa.edu.au/mrcdata/base/uwa\\_mrc.htm](http://www.psy.uwa.edu.au/mrcdata/base/uwa_mrc.htm)) with concreteness, imagability, and familiarity

Table 2  
*Participant Characteristics*

Measure	Experiment 1		Experiment 2		Experiment 3	
	<i>M</i> ( <i>SD</i> ) <sub>Young</sub> <i>n</i> = 24	<i>M</i> ( <i>SD</i> ) <sub>Older</sub> <i>n</i> = 24	<i>M</i> ( <i>SD</i> ) <sub>Young</sub> <i>n</i> = 23	<i>M</i> ( <i>SD</i> ) <sub>Older</sub> <i>n</i> = 23	<i>M</i> ( <i>SD</i> ) <sub>Young</sub> <i>n</i> = 27	<i>M</i> ( <i>SD</i> ) <sub>Older</sub> <i>n</i> = 27
Age	19.79 (2.73)	69.92 (5.78)	22.75 (3.30)	70.70 (4.66)	20.41 (2.12)	70.63 (4.86)
Education	14.13 (1.42)	16.96 (3.22)	15.30 (1.99)	14.83 (3.04)	14.00 (1.82)	15.91 (3.26)
DS	83.67 (11.09)	58.54 (11.99)***	89.74 (16.44)	61.04 (13.04)***	81.67 (10.60)	58.04 (12.71)***
RS	2.75 (.94)	2.04 (1.08)*	4.09 (1.83)	2.57 (1.59)**	2.56 (1.197)	1.89 (1.15)*
VOCAB <sup>a</sup>	28.00 (5.22)	39.04 (39.04)***	35.78 (6.04)	38.30 (8.03)	13.16 (5.56)	19.79 (8.14)***
PA	58.88 (1.33)	57.17 (3.16)*	59.00 (1.41)	56.30 (3.53)**		
PA—Interference	42.33 (42.33)	29.71 (12.71)***	46.61 (9.76)	30.00 (12.91)***		
CAL—FR	.11 (.14)	.25 (.20)**	.08 (.12)	.26 (.17)***	.14 (.10)	.37 (.20)***
CAL—P	.08 (.11)	.14 (.11)*	.06 (.09)	.13 (.11)*	.08 (.12)	.17 (.13)**
WCST—CC	3.73 (.99)	2.91 (1.69)*	4.14 (1.25)	3.00 (1.11)**	3.96 (.87)	1.81 (1.47)***
WCST—PE	8.00 (4.33)	8.69 (4.94)	5.64 (1.84)	8.36 (4.19)**	8.38 (4.88)	11.67 (5.59)*

*Note.* DS = Digit Symbol (Wechsler, 1997); RS = Reading Span (Daneman & Carpenter, 1980); PA = Paired Associate; CAL = Conditional Associative Learning (Levine, Stuss, & Milberg, 1997); FR = Forgotten Responses; P = Perseverations; WCST = Wisconsin Card Sorting Task (Grant & Berg, 1948); CC = Categories Completed; PE = Perseverative Errors.

<sup>a</sup> Experiment 1 and 2 Mill Hill Vocabulary Test (Raven, Raven, & Court, 1998), Experiment 3 Advanced Vocabulary Test (Ekstrom, French, Harman, & Dermen, 1976).

\*  $p \leq .05$ . \*\*  $p \leq .01$ . \*\*\*  $p \leq .001$ .

ratings of at least 500. The words in this pool had no apparent semantic relationships with each other and all began with a different initial letter. The words selected were *apple, brain, cigar, daisy, elbow, frost, glass, house, ivory, judge, knife, linen, movie, nurse, ocean, phone, queen, radio, skate, tiger, uncle, and world*. Twelve unique lists were created by randomly assigning words from this pool to the cue types in Table 1. Across these lists, the letters F, G, H, and J were assigned equally often as outcomes to each cue type in each task.

Stimuli were presented on an Apple iMac computer with a 20-in. screen. Single cues were presented in the center of the screen and compound cues were presented in the center with one word positioned above the other. The position of the two words in each compound cue was counterbalanced so that there were two presentations of each unique compound cue within each block. The four outcomes were displayed in horizontal boxes at the bottom of the screen (see Figure 1). Participants' responses were made on a standard computer keyboard using the F, G, H, and J keys.

**Experiment 1: Blocking.** The blocking task consisted of early and late training phases and a test phase. During each training trial, a single (early phase) or compound (late phase) cue appeared and participants made a prediction response, which was followed by feedback on response accuracy. The early training phase was designed to ensure that participants learned the A→X associations. The F→Y cue–outcome trials were included as fillers and were not shown in any other phase of the task. The late training phase was designed to replicate Kamin's (1969) blocking procedure, which relies upon pairing the previously trained cue with a novel cue. The target compound cue–outcome pairs in this phase were created by pairing old A stimuli with novel B stimuli (i.e., A1.B1→X1, A2.B2→X2). For these pairs, the existing A→X associations should block the acquisition of B→X associations. Two additional novel pairs (i.e., C1.D1→Y1; C2.D2→Y2) were presented as controls during the late training phase. There should be no blocking effect for the individual stimuli in these compound cues because they were both novel.

In the test phase, two additional training trials were given for each A.B and C.D compound cue to ensure that participants retained the outcomes associated with these cues during training. These trials were randomly intermixed with presentations of A.C (i.e., A.C and A.D) and B.D (i.e., B.C and B.D) compound cues to test for outcome preference. No feedback was given for these test trials. A should have a strong association with X, B should have little or no association with X, and C and D should have equally strong associations with Y. Thus, a blocking effect will be evident if participants show a strong preference to predict outcome X for A.C cues and a strong preference to predict outcome Y for B.D cues.

**Experiment 2: Highlighting.** The highlighting task consisted of one early training phase, two late training phases, and a test phase. For the compound cues presented in this task, the letter I indicates that the individual cue is an Imperfect predictor and the letter P indicates that the individual cue is a Perfect predictor. As in the blocking task, each unique compound cue–outcome pair was presented twice to accommodate counterbalancing of individual cue position. Likewise, each training trial consisted of cue presentation, outcome prediction, and response accuracy feedback.

In the early training phase, participants learned cue–outcome associations for compound cues composed of the imperfect

predictor and early perfect predictors (i.e., I1.PE1→E1 and I2.PE2→E2). In the late training phases, they learned associations for the same imperfect predictors combined with new late perfect predictors (e.g., I1.PL1→L1 and I2.PL2→L2). During the first late training phase, the compound I.PE and the I.PL cues were presented in the ratio of 3 to 1 to strengthen the previous associations for the early perfect predictors and to allow participants to begin forming associations for the late perfect predictors. This ratio was reversed in the second late training phase so that over the training trials, the total number of presentations for all cues was the same.

In the test phase, I.PE→E and I.PL→L trials were randomly intermixed with trials in which each imperfect predictor was presented alone and trials in which the early and late perfect predictors were presented together. No feedback was presented for any test trials. In the early phase of training, participants should acquire both I→E and PE→E associations. However, during the late phase of training, responding based on the earlier formed I→E association would produce errors for the I.PL cues, so to learn the I.PL→L associations, participants must shift their attention away from I to PL. As a result, the PL→L associations should be relatively stronger than the PE→E associations. Thus, the highlighting effect will be reflected in a greater preference for outcome L than outcome E when presented with the PE.PL cues.

**Procedure.** Experiment 1 (blocking) and Experiment 2 (highlighting) were approved by the Western Kentucky University (WKU) institutional review board (WKU HSRB#HS10-046, Learned Attention in Younger and Older Adults). Participants were tested individually during a single session. After providing informed consent and completing a Biographical and Health Questionnaire, they were seated at a comfortable viewing distance from the computer screen. For both the blocking and highlighting tasks, participants were told that the goal was to learn which of four keys (F, G, H, or J) to press when a particular word or pair of words was displayed on the computer screen. For each training trial, they viewed a cue and then made their outcome prediction by pressing a key on the keyboard. Cues remained on the screen until participants made a response and responses were followed by immediate feedback. If the response was correct, the phrase "Yes! The correct answer is [letter]" appeared between the cue stimuli and the response choices. If the response was incorrect, the phrase "Wrong! The correct answer is [letter]" was displayed. Training trials were followed immediately by test trials. In the blocking test, feedback was provided only after responses for the two additional training trials as described above; in the highlighting test, feedback was never provided after a response. In the latter case, responses were followed by the phrase "Your response has been recorded."

After finishing the learning task, participants completed tasks to measure speed of processing, working memory, executive function, associative learning, paired associate memory, and semantic knowledge (see Table 2). They were then debriefed and compensated for their time.

## Results and Discussion

Statistical analyses were conducted using null hypothesis significance testing (NHST) and Bayesian hypothesis testing. For NHST, we assumed an alpha level of .05 and, depending on the type of test, we used either Cohen's  $d$  or  $\eta_p^2$  as an estimate of

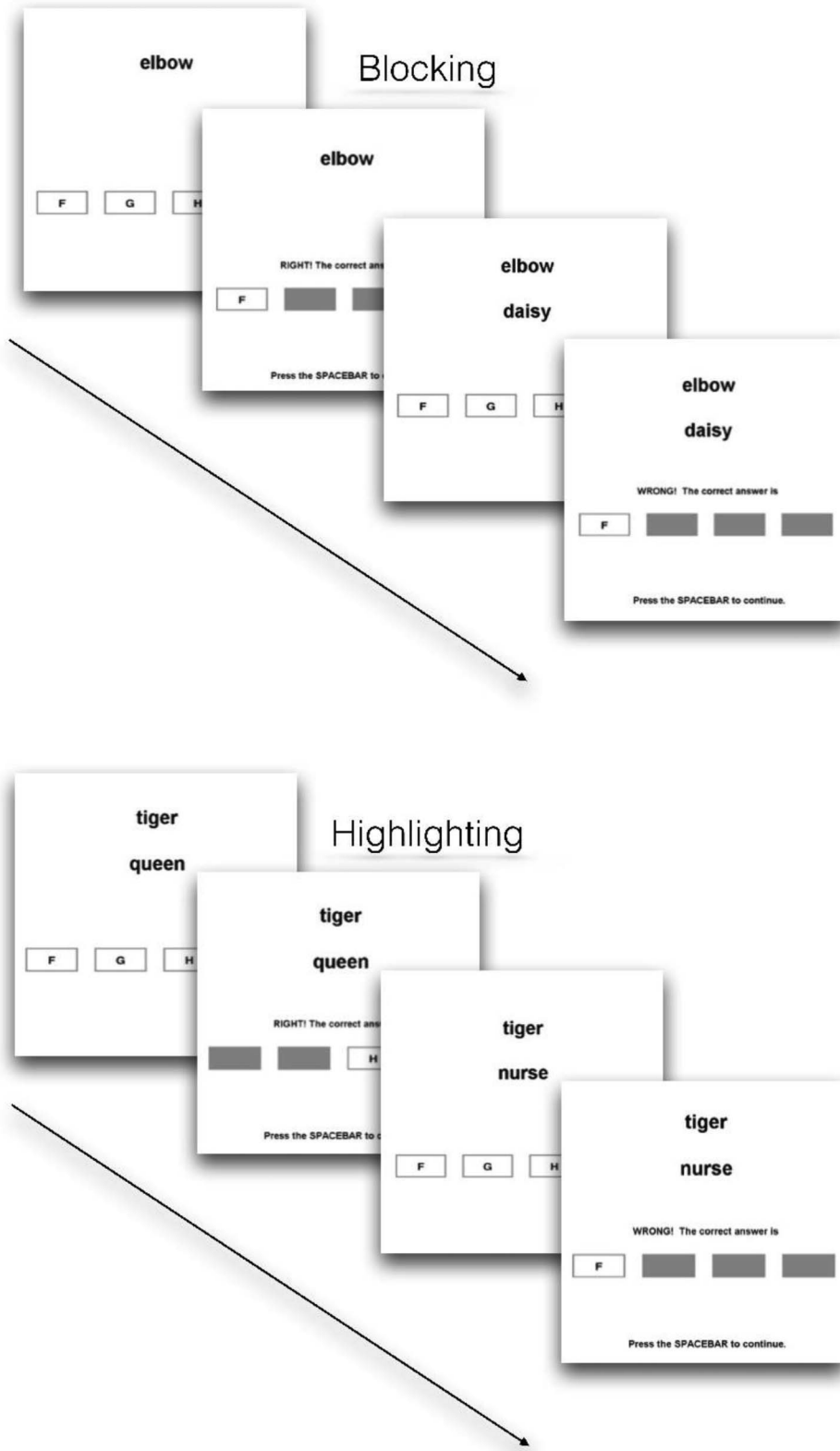


Figure 1. Stimuli used in the blocking and highlighting tasks showing how learning trials were formatted and presented to participants. The first two screens in the progression show Phase 1 training, and the latter two show Phase 2 training.

effect size. Information on the Bayesian analyses is provided in the online supplemental material.

**Experiment 1: Blocking.** Tests of the blocking effect require adequate learning of associations presented during the training phases of the blocking task. During the test phase, participants received eight additional A.B  $\rightarrow$  X and C.D  $\rightarrow$  Y training trials. According to the learning criterion used by Kruschke (e.g., Kruschke et al., 2005), they should respond correctly on six of these trials (75% correct) to be above chance performance. All participants met this learning criterion.

Our first analyses examined whether there were age differences in prediction accuracy during the training and test phases. Overall accuracy for the A  $\rightarrow$  X and F  $\rightarrow$  Y trials presented in the early training phase was higher for young ( $M = 89.23$ ,  $SD = 8.93$ ), than older adults ( $M = 80.87$ ,  $SD = 11.52$ ),  $F(1, 46) = 7.90$ ,  $MSE = 106.26$ ,  $p = .007$ ,  $\eta_p^2 = .15$ ,  $B_{10} = 6.26$ . But their overall accuracy for the A.B  $\rightarrow$  X and C.D  $\rightarrow$  Y trials presented in the late training phase (YA:  $M = 95.42$ ,  $SD = 3.16$ ; OA:  $M = 93.18$ ,  $SD = 4.78$ ),  $F(1, 46) = 3.67$ ,  $MSE = .1645$ ,  $p = .06$ ,  $\eta_p^2 = .07$ ,  $B_{10} = 1.25$ , and the test phase (YA:  $M = 97.92$ ,  $SD = 4.39$ ; OA:  $M = 96.35$ ,  $SD = 6.36$ ),  $F(1, 46) < 1$ ,  $MSE = 29.86$ ,  $p = .33$ ,  $\eta_p^2 = .02$ ,  $B_{10} = .43$ , did not differ. Thus, both groups learned the A.B and C.D associations.

Before examining response preferences for the blocked cue B, it is necessary to demonstrate that A  $\rightarrow$  X associations are stronger than C  $\rightarrow$  Y and D  $\rightarrow$  Y associations. Specifically, participants should have a stronger preference for X than Y when presented with the A.C compound test cues. The mean percent X and Y choices for these cues are presented in Figure 2. Participants also had the option of selecting outcomes that did not correspond with original cues (e.g., A1  $\rightarrow$  X2, C2  $\rightarrow$  Y1, etc.). These erroneous responses were made 2.86% of the time by young adults, and 3.13% of the time by older adults, and were not included in any analyses. A 2 (group)  $\times$  2 (outcome choice) ANOVA revealed that participants strongly preferred X over Y on A.C test trials and there was no age difference in this preference, group,  $F(1, 46) < 1.00$ ,  $MSE = 11.81$ ,  $p = .71$ ,  $\eta_p^2 = .003$ ,  $B_{10} = .23$ ; outcome choice,  $F(1, 46) = 25.02$ ,  $MSE = 860.20$ ,  $p < .001$ ,  $\eta_p^2 = .35$ ,  $B_{10} = 1.740e + 7$ ; Group  $\times$  Outcome Choice:  $F(1, 46) = 2.32$ ,  $p = .13$ ,  $\eta_p^2 = .05$ ,  $B_{10} = 1.87$ .

In contrast to the preference for X for the A.C compound cues, if the blocking effect is present, participants should prefer Y over X when presented with B.D compound cues. The mean percent X and Y choices for B.D cues are presented in Figure 2. Participants again had the option of selecting responses that corresponded to other cues (e.g., D1  $\rightarrow$  Y2, B2  $\rightarrow$  X1, etc.). Young adults responded erroneously on 6.51% of B.D trials and older adults responded erroneously on 12.24% of these trials. These responses were not included in any analyses. A 2 (group)  $\times$  2 (outcome choice) ANOVA revealed an age difference in overall X and Y choices,  $F(1, 46) = 5.13$ ,  $MSE = 45.71$ ,  $p = .03$ ,  $\eta_p^2 = .10$ ,  $B_{10} = .29$ , reflecting the slightly greater number of errors older adults made for the B.D trials. However, the Bayesian analysis indicated that the data were more likely under the hypothesis of no difference in young and older adults' overall X and Y choices than under the hypothesis of an age difference. More importantly, participants in both groups had an equally strong preference for Y over X on the B.D test trials, outcome choice,  $F(1, 46) = 17.15$ ,  $MSE = 562.70$ ,  $p < .001$ ,  $\eta_p^2 = .27$ ,  $B_{10} = 6.9035e + 4$ ; Group  $\times$  Outcome Choice,  $F(1, 46) = 1.28$ ,  $p = .26$ ,  $\eta_p^2 = .03$ ,  $B_{10} = .743$ . Thus, both young and older adults showed a large blocking effect for B and there was no age difference in this effect.

There are strong links between WM, executive function, and attentional control (e.g., Engle, 2002; Gazzaley & Nobre, 2012; McCabe, Roediger, McDaniel, Balota, & Hambrick, 2010). The presence of age differences in our measures of WM and executive function (see Table 2), combined with the absence of an age difference in the blocking effect raises the question of whether variations in attentional control are actually related to blocking. To answer this question, we calculated correlations between measures of blocking, WM, and executive function. To obtain a single outcome preference score for A.C cues, we subtracted the total number of times each participant gave a Y response from the total number of times each participant gave an X response and divided by the total number of A.C presentations (A minus C for choice or *AmCc*). To obtain a single outcome preference score for B.D cues, we subtracted the total number of X responses from the total number of Y responses and divided by the total number of B.D presentations (D minus B for choice or *DmBc*). We combined these two scores (i.e., *AmCc* + *DmBc*) to obtain a measure of

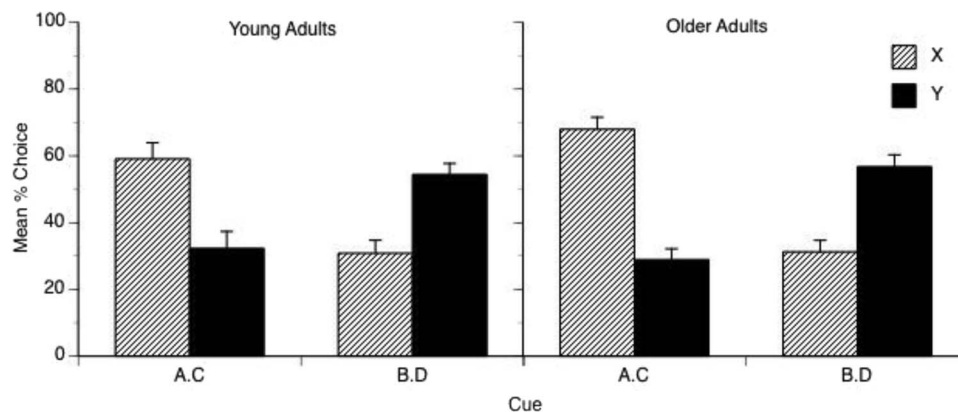


Figure 2. Mean percent X and Y outcome choices for young and older adults in the blocking task in Experiment 1. Error bars show standard error.

overall blocking (cf. Kruschke et al., 2005). There was no relationship between the blocking effect and WM, reading span:  $r(48) = -.10, p = .49, B_{10} = .23$ , or executive function, Wisconsin Card Sorting Task - Categories Completed (WCST-CC):  $r(48) = -.11, p = .46, B_{10} = .23$ . In contrast, prediction accuracy was positively correlated with measures of WM and executive function in the initial training phase, reading span:  $r(48) = .44, p = .002, B_{10} = 21.86$ ; WCST-CC:  $r(48) = .38, p = .008, B_{10} = 5.34$ , the late training phase, reading span:  $r(48) = .50, p = .000, B_{10} = 99.95$ ; WCST-CC:  $r(48) = .47, p = .001, B_{10} = 43.54$ , and at test, WM:  $r(48) = .31, p = .03, B_{10} = 1.75$ ; WCST-CC:  $r(48) = .39, p = .005, B_{10} = 7.65$ .<sup>1</sup> This suggests that although controlled attentional processes are related to prediction accuracy during associative learning, this form of attention may not be involved in the blocking effect.

**Experiment 2: Highlighting.** As with the blocking effect, testing the highlighting effect requires that participants learn the I.PE → E and I.PL → L associations during training. Participants received a total of eight I.PE → E and I.PL → L training trials during the test phase, so based on a 75% correct learning criterion, they should respond correctly on six of these test trials (Kruschke et al., 2005). Four participants (two young and two older) failed to reach the criterion. Two (one young and one older) were replaced with participants who met the criterion and two were excluded from further analysis, leaving 23 participants in each group.

We first examined prediction accuracy for young and older adults during the training and test phases. Age differences were present in prediction accuracy for the I.PE → E trials during the early training phase (YA:  $M = 95.64, SD = 3.12$ ; OA:  $M = 83.03, SD = 19.13$ ),  $F(1, 44) = 9.74, MSE = 187.86, p = .003, \eta_p^2 = .18, B_{10} = 12.13$ , and for the I.PE → E and I.PL → L trials during the late training phase (YA:  $M = 96.76, SD = 2.15$ ; OA:  $M = 88.34, SD = 10.96$ ),  $F(1, 44) = 13.05, MSE = 62.39, p = .001, \eta_p^2 = .23, B_{10} = 38.58$ . Overall accuracy for these trials was also higher for young adults than older adults in the test phase (YA:  $M = 98.10, SD = 3.49$ ; OA:  $M = 94.29, SD = 7.03$ ),  $F(1, 44) = 5.40, MSE = 30.80, p = .02, \eta_p^2 = .11, B_{10} = 2.45$ ; however, older adults' predictions were very accurate and the Bayesian analysis indicated that evidence for an age difference was weak (Raftery, 1995).

Turning now to analyses of the highlighting effect, we first examined outcome choices for I trials. Each participant received four I trials in the test phase of the experiment and the mean percent early (E) and late (L) outcome choices for these trials are shown in Figure 3. Participants also had the option of choosing outcomes that did not correspond to original cues (i.e., I1 → L2, I1 → E2, etc.). These errors accounted for 8.70% of responses for young adults and 10.87% of responses for older adults and these responses were not included in any analyses. A 2 (group) × 2 (outcome choice) ANOVA indicated that both young and older adults had a reliably greater preference for the E outcome during I test trials, group,  $F(1, 44) < 1.00, MSE = 145.75, p = .52, \eta_p^2 = .009, B_{10} = .25$ ; outcome choice,  $F(1, 44) = 40.14, MSE = 1110.42, p < .001, \eta_p^2 = .48, B_{10} = 2.488e + 10$ ; Group × Outcome Choice,  $F(1, 44) < 1.00, p = .94, \eta_p^2 = .00, B_{10} = .52$ .

In contrast to their preference for E when presented with cue I, participants should prefer L when presented with the compound cue PE.PL. Each participant was presented with four PE.PL trials in the test phase. The mean percent E and L outcome choices for these trials are shown in Figure 3. As with I test trials, participants

occasionally chose outcomes that did not correspond to the test items (e.g., PE1.PL1 → E2, PE1.PL1 → L2, etc.). These responses accounted for 1.09% of young adult and 2.17% of older adult responses and were not included in any analyses. A 2 (group) × 2 (outcome choice) ANOVA for the outcome choices for the PE.PL test trials indicated that young and older adults alike preferred L over E for the PE.PL test trials, group,  $F(1, 44) < 1.00, MSE = 19.76, p = .56, \eta_p^2 = .008, B_{10} = .23$ ; outcome choice,  $F(1, 44) = 12.52, MSE = 1762.60, p = .001, \eta_p^2 = .22, B_{10} = 6213.39$ ; Group × Outcome Choice,  $F(1, 44) < 1.00, p = .85, \eta_p^2 = .001, B_{10} = .28$ . Thus, both young and older adults showed a strong highlighting effect.

As before, age differences were observed for measures of WM and executive function, but not for highlighting. We therefore examined correlations between measures of highlighting, WM, and executive function. To obtain a single outcome preference score for the I test trials, we subtracted the total number of L responses from the total number of E responses and divided by the total number of I test trials ( $Ic$ ). To obtain a single outcome preference score for PE.PL, we subtracted the total number of E responses from the total number of L responses and divided by the total number of PE.PL test trials ( $PEmPLc$ ). We combined these two scores (i.e.,  $Ic + PEmPLc$ ) to obtain a measure of the overall highlighting effect (Kruschke et al., 2005). Highlighting was unrelated to WM, reading span:  $r(46) = -.008, p = .96, B_{10} = .18$ , or executive function, WCST-CC:  $r(46) = -.16, p = .31, B_{10} = .31$ . However, prediction accuracy in the early training phase, reading span:  $r(46) = .35, p = .02, B_{10} = 2.96$ ; WCST-CC:  $r(46) = .34, p = .02, B_{10} = 2.13$ , the late training phase, reading span:  $r(46) = .44, p = .002, B_{10} = 18.97$ ; WCST-CC:  $r(46) = .43, p = .003, B_{10} = 11.69$ , and the test phase, reading span:  $r(46) = .27, p = .07, B_{10} = .92$ ; WCST-CC:  $r(46) = .32, p = .04, B_{10} = 1.56$ , was again associated with WM and executive function (see Footnote 1).

These findings show that blocking and highlighting are robust effects for both young and older adults. The absence of an age difference in the blocking effect in Experiment 1 is consistent with previous research conducted by Hannah et al. (2012) and shows that older adults can learn to ignore irrelevant cues during associative learning. The highlighting effect we observed for older adults in Experiment 2 replicates and extends the Kruschke et al. (2005) findings with young adults. In that study, outcome predictions and visual gaze indicated that young adults distributed attention approximately equally among two cues and an outcome (I.PE → E) in initial training, but in later training, they shifted attention away from

<sup>1</sup> Because age is highly correlated with WM and executive function, we were interested in whether age effects in prediction accuracy were largely due to differences in these variables. Therefore, when there was a significant age difference in prediction accuracy, we conducted an ANCOVA for this effect controlling for performance on WM and executive function measures. This reduced age differences in accuracy in the early learning phase for the blocking experiment, Group,  $F(1, 44) = 2.89, MSE = 93.26, p = .10, \eta_p^2 = .06$ ; the early learning phase for the highlighting experiment, Group,  $F(1, 40) = 3.27, MSE = 194.25, p = .08, \eta_p^2 = .08$ ; the late learning phase for the highlighting experiment, Group,  $F(1, 40) = 3.08, MSE = 57.35, p = .09, \eta_p^2 = .07$ ; the test phase for the highlighting experiment, Group,  $F(1, 40) = 1.78, MSE = 31.80, p = .19, \eta_p^2 = .04$ ; and during learning in the dot probe experiment, Group,  $F(1, 49) = 2.46, MSE = 300.03, p = .12, \eta_p^2 = .05$ .

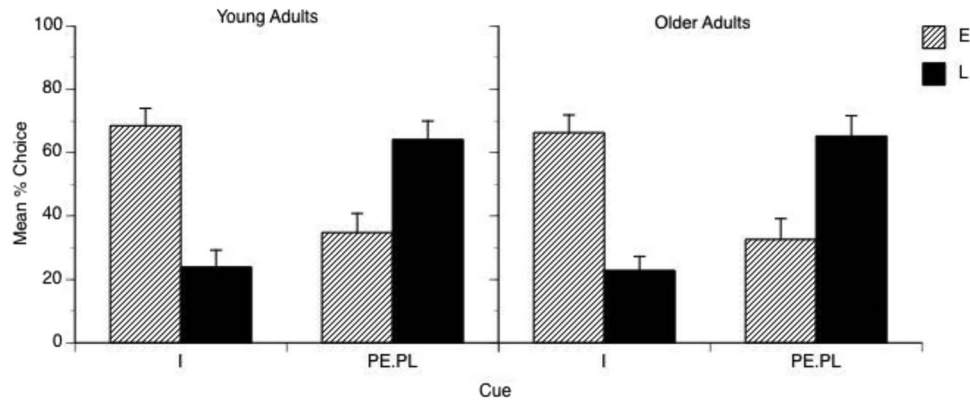


Figure 3. Mean percent early (E) and late (L) outcome choices for young and older adults in the highlighting task in Experiment 2. Error bars show standard error.

the early cue (I) toward the more predictive, novel cue (PL). Our findings suggest that older adults are equally capable of performing these attentional processes during associative learning.

Although the absence of age differences in blocking and highlighting was not entirely unexpected, null effects can be problematic. We are confident for several reasons that our findings reflect true age equivalence in these effects. First, we replicated the strong blocking and highlighting effects reported in previous research (e.g., Hannah et al., 2012; Kruschke, 2009; Kruschke et al., 2005; Sewell & Lewandowsky, 2012). Second, the distributions of blocking and highlighting scores for the two age groups were virtually identical (see Supplemental Figures S1a and S1b in the online supplemental material). Third, a post hoc power analysis indicated that while very large samples would be required to detect the weak age effects in blocking and highlighting,<sup>2</sup> our experimental design provided sufficient power to detect age differences in prediction accuracy and individual difference tests (see Table 2). And finally, Bayesian analyses indicated that the data for both blocking and highlighting were more likely to occur under a model without an Age  $\times$  Choice interaction than a model with this interaction.

The age equivalence we observed in blocking and highlighting suggests that aging spares the ability to modulate attention to stimuli during associative learning and also suggests that these effects involve similar attentional shifting processes (Kruschke, 2009; Kruschke et al., 2005). However, these attentional processes may not be voluntary. Attentional control is strongly linked to WM and executive function (Engle, 2002; Gazzaley & Nobre, 2012; McCabe et al., 2010), both of which show age-related decline (Reuter-Lorenz, Festini, & Jantz, 2016). This was likewise observed in the present research. Moreover, age, WM, and executive function were related to prediction accuracy during learning, but none of these variables were related to blocking and highlighting. Though surprising, this suggests that these effects involve a form of attention that is unrelated to WM (Sewell & Lewandowsky, 2012) and that learned predictiveness influences automatic rather than controlled attentional processes (Le Pelley, Vadillo, & Luque, 2013). We address this issue in the next experiment.

### Experiment 3

The findings from Experiments 1 and 2 suggest that young and older adults learn to direct attention toward predictive stimuli and away from irrelevant or less predictive stimuli during associative learning and further, that this attentional modulation may not be voluntary or controlled. However, these attentional effects were measured indirectly through outcome choice. Thus, it is not entirely clear whether experienced predictiveness actually influenced attention to the cues or whether it simply affected their associability (cf. Le Pelley et al., 2013). Moreover, any involuntary attentional effect for the cues was obscured by the overt attention that was required to formulate the outcome response.

To trace the development of an automatic attentional bias to predictive stimuli outside the learning context, we alternated a category learning task with a dot probe spatial cueing task developed by Le Pelley et al. (2013). In the learning task, participants saw displays with two squares, one green and one with oblique lines, on either side of a fixation point and learned whether the shade of green or the thickness of the lines predicted category membership. In the dot probe task, they saw the same displays, but after a delay, a small white square appeared over one of the squares and they indicated, as quickly as possible, the location of this probe. Le Pelley et al. found that young adults' response latencies to the dot probe were significantly shorter when its location was cued by a stimulus experienced as predictive during category learning and this effect increased as learning performance improved. Moreover, this attentional bias was present only when the delay between stimulus and probe onset was brief (i.e., 250 vs. 1,000 ms), suggesting that it reflected an involuntary rather than voluntary response to the stimulus.

<sup>2</sup> We conducted a post hoc power analysis to rule out the possibility that power was not sufficient to detect the much smaller age-related differences in blocking and highlighting. The results of this analysis indicated that to detect an age effect as small as the largest effect in our study while achieving power of at least 0.8 would require more than 250 participants. To detect the smallest effect that was larger than zero with power of 0.8 would require a sample size greater than 1,600.



Spatial cuing is a well-established methodology for investigating older adults' attentional responses to emotional stimuli (e.g., Isaacowitz, Wadlinger, Goren, & Wilson, 2006). Moreover, other research suggests that older adults have little difficulty using cues to guide their attention to a location. For example, valid endogenous cues (i.e., arrows) produce similar benefits in young and older adults' target detection (Hartley, Kieley, & Slabach, 1990). We therefore expected that if older adults learn to modulate their attention in the same way as young adults, stimuli experienced as predictive during category learning should be more likely to automatically capture their attention during the dot probe task than stimuli experienced as nonpredictive.

## Method

**Participants.** Twenty-eight young and 28 older adults were recruited and screened for participation in the same way as before. Because of the nature of the stimuli in this experiment, volunteers were also screened for colorblindness using the Ishihara Color Vision Test (Ishihara, 2006). The data for one young and one older adult were excluded from analysis due to excessive errors in the dot probe task and use of incorrect response keys in the category learning task, respectively. Demographic and cognitive characteristics for the groups are shown in Table 2.

**Stimuli and task design.** Stimuli consisted of colored squares with sides subtending a  $9^\circ$  visual angle from a viewing distance of 60 cm. Two squares were filled with either a light or dark shade of green (GreL and GreD) and two squares were filled with either thick or thin blue oblique lines (LinK and LinN) sloping downward from the left on a black background (see Supplemental Figure S2 in the online supplemental materials). In each trial, a stimulus pair consisting of one green square and one line-filled square appeared centrally on a white background. Across trials, stimulus pairs representing all four combinations of the two shades of green and two line thicknesses were presented. The squares were preceded and accompanied by a fixation cross that subtended a  $.67^\circ \times .67^\circ$  visual angle in the center of the screen. There was a distance of  $9^\circ$  between the center of the fixation cross and the center of each square. The dot probe was a small white square with sides subtending a  $1.35^\circ$  visual angle that appeared in the center of a green or line-filled square. On occasion, an arrow pointing to either the right ( $>$ ) or left ( $<$ ) replaced the fixation cross and no squares or dot probe were presented. Stimuli were presented on an iMac computer with a 20-in. screen.

In the category learning task, either the shade of green or the thickness of the lines predicted category membership throughout the task. The category dimensions were combined with two category responses, UP and DOWN, to produce four Predictive Category  $\times$  Category Response conditions. Participants in each age group were randomly assigned to one of these conditions. There were 32 blocks in the category learning task divided into eight phases of four blocks. In a block, each of the four stimulus pairs was presented twice to counterbalance the right/left position of the green and line-filled squares. These eight trials were randomly arranged in each block.

There were nine blocks in the dot probe task: one occurred prior to the category learning task to establish baseline dot probe response times and the remaining eight alternated with the phases of the category learning task (i.e., category learning, dot probe, cat-

egory learning, dot probe . . . etc.). There were 20 randomly arranged trials in each block of the dot probe task. The four stimulus pairs in their two counterbalanced left/right positions were presented twice for a total of 16 trials. The dot probe appeared once on each of the four squares in each of the left/right positions, ensuring that it was equally likely to appear on predictive and nonpredictive squares. There were also four arrow-only trials (two right and two left). Responses to all stimuli were made using labeled keys on the numeric keypad of the iMac keyboard (4 = left, 6 = right, 8 = up, 2 = down).

**Procedure.** Experiment 3 received approval from the WKU institutional review board (WKU IRB#16-486, Age Differences in Causal Learning). Participants were tested individually during a single session. After completing informed consent, colorblindness screening, and the Biographical and Health Questionnaire, participants were seated at a distance of 60 cm from the computer screen and given instructions for the experimental task. They were told that they would be performing two different tasks during the experiment and that the first would be a dot probe target detection task. On each trial, the fixation cross appeared in the middle of the display and was followed 500 ms later with color and line-filled squares that appeared on either side of the cross. The dot probe appeared 250 ms later, centered on one of the squares. Participants were told to respond to the left/right location of the dot probe as quickly as possible by pressing the key labeled *left* or the key labeled *right*. Participants were also told that it was important to keep their attention on the fixation cross at all times and that they would perform better if they ignored the colored squares. To further ensure that their attention remained on the fixation cross, they were told that on random occasions during the dot probe task an arrow pointing either left or right would replace the fixation cross after 500 ms and no stimulus pair or dot probe would appear. They were instructed to indicate as quickly as possible the direction the arrow was pointing. All stimuli remained on the screen until a response was made. Responses occurring before the onset of a dot probe prompted the feedback message "Do not anticipate stimuli" that appeared for 1,000 ms.

After the baseline dot probe task, participants received instructions for the category learning task. They were told that they would again see displays with a fixation cross and colored squares, and that it was still important to keep their attention on the cross. However, they were told that their goal was to learn which of two category responses, up or down, was correct for the pair of squares. They were instructed to use the labeled *up/down* keys to indicate their category decisions. Participants were told that accuracy rather than speed was important. However, if they made no response within 1,500 ms they received a prompt centered over the fixation cross and stimulus pair that consisted of a white square with the words "*up or down?*" surrounded by a black border with sides subtending a  $8.39^\circ$  visual angle. The prompt disappeared after a response was given. Immediately after a response, the message "That is correct!" or the message "ERROR! The correct response was [up/down]," appeared below the stimuli for 2,000 ms. For the remaining dot probe blocks and category learning phases, shifts from one task to the next were signaled by a display that indicated which task participants were about to perform and reminded them about how to make the responses for that task.

Afterward, participants completed tasks to measure speed of processing, WM, executive function, associative learning, and

semantic knowledge (see Table 2). They were then debriefed and compensated for their time.

## Results and Discussion

Data for the four Predictive Category  $\times$  Category Response conditions were pooled for all analyses of category learning accuracy and dot probe response time. We conducted statistical analyses using both NHST and hierarchical Bayesian analyses (see the online supplemental material).

**Category learning.** Young and older adults' mean category prediction accuracy across blocks is shown in Figure 4. A 2 (group)  $\times$  32 (block) ANOVA showed that accuracy improved over blocks for both groups,  $F(31, 1612) = 18.77$ ,  $MSE = 207.10$ ,  $p < .001$ ,  $\eta_p^2 = .26$ ,  $B_{10} = 4.749e + 83$ . Young adults' prediction accuracy was higher than older adults' across all blocks, group,  $F(1, 52) = 8.19$ ,  $MSE = 9452.00$ ,  $p = .006$ ,  $\eta_p^2 = .14$ ,  $B_{10} = 7.18$ ; Group  $\times$  Block,  $F(31, 1612) < 1.00$ ,  $p = .99$ ,  $\eta_p^2 = .007$ ,  $B_{10} = 6.644e-5$ . However, mean accuracy in the last block of the task was significantly greater than chance (50%) for both groups, YA,  $t(26) = 18.50$ ,  $p < .001$ ,  $d = 3.56$ ,  $B_{10} = 5.956e + 13$ ; OA,  $t(26) = 7.66$ ,  $p < .001$ ,  $d = 1.47$ ,  $B_{10} = 717320.00$ . Thus, both young and older adults learned the rule (i.e., color of the square or the size of the oblique lines in the squares) that predicted category membership.

To determine whether prediction accuracy during category learning was related to WM or executive functioning, we calculated zero order correlations between overall accuracy scores and these variables. There was a positive correlation between prediction accuracy and executive functioning, WCST-CC:  $r(54) = .31$ ,  $p = .02$ ,  $B_{10} = 2.12$ , and between accuracy and WM, though this latter relationship did not reach statistical significance, reading span:  $r(54) = .23$ ,  $p = .09$ ,  $B_{10} = .66$ .<sup>1</sup>

**Dot probe response.** Anticipation responses and errors were removed from the reaction time (RT) data. Across blocks, this constituted only 0.10% of the young adults' data and only 0.06% of the older adults' data. Median RTs were obtained for each participant for the predictive and the nonpredictive trials in each block of the task. To account for overall speed differences between

young and older adults, facilitation ratios were obtained for each block by subtracting the median RT for the predictive and non-predictive trials after learning from the median baseline RT for these trials before learning and then dividing by the respective median baseline RT (cf., Spieler, Balota, & Faust, 1996). Excluding the baseline dot probe phase, the data for each consecutive pair of dot probe blocks were then combined to produce four epochs. Young and older adults' mean facilitation ratios are shown in Figure 5.

Overall facilitation ratios increased over epochs and this did not vary by age, group,  $F(1, 52) < 1.00$ ,  $MSE = .13$ ,  $p = .98$ ,  $\eta_p^2 = .00$ ,  $B_{10} = .29$ ; Epoch,  $F(3, 156) = 7.27$ ,  $MSE = .014$ ,  $p < .001$ ,  $\eta_p^2 = .127$ ,  $B_{10} = 198.15$ ; Group  $\times$  Epoch,  $F(3, 156) = 1.31$ ,  $p = .27$ ,  $\eta_p^2 = .025$ ,  $B_{10} = .13$ , showing that response latencies for both young and older adults improved with practice. More importantly, for both age groups locating the dot probe over a square that predicted category membership produced greater facilitation in RT than locating the probe over a square that was nonpredictive and this effect was present throughout the learning task, cue,  $F(1, 52) = 6.42$ ,  $MSE = .04$ ,  $p = .01$ ,  $\eta_p^2 = .11$ ,  $B_{10} = 577.94$ ; Group  $\times$  Cue,  $F(1, 52) < 1.00$ ,  $p = .74$ ,  $\eta_p^2 = .002$ ,  $B_{10} = .17$ ; Epoch  $\times$  Cue,  $F(3, 156) = 1.68$ ,  $MSE = .003$ ,  $p = .17$ ,  $\eta_p^2 = .031$ ,  $B_{10} = .04$ ; Group  $\times$  Epoch  $\times$  Cue,  $F(3, 156) < 1.00$ ,  $p = .487$ ,  $\eta_p^2 = .016$ ,  $B_{10} = .06$ .

Zero order correlations between the average difference in facilitation ratios for predictive and nonpredictive cues and measures of WM and executive function indicated there was no relationship between the increase in facilitation observed for predictive cues and WM, reading span:  $r(54) = .16$ ,  $p = .26$ ,  $B_{10} = .32$ , or executive function, WCST-CC:  $r(53) = .11$ ,  $p = .43$ ,  $B_{10} = .23$ .

In summary, although age differences were observed in prediction accuracy during category learning, the experienced predictiveness of stimuli during learning affected young and older adults' dot probe responses in the same way. Like young adults, older adults responded more rapidly to the dot probe when its location was cued by a stimulus that was predictive of category membership than when its location was cued by a stimulus that was irrelevant to category membership. The findings of this experiment confirm in a more direct way than Experiments 1 and 2 that older adults experience little or no decline in modulating attention to relevant and irrelevant stimuli during associative learning.

The findings also replicate earlier research suggesting that the experienced predictiveness of stimuli influences automatic attentional capture (Le Pelley et al., 2013, Experiment 2). As Le Pelley et al. noted, any facilitation for the predictive cues with a brief 250-ms SOA likely involves the operation of involuntary attentional processes. Moreover, both young and older adults shifted attention to predictive cues even though this provided no advantage in probe detection and they were explicitly instructed that their performance would be better if they maintained central fixation throughout the task.

## General Discussion

In three experiments, we studied older adults' ability to direct attention toward relevant, predictive stimuli and away from irrelevant, less predictive stimuli during associative learning. In the first two experiments, we used standard blocking and highlighting tasks and found that although there were age differences in the

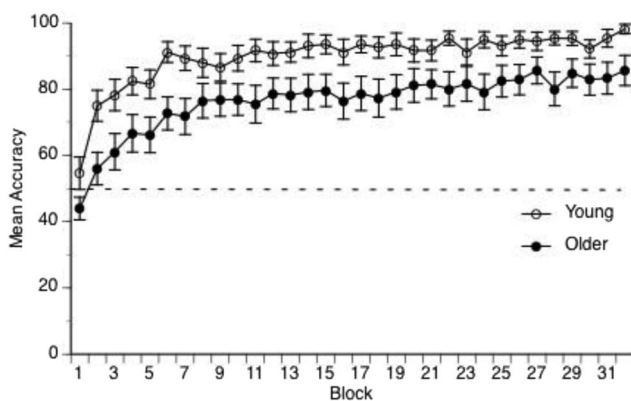


Figure 4. Mean percent correct responses for young and older adults across blocks in the category learning task in Experiment 3. Error bars show standard error. Chance performance is 50%. Every four blocks constituted one phase of the category learning task.

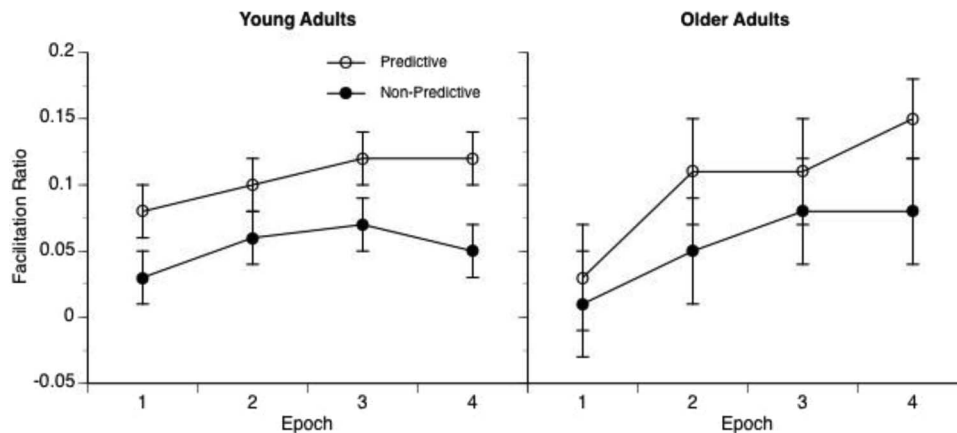


Figure 5. Mean dot probe facilitation ratios for young and older adults across epochs in Experiment 3. Facilitation ratio = [(predictive baseline RT—predictive learning RT)/predictive baseline RT] and [(nonpredictive baseline RT—nonpredictive learning RT)/nonpredictive baseline RT]. Error bars show standard error.

accuracy of cue–outcome prediction, like young adults, older adults learned to shift attention toward predictive stimuli and ignore irrelevant or less predictive stimuli. Prediction accuracy was related to WM and executive function, but blocking and highlighting were not, suggesting that these attentional effects did not involve voluntary control processes. The third experiment provided further support for this idea. We alternated a category learning task with a dot probe task to more directly assess attentional biases to stimuli outside the learning context. There were again age differences in prediction accuracy, but like young adults, older adults responded more rapidly to the location of a dot probe cued by a stimulus experienced as predictive during learning than one cued by a stimulus experienced as nonpredictive. These findings provide converging evidence that even though there is an age-related decline in prediction accuracy, there is no decline in the ability to modulate attention based on the predictive relevance of cues experienced during associative learning.

The outcome of this study is consistent in most respects with Mackintosh's (1975) attentional theory that suggests greater attention will be devoted to stimuli that are accurate predictors of an outcome than to those that are less predictive. The fact that there were no age differences in blocking, highlighting, or dot probe facilitation for predictive stimuli suggests that this attentional process is particularly robust. However, Mackintosh also proposed that attentional modulation improves the efficiency of error-based associative learning by increasing the associability of stimuli. The present results are not entirely consistent with this idea—age differences persisted in prediction accuracy despite older adults' successful attentional modulation for predictive and nonpredictive stimuli.

Recent hybrid theories of attention (e.g., Le Pelley, 2004; Pearce & Mackintosh, 2010) provide a more nuanced view of the relationship between attention, predictiveness, and associability. According to these theories, two independent attentional mechanisms operate during associative learning. Both are driven by prediction error (i.e., Rescorla & Wagner, 1972), but their operational rules differ. The first, attention to action (Holland & Maddux, 2010), is based on an assessment of the relative predictability of all stimuli in the learning environment. In line with Mackintosh (1975),

attention to action increases for stimuli generating the least prediction error. This type of attention promotes rapid detection and execution of appropriate responses for reliable predictors. It develops quickly, requires minimal capacity, and enhances the associability of these stimuli during new learning by automatically bringing them into the focus of attention (Pearce & Mackintosh, 2010). The second, attention to learning (Holland & Maddux, 2010), operates via an assessment of the uncertainty regarding outcome occurrence for a specific stimulus. More attention will be directed to stimuli that are followed by unexpected or surprising outcomes (i.e., learning is incomplete) than to those that are followed by well-predicted outcomes (Pearce & Hall, 1980). Attention to learning increases the salience of the stimulus and thus its associability but requires controlled processing and is limited in capacity (Pearce & Mackintosh, 2010).

It is likely that the blocking, highlighting, and dot probe effects in our study reflect attention to action. This is supported by the absence of any relationship between these effects and WM or executive function (cf. Sewell & Lewandowsky, 2012) as well as evidence that dot probe facilitation occurred automatically for stimuli experienced as predictive (cf. Le Pelley et al., 2013). Our findings therefore suggest that older adults retain the ability to automatically assess the relative predictability of environmental stimuli and direct attention to those that are the current best predictors. A number of other findings in the literature on aging and learning also suggest that attention to action is spared with age. For example, there are no age differences in the development of a visual search habit for targets that have a high probability of appearing in a particular quadrant of an array of distractors (Jiang, Koutstaal, & Twedell, 2016), older adults improve their target detection by acquiring and using tacit knowledge of spatial associations between targets and distractors in repeated search arrays (e.g., Howard, Howard, Dennis, Yankovich, & Vaidya, 2004; Merrill, Conners, Roskos, Klinger, & Klinger, 2013), and older adults show little decline in speeded visual search and response for stimuli associated with high versus low reward values (Störmer, Eppinger, & Li, 2014).

In contrast, attention to learning may not remain intact with age. Although we did not directly examine this form of attention in the

present study, the relationships we observed between prediction accuracy, WM, and executive function provide some evidence for the operation of a limited capacity attentional mechanism in our associative learning tasks. Moreover, an age-related deficit in attention to learning could explain why older adults continued to make prediction errors even when they were clearly able to assess the relative predictiveness of the cues. Specifically, this deficit reduced the salience of cue representations and thus their associability. In support of this idea, recent work by Nassar et al. (2016) has linked age-related deficits in associative learning to insufficient representation and use of uncertainty to guide learning. Further support comes from studies showing that age differences are exacerbated in learning contexts that make the resolution of uncertainty more difficult (e.g., Chasseigne et al., 2004; Eppinger, Hämmerer, & Li, 2011; Mutter & Plumlee, 2014).

This study has produced new findings that extend our current view of aging, attention, and associative learning. It also raises interesting questions for future research. First, our findings show that attention to action is preserved with age and further suggest that attention to learning could be impaired. Confirmation of an age-related dissociation in these two attentional mechanisms during associative learning will require studies in which both relative predictiveness and uncertainty regarding outcomes are manipulated in the same learning task. Second, research with young adults has shown that attention to action affects the associability of stimuli (Le Pelley et al., 2011; Mitchell, Griffiths, Seetoo, & Lovibond, 2012) and it is important to determine whether this is also the case for older adults. Finally, both attention to action and attention to learning are “tuned” by prediction error (e.g., Pearce & Mackintosh, 2010). There is considerable evidence that prediction error signals are generated by midbrain dopamine neurons in response to discrepancies between actual and expected outcomes (e.g., Schultz & Dickinson, 2000; Steinberg et al., 2013) and researchers have begun the work of mapping the distinct neural pathways that allow these signals to regulate the amount of learning and different types of attention to environmental stimuli (Nasser, Calu, Schoenbaum, & Sharpe, 2017). Dopamine neuron error signaling becomes less efficient with age (Bäckman, Nyberg, Lindenberger, Li, & Farde, 2006) and research on how this age-related change affects the operation of these neural pathways may shed further light on older adults’ attentional modulation during associative learning.

## References

- Bäckman, L., Nyberg, L., Lindenberger, U., Li, S. C., & Farde, L. (2006). The correlative triad among aging, dopamine, and cognition: Current status and future prospects. *Neuroscience and Biobehavioral Reviews*, *30*, 791–807. <http://dx.doi.org/10.1016/j.neubiorev.2006.06.005>
- Beesley, T., & Le Pelley, M. E. (2011). The influence of blocking on overt attention and associability in human learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *37*, 114–120. <http://dx.doi.org/10.1037/a0019526>
- Chasseigne, G., Ligneau, C., Grau, S., Gall, A. L., Roque, M., & Mullet, E. (2004). Aging and probabilistic learning in single- and multiple-cue tasks. *Experimental Aging Research*, *30*, 23–45. <http://dx.doi.org/10.1080/03610730490251469>
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning & Verbal Behavior*, *19*, 450–466. [http://dx.doi.org/10.1016/S0022-5371\(80\)90312-6](http://dx.doi.org/10.1016/S0022-5371(80)90312-6)
- Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). *Manual for kit of factor-referenced cognitive tests* (Vol. 102). Princeton, NJ: Educational Testing Service.
- Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*, 19–23. <http://dx.doi.org/10.1111/1467-8721.00160>
- Eppinger, B., Hämmerer, D., & Li, S. C. (2011). Neuromodulation of reward-based learning and decision making in human aging. *Annals of the New York Academy of Sciences*, *1235*, 1–17. <http://dx.doi.org/10.1111/j.1749-6632.2011.06230.x>
- Gazzaley, A., Cooney, J. W., Rissman, J., & D’Esposito, M. (2005). Top-down suppression deficit underlies working memory impairment in normal aging. *Nature Neuroscience*, *8*, 1298–1300. <http://dx.doi.org/10.1038/nn1543>
- Gazzaley, A., & Nobre, A. C. (2012). Top-down modulation: Bridging selective attention and working memory. *Trends in Cognitive Sciences*, *16*, 129–135. <http://dx.doi.org/10.1016/j.tics.2011.11.014>
- Grant, D. A., & Berg, E. A. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology*, *38*, 404–411. <http://dx.doi.org/10.1037/h0059831>
- Hannah, S. D., Allan, L. G., & Young, M. E. (2012). Age differences in contingency judgement linked to perceptual segregation. *The Quarterly Journal of Experimental Psychology*, *65*, 1195–1213. <http://dx.doi.org/10.1080/17470218.2011.649293>
- Hartley, A. A., Kieley, J. M., & Slabach, E. H. (1990). Age differences and similarities in the effects of cues and prompts. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 523–537. <http://dx.doi.org/10.1037/0096-1523.16.3.523>
- Holland, P. C., & Maddux, J. M. (2010). Brain systems of attention in associative learning. In C. J. Mitchell & M. E. Le Pelley (Eds.), *Attention and learning* (pp. 305–349). New York, NY: Oxford University Press.
- Howard, J. H., Jr., Howard, D. V., Dennis, N. A., Yankovich, H., & Vaidya, C. J. (2004). Implicit spatial contextual learning in healthy aging. *Neuropsychology*, *18*, 124–134. <http://dx.doi.org/10.1037/0894-4105.18.1.124>
- Isaacowitz, D. M., Wadlinger, H. A., Goren, D., & Wilson, H. R. (2006). Is there an age-related positivity effect in visual attention? A comparison of two methodologies. *Emotion*, *6*, 511–516. <http://dx.doi.org/10.1037/1528-3542.6.3.511>
- Ishihara, S. (2006). *The series of plates designed as a test for color deficiency*. Tokyo, Japan: Kanehara Trading, Inc.
- Jiang, Y. V., Koutstaal, W., & Twedell, E. L. (2016). Habitual attention in older and young adults. *Psychology and Aging*, *31*, 970–980. <http://dx.doi.org/10.1037/pag0000139>
- Kamin, L. J. (1969). Predictability, surprise, attention and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior* (pp. 279–296). New York, NY: Appleton-Century-Crofts.
- Kim, S., & Rehder, B. (2011). How prior knowledge affects selective attention during category learning: An eyetracking study. *Memory & Cognition*, *39*, 649–665. <http://dx.doi.org/10.3758/s13421-010-0050-3>
- Kruschke, J. K. (2003). Attention in learning. *Current Directions in Psychological Science*, *12*, 171–175. <http://dx.doi.org/10.1111/1467-8721.01254>
- Kruschke, J. K. (2009). Attentional highlighting: A canonical experiment. *Psychology of Learning and Motivation*, *51*, 153–185. [http://dx.doi.org/10.1016/S0079-7421\(09\)51005-5](http://dx.doi.org/10.1016/S0079-7421(09)51005-5)
- Kruschke, J. K., Kappenman, E. S., & Hetrick, W. P. (2005). Eye gaze and individual differences consistent with learned attention in associative blocking and highlighting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 830–845. <http://dx.doi.org/10.1037/0278-7393.31.5.830>
- Le Pelley, M. E. (2004). The role of associative history in models of associative learning: A selective review and a hybrid model. *The Quar-*

- terly *Journal of Experimental Psychology B: Comparative and Physiological Psychology*, 57, 193–243. <http://dx.doi.org/10.1080/02724990344000141>
- Le Pelley, M. E., Beesley, T., & Griffiths, O. (2011). Overt attention and predictiveness in human contingency learning. *Journal of Experimental Psychology: Animal Behavior Processes*, 37, 220–229. <http://dx.doi.org/10.1037/a0021384>
- Le Pelley, M. E., Vadillo, M., & Luque, D. (2013). Learned predictiveness influences rapid attentional capture: Evidence from the dot probe task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39, 1888–1900. <http://dx.doi.org/10.1037/a0033700>
- Levine, B., Stuss, D. T., & Milberg, W. P. (1997). Effects of aging on conditional associative learning: Process analyses and comparison with focal frontal lesions. *Neuropsychology*, 11, 367–381. <http://dx.doi.org/10.1037/0894-4105.11.3.367>
- Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, 82, 276–298. <http://dx.doi.org/10.1037/h0076778>
- Mackintosh, N. J., & Turner, C. (1971). Blocking as a function of novelty of CS and predictability of UCS. *The Quarterly Journal of Experimental Psychology*, 23, 359–366. <http://dx.doi.org/10.1080/14640747108400245>
- Mathot, S. (2017, May 15). *Bayes like a baw: Interpreting Bayesian repeated measures in JASP* [Web log post]. Retrieved from <https://www.cogsci.nl/blog/interpreting-bayesian-repeated-measures-in-jasp>
- McCabe, D. P., Roediger, H. L., III, McDaniel, M. A., Balota, D. A., & Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: Evidence for a common executive attention construct. *Neuropsychology*, 24, 222–243. <http://dx.doi.org/10.1037/a0017619>
- Merrill, E. C., Conners, F. A., Roskos, B., Klinger, M. R., & Klinger, L. G. (2013). Contextual cueing effects across the lifespan. *The Journal of Genetic Psychology: Research and Theory on Human Development*, 174, 387–402. <http://dx.doi.org/10.1080/00221325.2012.694919>
- Mitchell, C. J., Griffiths, O., Seetoo, J., & Lovibond, P. F. (2012). Attentional mechanisms in learned predictiveness. *Journal of Experimental Psychology: Animal Behavior Processes*, 38, 191–202. <http://dx.doi.org/10.1037/a0027385>
- Mutter, S. A., Atchley, A. R., & Plumlee, L. M. (2012). Aging and retrospective reevaluation of causal learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38, 102–117. <http://dx.doi.org/10.1037/a0024851>
- Mutter, S. A., DeCaro, M. S., & Plumlee, L. F. (2009). The role of contingency and contiguity in young and older adults' causal learning. *The Journals of Gerontology: Series B: Psychological Sciences and Social Sciences*, 64, 315–323. <http://dx.doi.org/10.1093/geronb/gbp004>
- Mutter, S. A., Haggblom, S. J., Plumlee, L. F., & Schirmer, A. R. (2006). Aging, working memory, and discrimination learning. *The Quarterly Journal of Experimental Psychology*, 59, 1556–1566. <http://dx.doi.org/10.1080/17470210500343546>
- Mutter, S. A., & Poliske, R. M. (1994). Aging and illusory correlation in judgments of co-occurrence. *Psychology and Aging*, 9, 53–63. <http://dx.doi.org/10.1037/0882-7974.9.1.53>
- Mutter, S. A., & Plumlee, L. F. (2014). The effects of age on associative and rule-based causal learning and generalization. *Psychology and Aging*, 29, 173–186. <http://dx.doi.org/10.1037/a0035930>
- Mutter, S. A., Strain, L. M., & Plumlee, L. F. (2007). The role of age and prior beliefs in contingency judgment. *Memory & Cognition*, 35, 875–884. <http://dx.doi.org/10.3758/BF03193462>
- Mutter, S. A., & Williams, T. W. (2004). Aging and the detection of contingency in causal learning. *Psychology and Aging*, 19, 13–26. <http://dx.doi.org/10.1037/0882-7974.19.1.13>
- Nassar, M. R., Bruckner, R., Gold, J. I., Li, S. C., Heekeren, H. R., & Eppinger, B. (2016). Age differences in learning emerge from an insufficient representation of uncertainty in older adults. *Nature Communications*, 7, 11609. <http://dx.doi.org/10.1038/ncomms11609>
- Nasser, H. M., Calu, D. J., Schoenbaum, G., & Sharpe, M. J. (2017). The dopamine prediction error: Contributions to associative models of reward learning. *Frontiers in Psychology*, 8, 244. <http://dx.doi.org/10.3389/fpsyg.2017.00244>
- Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, 87, 532–552. <http://dx.doi.org/10.1037/0033-295X.87.6.532>
- Pearce, J. M., & Mackintosh, N. J. (2010). Two theories of attention: A review and a possible integration. In C. J. Mitchell & M. E. Le Pelley (Eds.), *Attention and associative learning: From brain to behavior* (pp. 11–39). New York, NY: Oxford University Press.
- Raftery, A. E. (1995). Bayesian model selection in social research. In P. V. Marsden (Ed.), *Sociological methodology* (pp. 111–196). Cambridge, MA: Blackwell.
- Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's Progressive Matrices and Mill Hill Vocabulary Scales: Section 5. The Mill Hill Vocabulary Scale*. San Antonio, TX: Pearson Education.
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current theory and research* (pp. 64–99). New York, NY: Appleton-Century-Crofts.
- Reuter-Lorenz, P. A., Festini, S. B., & Jantz, T. K. (2016). Executive functions and neurocognitive aging. In K. W. Schaie & S. L. Willis (Eds.), *Handbook of the psychology of aging* (8th ed., pp. 245–262). Boston, MA: Elsevier. <http://dx.doi.org/10.1016/B978-0-12-411469-2.00013-3>
- Schmitz, T. W., Cheng, F. H., & De Rosa, E. (2010). Failing to ignore: Paradoxical neural effects of perceptual load on early attentional selection in normal aging. *The Journal of Neuroscience*, 30, 14750–14758. <http://dx.doi.org/10.1523/JNEUROSCI.2687-10.2010>
- Schultz, W., & Dickinson, A. (2000). Neuronal coding of prediction errors. *Annual Review of Neuroscience*, 23, 473–500. <http://dx.doi.org/10.1146/annurev.neuro.23.1.473>
- Sewell, D. K., & Lewandowsky, S. (2012). Attention and working memory capacity: Insights from blocking, highlighting, and knowledge restructuring. *Journal of Experimental Psychology: General*, 141, 444–469. <http://dx.doi.org/10.1037/a0026560>
- Spieler, D. H., Balota, D. A., & Faust, M. E. (1996). Stroop performance in healthy younger and older adults and in individuals with dementia of the Alzheimer's type. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 461–479. <http://dx.doi.org/10.1037/0096-1523.22.2.461>
- Steinberg, E. E., Keiflin, R., Boivin, J. R., Witten, I. B., Deisseroth, K., & Janak, P. H. (2013). A causal link between prediction errors, dopamine neurons and learning. *Nature Neuroscience*, 16, 966–973. <http://dx.doi.org/10.1038/nn.3413>
- Störmer, V., Eppinger, B., & Li, S. C. (2014). Reward speeds up and increases consistency of visual selective attention: A lifespan comparison. *Cognitive, Affective & Behavioral Neuroscience*, 14, 659–671. <http://dx.doi.org/10.3758/s13415-014-0273-z>
- Wechsler, D. (1997). *Wechsler Adult Intelligence Scale* (3rd ed.). San Antonio, TX: Psychological Corporation.

Received February 5, 2018

Revision received May 22, 2018

Accepted May 30, 2018 ■