They Can Take a Hint: Older Adults Effectively Integrate Memory Cues During Recognition

Alex Konkel, Diana Selmeczy, and Ian G. Dobbins
Washington University

Adaptively biasing recognition judgments in light of environmental cues improves net accuracy. Based on previous work suggesting that strategically shifting biases on a trial-wise basis should be cognitively demanding, the authors predicted that older adults would not achieve the same accuracy benefits from environmental cues as the young. However, despite showing clear declines in cognitive control as indexed by complex span, older adults demonstrated similar accuracy gains and similar alterations of response probabilities with cues of 75% reliability (Experiment 1) and more complex cues spanning 3 levels of reliability (Experiment 2). Despite preserved gains in accuracy, older adults clearly demonstrated disproportionate slowing that was specific to trials in which cues were invalid. This slowing may reflect impairments in behavioral inhibition that could impinge upon accuracy were responding increasingly sped and future work manipulating response speed and measures of inhibition may yield further insights.

Keywords: recognition memory, aging, criterion, cognitive control, metacognition

Supplemental materials: http://dx.doi.org/10.1037/pag0000058.supp

Memory failures range from annoying, such as misplacing the TV remote, to serious, such as having difficulty remembering if medication was taken earlier in the day. The latter example is particularly troubling given that memory difficulties increase with age (Hasher & Zacks, 1988; Light, 1991). However, some of this decline in memory might be offset through judicious use of environmental supports, and older adults are known to increasingly rely on external aids like notes and lists (Bolla, Lindgren, Bonacorsy, & Bleecker, 1991) and to generally turn to contextual/environmental information to bolster performance (Hasher & Zacks, 1988).

One way in which older adults may rely on the environment could be through the memories of others. Continuing the medication example, an older adult may ask a spouse or other caregiver if they remember the pill being taken, attempting to combine or integrate these two imperfect memories of the prior event. It is unclear how effective older adults might be when integrating the judgments of others into their own memory decisions, which is the focus of the present work.

Experimental, this situation has been examined using an explicit memory cueing (EMC) task (Dobbins, Jaeger, Stader, & Simons, 2012; Jaeger, Konkel, & Dobbins, 2013; Jaeger, Cox, & Dobbins, 2012; Jaeger, Lauris, Selmeczy, & Dobbins, 2012; O'Connor, Han, & Dobbins, 2010; Selmeczy & Dobbins, 2013). This paradigm is similar to a typical recognition memory task except that during test, most recognition probes are immediately preceded by a cue informing the participant that the upcoming item is “likely old” or “likely new” (uncued trials are also intermixed to assess baseline recognition memory). The cues are probabilistically correct but above-chance and thus can be used to increase performance and young adults have been shown to appreciably increase their net recognition accuracy when such cues are available.

The improvement in performance for cued versus uncued recognition is anticipated by a simple signal detection (Macmillan & Creelman, 2005) decision model in which observers use the cues to bias recognition decisions on a trial-by-trial basis through repeated alterations of their recognition decision criterion (see Selmeczy & Dobbins, 2013 for more detail). Under this model, observers use a lax/liberal recognition criterion when cued “likely old,” a neutral criterion when uncued, and a strict/conservative criterion when cued “likely new.”

There are reasons to believe that the use of external cues during recognition, as required by the EMC paradigm, may be a demanding decision strategy. First, effective use of the cues is related to metamnemonic ability (Selmeczy & Dobbins, 2013); participants with better insight into the quality of their memory have larger gains in performance on cued trials compared to uncued trials. This result suggests that they are jointly considering the veracity of their own memory evidence and the recommendation of the cue, and adjudicating these on every trial. Second, the EMC paradigm uses a randomized presentation of the cues and thus the participant must...
continuously alter the decision criterion while carrying out the recognition judgments throughout testing. In paradigms that do not involve trial-wise cueing, it appears that participants will strategically alter their criterion to try to take advantage of manipulations of base rates (Estes & Maddox, 1995; Kantner & Lindsay, 2010; Rhodes & Jacoby, 2007) or to comply with experimenter instructions to avoid certain types of errors (Miller, Handy, Cutler, Inati, & Wolford, 2001; Postma, 1999). Criterion shifting is also described as a strategic alteration even when the experimenters do not use overt or explicit methods to induce shifts in participants; for example, Benjamin and Bawa (2004) noted that participants “strategically adjust their recognition criterion in response to assessments of relative test difficulty” (p. 165). Dobbins and Kroll (2005) found that participants will shift criteria as a function of the perceived subjective memorability of each recognition probe under self-paced testing conditions but fail to do so under speeded response conditions, which is also consistent with the perspective of shifting as strategic and requiring resources to implement.

However, some studies have failed to find criterion shifting in situations in which it would improve accuracy, and in these cases, it is often assumed that participants find shifting too effortful or demanding to implement on a trial-wise basis. For example, Wixted and Stretch (2000) examined the case for a criterion shift account of false memory performance and summarized the criterion shifting literature in general as showing that participants are unwilling to shift their criterion from trial to trial. The authors note that in such cases, “dozens of criterion shifts would be required throughout the course of a single recognition test. That kind of mental effort, which requires assessing the status of each item and adjusting the criterion accordingly, may be something that participants are generally unwilling to exert” (Wixted & Stretch, 2000, p. 376).

Overall, the extant literature seems to suggest that repeated (and appropriate) criterion shifting is a cognitively taxing, strategic, process. In particular, it may rely on working memory capacity as participants must maintain and evaluate many relevant criteria with respect to the contents of memory during the recognition task.

Despite the widespread assumption that criterion regulation may be cognitively demanding, little research has attempted to directly relate criterion regulation to potentially relevant skills such as cognitive control or working memory. To the extent that criterion regulation is cognitively taxing, it should be impaired in groups (such as older adults) with known cognitive control/working memory deficits as indexed by tasks such as complex span. However, the literature on aging and criterion regulation is sparse and mixed. In terms of test-wide criterion placement, older adults have been found to be generally more liberal (Huh, Kramer, Gazzaley, & Delis, 2006), generally more conservative (Poon & Fozard, 1980), or to have the same criterion placement (Baron & Surdy, 1990) as young adults. Similarly, older adults have been found to shift their criterion as adaptively as (Criss, Aue, & Kılıç, 2014; Pendergrass, Olffman, Schmalstig, Seder, & Light, 2012) and less adaptively than (Baron & Surdy, 1990) young adults. However, these studies only examined overall criterion placement during the course of a single test, or they involved a very small number of potential shifts across test sessions. In contrast, the current study examines the ability of participants to continually adjust the criterion, on a trial-wise basis, to effectively incorporate external recommendations into their recognition memory judgments.

Given that older adults typically have declines in cognitive control abilities relative to young adults (Hasher & Zacks, 1988; Lustig, Hasher, & Zacks, 2007; Park, Polk, Mikels, Taylor, & Marshuetz 2001), we predicted that they would also have difficulty with criterion regulation and may not benefit from external cues as much as young adults during recognition judgments. Specifically, we predicted that a breakdown in cognitive control could lead to older adults deciding that integrating the cue with their memory is too difficult and result in either not shifting at all (e.g., Wixted & Stretch, 2000) or simply parroting the cue. In either case, older adults’ performance should be better than either of their baseline performance or the cue in isolation because they are integrating two sources of information. Thus overall, the assumption based on prior literature that criterion regulation is quite cognitively demanding along with the assumed working memory capacity differences across the elderly and young led to a prediction that older adults would be noticeably impaired in the EMC task.

Finally, an additional question of interest was whether metacognitive monitoring would be related to the ability to benefit from cues in the elderly. We expected to replicate our previous findings and demonstrate that younger adults with higher metacognitive ability would benefit more from environmental cues (Selmeczy & Dobbins, 2013); however, a prediction for older adults was less clear. Although older adults have demonstrated intact monitoring for various metacognitive judgments including judgments of learning and feeling of knowing (Hertzig & Dunlosky, 2011), they have shown worse monitoring performance in other memory domains such as cued recall (Kelley & Sahakyan, 2003; Rhodes & Kelley, 2005). In addition, Dodson, Bawa, and Krueger (2007) showed that older adults demonstrated intact monitoring during item recognition but declines in monitoring during source identification mainly because of high confidence errors. Thus, since the current study uses item recognition it is possible that older adults may demonstrate intact monitoring; however, given the previously outlined prediction that older adults will show deficits in cue integration we predicted that the relationship between metacognitive monitoring and cueing benefit might be weakened or nonexistent for older adults.

**Experiment 1**

**Methods**

**Participants.** Thirty-three young adults from Washington University participated for fulfillment of course requirements or compensation of $10 per hour (average age = 20.79, age range = 18–35; 25 female). Twenty-eight older adults (average age = 71.8, age range = 65–78; 20 female) were recruited from the department’s older adult participant pool and were compensated $10 per hour.

**Materials.** Four hundred words were randomly selected for each participant from a larger pool of over 1,200 words with an average of seven letters, 2.3 syllables, and a log HAL frequency of 7.74.

**Procedure.** After giving informed consent and completing a brief demographics questionnaire, the study was verbally de-
scribed with more detailed instructions presented via computer over the course of the experiment. The experiment was generated using MatLab and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997).

During the study, 100 words were serially presented and the participants were asked to count the number of syllables in each word (one, two, or three or more) while remembering the words for a later memory test. Each study trial consisted of a 1,000 ms blank screen, the question “How many syllables does this word have?” for 500 ms, and then the word appeared along with a reminder of the response options until a response was entered. Syllable counting was used to promote moderate levels of performance and help control individual differences in encoding strategies.

At the beginning of the test phase, participants were informed that they would see serially presented and randomly intermixed studied (old) or novel (new) words, with some preceded by a computer-generated hint (likely old or likely new) that was 75% accurate. It was emphasized that the hints were useful and should be used to help their performance. In a self-paced manner participants first indicated whether words were studied or not, followed by a confidence rating from 50% (guessing) to 100% (absolutely sure) in 10% increments. The test consisted of all 100 studied words plus 100 novel lures. Forty of these trials (20 old and 20 new) were uncued trials where no hint was given (the screen displayed “????”). The remaining 160 trials were cued with 75% of the cues accurately forecasting the upcoming memory probes and 25% inaccurately forecasting the probe. Each test trial consisted of a 1,000 ms blank screen followed by the cue for 1,000 ms, and then the probe along with the cue until the confidence rating was completed.

Following the test, participants saw a screen encouraging them to take a short break until they pressed a button to proceed. They then completed a second study/test block with a new random selection of words yielding a total of 400 test trials, 200 of which constituted old items.

After completing the cued recognition task, participants completed an automated OSPAN task (Unsworth, Heitz, Schrock, & Engle, 2005) presented using ePrime software (Psychology Software Tools, Pittsburgh, PA). OSPAN was collected to ensure that participants’ working memory was estimated using a traditional absolute OSPAN score, which is the sum of all the perfectly recalled sets (Unsworth et al., 2005). As expected, older adults showed a large (Cohen’s $d = 1.30$) decline in OSPAN score relative to young adults, 24.8 vs. 46.5, $t(59) = 5.07, p < .001$. Thus, our older sample demonstrates the typical pattern of working memory decline.

### Conditional response proportions

Although the prior literature has focused on criterion and the Signal Detection framework, we wished to demonstrate that the results are not reliant on the assumptions of that model (see also Jaeger, Cox, & Dobbins, 2012). To maintain consistency, however, commonly used signal detection measures are also reported in Table 1 and all of the results below obtain when using $d’$ and $C$ instead of relative proportions.

To examine the influence of the cues we calculated the proportion of correct responses for old (hit) and new (correct rejection) items across the cued conditions, relative to baseline performance. For example, under the likely old cues applied to old materials, we first calculated the hit rate during these validly cued trials and then subtracted the hit rate during baseline, uncued performance. Thus the sign of the scores represents either the decline or increase in accuracy driven by the cues, relative to baseline performance. The findings are illustrated in Figure 1 where black and gray points

### Table 1

<table>
<thead>
<tr>
<th>CR</th>
<th>HR</th>
<th>$d'$</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Younger</td>
<td>Uncued/baseline</td>
<td>.70 (.15)</td>
<td>.81 (.13)</td>
</tr>
<tr>
<td></td>
<td>Cued</td>
<td>.77 (.10)</td>
<td>.84 (.08)</td>
</tr>
<tr>
<td></td>
<td>Likely old cue</td>
<td>.54 (.19)</td>
<td>.90 (.06)</td>
</tr>
<tr>
<td></td>
<td>Likely new cue</td>
<td>.84 (.10)</td>
<td>.68 (.17)</td>
</tr>
<tr>
<td>Older</td>
<td>Uncued/baseline</td>
<td>.83 (.11)</td>
<td>.74 (.16)</td>
</tr>
<tr>
<td></td>
<td>Cued</td>
<td>.85 (.08)</td>
<td>.81 (.12)</td>
</tr>
<tr>
<td></td>
<td>Likely old cue</td>
<td>.69 (.16)</td>
<td>.85 (.11)</td>
</tr>
<tr>
<td></td>
<td>Likely new cue</td>
<td>.90 (.08)</td>
<td>.70 (.19)</td>
</tr>
</tbody>
</table>

Note. CR = correct rejection rate; HR = hit rate.
reflect valid invalid cueing trials respectively (for raw proportions see Table 1). The plot demonstrates that both groups were clearly influenced by the cues, demonstrating both valid cueing gains and invalid cueing losses. However, the young group appeared more robustly influenced than the older adults. This can be seen by examining the influence of valid versus invalid cues for new materials (left two points in each panel) across the groups, and analogously the influence of valid versus invalid cues for old materials (right two points within each panel) across the groups. In both cases, the size of the cueing effect appears somewhat larger for the young group.

These impressions were confirmed with a Group (older vs. young) × Response (hit vs. correct rejection) × Cue (likely old vs. likely new) mixed analysis of variance (ANOVA) performed on the proportional gains and losses in response to the cueing. The data yielded a significant three way interaction among the factors, F(1, 59) = 5.04, MSE = 0.02, $\eta^2_g = .08, p = .03$. To decompose this effect we examined correct rejections and hits separately. For correct rejections, a Group × Cue mixed ANOVA yielded a two-way interaction, $F(1, 59) = 5.50$, MSE = 0.01, $\eta^2_g = .09, p = .02$, demonstrating that the change in correct rejection rates as a function of cueing was larger for the young than older adults. Post hoc pairwise comparisons demonstrated that gains with valid cueing were larger for the younger adults ($p < .001$), whereas losses with invalid cueing were comparable across groups ($p = .52$) via Fisher’s least significance difference. An analogous Group × Cue mixed ANOVA on hit rates trended toward a significant interaction, $F(1, 59) = 3.04$, MSE = 0.01, $\eta^2_g = .05$ $p = .09$, and follow-up pairwise comparisons revealed greater losses for the young versus older groups ($p = .002$) but no difference in gains ($p = .51$). Thus overall, the data suggest moderate differences in the conditional response proportions. Although both groups are clearly influenced by the cues the younger adults gained somewhat more for valid cueing of new materials but they also lost somewhat more for the invalid cueing of old materials. Thus they appeared to be somewhat more responsive to the likely new cue than the older adults.

Overall, conditional response proportions demonstrate that both groups are reliably influenced by the cues; however, the young adults appear influenced to a larger degree than the older adults based solely on these proportions. Finally, it is important to note that Figure 1 suggests a net improvement in performance under the cues for both groups even though the gains and losses appear roughly centered on zero. This is because the valid cues (black points) occur on 75% of the trials whereas the invalid cues (gray points) appear on only 25% of the trials. We confirm this by considering net accuracy below. In addition, both Figure 1 and Table 1 demonstrate that the observers are not simply parroting the cues, if observers were parroting the cues, accuracy would be at ceiling and floor for valid and invalid trials respectively, which is clearly not the case.

**Net response accuracy.** The analysis above conditionlizes accuracy as a function of the item and cue combinations. As noted however, invalid cueing only occurs on a minority (25%) of trials and thus the net accuracy of an observer can be quite good even with a noticeable decline in performance during invalid cueing. Here we compare each participant’s net success rate across uncued and cued conditions to determine whether success was more likely under the former and whether this was comparable across the groups. A Group × Cue Provision (cued or uncued) mixed ANOVA on the proportion of correct responses did not yield a significant group effect, $F(1, 59) = 2.07$, MSE = 0.01, $\eta^2_g = .03$,
interaction, 

rately. For correct rejections, there was a robust Group
demonstrating the greater alteration of median RT as a function of

Median reaction time during correct responding. The condi-
tional and net accuracy data suggest only a modest differ-
ence between the groups during the cueing paradigm. Both groups were
using the cues and improve net performance by about 5% when cues are present versus absent. However, the older adults appeared
somewhat less influenced by the cues, raising the possibility that they used the cues more cautiously. If so, this might be evident in
their RT data. To examine this we considered the median RT of
correct responses for each participant under the various cueing
conditions, analogous to the conditional accuracy analysis above.

Again, the data are relative to baseline performance but raw RTs are presented in Table 2. Thus the likely new (R) condition for new
items represents the difference between median RT for correct rejections under a valid Likely New cue and uncued, baseline
correct rejections; a situation in which one would expect quicker responding and hence negative value. The data are illustrated in
Figure 2. Across both groups there is evidence for slowing during invalidly (gray points) cued trials relative to baseline and some-
what less prominent speeding during validly (black points) cued trials relative to baseline. However, these cueing effects were
much more pronounced for the older adults, particularly in the case of the slowing incurred by invalid cueing. These impressions were
tested via a Group (older vs. young) × Response (hit vs. correct rejection) × Cue (likely old vs. likely new) mixed ANOVA that
yielded a significant three-way interaction, $F(1, 59) = 12.41$, $MSe = 0.62$, $\eta_p^2 = .17$, $p < .001$. This interaction was broken
down by considering the performance during correct rejections (left two points in each panel) and hits (right two points) sepa-
rately. For correct rejections, there was a robust Group × Cue interaction, $F(1, 59) = 9.26$, $MSe = 0.56$, $\eta_p^2 = .14$, $p = .003$, demonstrating the greater alteration of median RT as a function of cueing in the older adults. Turning to RT during hits, a Group × Cue mixed ANOVA again demonstrated a significant interaction, $F(1, 59) = 11.35$, $MSe = 0.24$, $\eta_p^2 = .16$, $p = .001$, because of the larger differential response of the older adults compared to the young adults.1

The contribution of metacognitive awareness to cueing gains. As previously demonstrated (Selmeczy & Dobbins, 2013), metacognitive ability is a significant predictor of cue ben-

We followed the hierarchical regression analysis reported in Selmeczy and Dobbins (2013). Cued and uncued accuracy was measured using the net percent correct under these conditions. In Step 1, we predicted cued accuracy using uncued accuracy in order to control for individual differences in baseline skill. Importantly, in Step 2 gamma was added to see if it predicted any additional variance in cued performance that could not be accounted for by baseline recognition. In Step 3 age group was added and in Step 4 higher order interactions with age group were added to examine age related differences. In Step 1, uncued accuracy was a significant predictor of cued performance, $b = .61, t(56) = 10.11, p < .001$. Critically, adding gamma in Step 2 significantly increased fit, explaining an additional 6.3% of the variance—$F(1, 55) = 12.02$, $p = .001$; increase in $R^2$ squared using Somers D = 5.1%—suggesting that metacognitive monitoring is a significant predictor of cueing benefit. Furthermore, in Step 3 adding the main effect of age group did not increase fit ($r < 1$) and in Step 4 age group also did not interact with any other variables ($ns < 1$). Thus replicating previous work, metacognitive monitoring was a significant predic-
tor of cueing benefit and additionally the magnitude of this relation-
ship was similar across age groups. Further analyses of con-
fidence demonstrate largely similar patterns of mean confidence

<table>
<thead>
<tr>
<th>CR</th>
<th>Hit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uncued/baseline</td>
<td>1.91 (.51)</td>
</tr>
<tr>
<td>Likely old cue</td>
<td>2.30 (.83)</td>
</tr>
<tr>
<td>Likely new cue</td>
<td>1.71 (.45)</td>
</tr>
<tr>
<td>Older</td>
<td>3.22 (1.10)</td>
</tr>
<tr>
<td>Likely old cue</td>
<td>4.18 (1.90)</td>
</tr>
<tr>
<td>Likely new cue</td>
<td>2.76 (.82)</td>
</tr>
</tbody>
</table>

Note. CR = correct rejection rate.

1 Although Figure 2 suggests a very prominent slowing during invalid cueing for the older adults, one possibility is that the slowing is nonetheless similar if one takes differences in baseline RTs into account. For example, if all individuals slowed by 25% relative to their baseline speed in the face of invalid cueing, then the slowing of the of the older adults would appear much larger because their baseline RTs are also much larger. To address this, we converted the slowing during invalid cueing into a proportional score; namely, (invalid RT−baseline RT)/baseline RT using the median values for each group during correct responding. Using this measure, older adults were disproportionately slowed during the invalid cueing of old materials, $t(59) = 2.33, p = .023$ but not the invalid cueing of new materials ($p > .34$).
across age groups as well, and are described in the Supplemental Materials.

Summary

The data demonstrate that although both groups are actively and effectively using the cues, and hence demonstrating similar net performance increases when cued versus uncued, they differed somewhat in how these cues altered behavior. Whereas the cue influences were more prominent in conditional response probabilities for the young than the older adults, the reverse occurred in the RT data, where invalid cue-induced slowing was particularly prominent and disproportionate in the older adults. The modest differences of the groups in response probabilities were in contrast to the OSPAN performance, which demonstrated that the older participants had markedly lower working memory abilities. If the ability to dynamically integrate cues during recognition was dependent upon the types of control abilities indexed by the OSPAN, then it seems likely that cue integration abilities should have differed more clearly across the groups. Finally, older adults also exhibited intact metacognitive monitoring and this measure was significantly related to cueing benefit similarly for both age groups.

Experiment 2

Rationale

Although there were modest differences in the RT and conditional response data of the older and younger groups, both appeared to effectively incorporate the cues in Experiment 1 and in no way did the accuracy data suggest impairment in the older adults. Because we were surprised by this initial demonstration of relatively intact cue utilization in the older adults, we designed an even stronger test of the hypothesis that adaptive incorporation of the cues taxes cognitive control, by more than doubling the complexity of the cueing through incorporating three levels of cue reliability (60%, 70%, and 80%). These reliabilities, along with the direction of cue (likely old or likely new), were again administered in a trial-wise fashion. Thus, participants were not only required to bias judgments in a particular direction, but now should also alter the degree of dependency upon the external cue if they wished to gain as much as possible from the cueing. In contrast to Experiment 1 where there were three possible cues (i.e., uncued, likely old, and likely new), the current design results in seven possible cues randomly intermixed across trials (i.e., uncued, likely old, and likely new), the current design results in seven possible cues randomly intermixed across trials (i.e., uncued, likely old, and likely new), the current design results in seven possible cues randomly intermixed across trials (i.e., uncued, likely old, and likely new), the current design results in seven possible cues randomly intermixed across trials (i.e., uncued, likely old, and likely new), the current design results in seven possible cues randomly intermixed across trials (i.e., uncued, likely old, and likely new), the current design results in seven possible cues randomly intermixed across trials (i.e., uncued, likely old, and likely new, likely new at 60%, 70%, and 80% reliabilities).

Methods

Participants. Thirty-six young adults (average age = 19.7, age range = 18–26; 21 female) were recruited from the student participant pool at Washington University; they participated in exchange for partial fulfillment of course requirements. Thirty-three older adults (average age = 70.2, age range = 53–87; 24 female) were recruited from the older adult participant pool maintained by the department; they were compensated $10 per hour of their time. The study was approved by the Human Subjects Committee and Internal Review Board at Washington University.

Materials. Five hundred twenty words were randomly selected for each participant from the same pool used in Experiment 1. These words were again randomly assigned to conditions.
Procedure. Informed consent and a demographics form were completed as in Experiment 1, and similarly the study was described in brief before the participant began the cued recognition task.

The study phase was identical to Experiment 1 except that it consisted of 130 trials. The increased number of trials was due to a desire to make sure there were a sufficient number of trials in the smallest potential bin for analysis (invalidly cued old/new trials in the 80% cue accuracy condition).

The test phase was similar to Experiment 1 in structure but consisted of 260 trials. The number of uncued trials was reduced to 20 to accommodate additional cued trials while keeping test time at a reasonable level. To clearly differentiate the levels of cue reliability for the participants, the cue reliability (60%, 70%, or 80%) was indicated in the cue prompt (e.g., “likely new 70%”).

The different cue levels were also presented in different font sizes (larger for more accurate) and colors to further make the reliability distinctions salient to the participants. Eighty cued trials (consisting of 40 old and 40 new words) were presented at each level of cue reliability. There were two study/test blocks as in Experiment 1.

Following the cueing task, participants again completed the OSPAN. Because of the experimental time constraint (two hours) and the longer cued recognition task, not all participants completed the OSPAN. Thirty-five of 36 young adults completed the OSPAN, as did 21 of 33 older adults. In addition, to further characterize our older adult sample, participants also completed a sustained attention to response task (SART) (Jackson & Balota, 2011; see the Supplemental Material for more details).

Results

Once again preliminary analyses examined participants’ working memory estimated using a traditional absolute OSPAN score (i.e., sum of all the perfectly recalled sets; Unsworth et al., 2005).

As expected, older adults showed a large (Cohen’s $d = 1.53$) and significant decline in OSPAN score relative to young adults, 30.4 vs. 53.8, $t(53) = 5.76, p < .001$. The large group differences found in OSPAN again serve to demonstrate that our older adult sample exhibited the typical age related declines in working memory.

Conditional response proportions. Cue influences were again examined by calculating the change in correct response rates for old (hit) and new (correct rejection) items across the cued conditions relative to baseline performance. The findings are illustrated in Figure 3 with each row corresponding to a different level of cue reliability (raw proportions are in Table 3). The plot demonstrates that both groups were clearly influenced by the cues demonstrating both valid cueing gains and invalid cueing losses and the degree of influence does not appear to noticeably differ across the groups. The plot also suggests a tendency for the difference between gains on validly cued trials (black points) and losses on invalidly cued trials (gray points) to increase with increasing cue reliability, which is expected if participants appropriately rely more heavily on more reliable cues.

These impressions were confirmed with a four-way mixed ANOVA with factors of group (older vs. young), response (hit vs. correct rejection), cue direction (likely old vs. likely new), and cue reliability (60%, 70%, or 80%). Criticality, the group factor failed to reliably interact with the other factors in isolation or combination (all $Fs < 1$) except in one case; namely, a Group × Response interaction, $F(1, 66) = 4.01, MSe = 0.10, \eta_p^2 = .06, p = .05$. This resulted because the average conditional correct rejection scores relative to baseline were lower for the older than younger adults ($p = .04$), but there was no reliable difference in average conditional hit scores. However, because this interaction collapses across both the reliability of the cues and their validity, it is difficult to meaningfully interpret and incidental to our main interest in the cueing phenomena. Aside from this there were only two other significant outcomes. There was a reliable interaction of Response × Cue Direction, $F(1, 66) = 113.51, MSe = 0.16, \eta_p^2 = .63, p < .001$, that simply reflects the fact that correct rejections and hits were modulated in a different direction by the likely old and likely new cues. In other words, observers were sensitive to cue direction. In addition, the size of this modulation was influenced by the stated reliability of the presented cues, yielding a significant Response × Cue Direction × Cue Reliability interaction, $F(2, 132) = 23.29, MSe = 0.01, \eta_p^2 = .26, p < .001$. Thus, as suggested in Figure 3, the difference between cue-induced gains and losses is increased as the cues become more reliable.

Overall, the data demonstrate a robust modulation of performance by the cues dependent upon their validity and a smaller but reliable modulation as a function of cue reliability. However, as Figure 3 demonstrates, conditional accuracy scores are quite similar across the older and younger groups.

Net response accuracy. As noted earlier, because invalid cueing only occurs on a minority of trials, net performance can be quite good even with marked declines during invalid cueing. Here we compare each participant’s net success rates across the four types of cues; namely uncued, 60%, 70%, and 80% cue reliability conditions. A $2 \times 4$ Group × Cue Reliability mixed ANOVA on the total proportion of correct responses under each cue condition did not yield a significant group effect ($F < 1$) nor did group interact with cue reliability ($F < 1$). Instead, there was only a large main effect of cue reliability, $F(3, 198) = 32.84, MSe = 0.003, \eta_p^2 = .33, p < .001$, demonstrating greater net success for increasingly reliable cues (see Figure 4). Thus it is clear that participants are appropriately increasingly weighting the cue information as its stated reliability increases.

Median reaction time during correct responding. Neither conditional nor net accuracy data suggested any notable differences between the age groups. However, Experiment 1 demonstrated disproportionate slowing during invalid cueing for the older adults so we again examined the median RTs of correct responses in relation to baseline performance (see Figure 5). As with the analysis of conditional response proportions we focused on group (older vs. young), response (hit vs. correct rejection), cue direction (likely old vs. likely new), and cue reliability (60%, 70%, or 80%) as factors with a dependent variable of relative costs (e.g., slowing) or gains (e.g., speeding) in median RTs as a function of the cues (for raw median RTs, see Table 4). The analysis yielded main effects of group, $F(1, 63) = 8.55, MSe = 4.23, \eta_p^2 = .12, p = .005$, and cue direction, $F(1, 63) = 8.15, MSe = 0.64, \eta_p^2 = .11, p = .006$, as well as a Cue Direction × Response interaction, $F(1, 63) = 50.70, MSe = 3.01, \eta_p^2 = .45, p < .001$. The latter effect merely reflects that, on average, subjects slowed for invalid versus valid cueing. However, all of these effects were conditioned by a three-way interaction between group, cue, and response ($F(1,
63) = 14.38, MSe = 3.01, \( \eta^2_p = .19, p < .001 \). As shown in Figure 5, this occurred because older adults slow to a much larger extent than the young during invalid cueing, yet speeding appears more similar during valid cueing. Indeed when comparing the speeding that occurs with valid cueing across the groups, not a single one of six conditions (black points in Figure 5) reliably differed across the groups (all \( t \)s < 1). In contrast, every one of the six conditions reflecting invalid cueing (gray points in Figure 5) reliably differed across the groups (minimum \( t(64) = 2.51, p = .02 \). Thus, relative to their own baselines, the elderly and the young speed comparably when validly cued, but the elderly slow to a much greater extent when invalidly cued.2

The contribution of metacognitive awareness to cueing gains. As in Experiment 1, gamma was calculated from uncued trials and it was not significantly different between age groups, \( r(66) = 1.64, p = .11 \) (older adult \( M = .32 \), young adult \( M = .46 \);

2 Simple visual inspection of Figure 5 already demonstrates that the slowing of the older adults is disproportionate with respect to the young because in five of the six invalid cueing conditions, the young do not appear to reliably slow at all. Nonetheless, we directly compared the groups across all six conditions using proportional slowing scores as in Experiment 1. Once again, we converted the slowing during invalid cueing into a proportional score; namely, (invalid RT-baseline RT)/baseline RT using the median values for each group across each cue reliability during correct responding. For invalidly cued old items, the \( t \) tests from lowest to highest cue reliability revealed disproportionate slowing, \( t(65) = 3.20, p = .002; t(65) = 2.60, p = .01; t(65) = 2.47, p = .02 \). For invalidly cued new items, \( t \) tests from lowest to highest cue reliability also indicated disproportionate slowing, \( t(65) = 3.42, p = .001; t(65) = 3.56, p < .001; t(65) = 2.10, p = .04 \). Thus, every comparison supports the conclusion that the older adults are slowed disproportionately compared to the young.

Figure 3. Conditional gain and loss rates across the groups and cue conditions for correct responses. Black points reflect valid cue conditions and gray points indicate invalid cue conditions. New items represent correct rejections and old items represent hits. The error bars reflect ±2 standard errors of the cell mean. The condition labels are appended with “(R)” to note that these are proportions adjusted relative to uncued performance. Each row corresponds to a different level of cue reliability, which is indicated in the right margin of the plot.
Table 3
Average Response Rates, Accuracy, and Criterion With Standard Deviations in Parentheses

<table>
<thead>
<tr>
<th>CR</th>
<th>HR</th>
<th>d'</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Younger</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uncued/baseline</td>
<td>.73 (.15)</td>
<td>.73 (.15)</td>
<td>1.42 (.96)</td>
</tr>
<tr>
<td>60% cues</td>
<td>Cued</td>
<td>.75 (.12)</td>
<td>.73 (.12)</td>
</tr>
<tr>
<td>Likely old cue</td>
<td>.60 (.24)</td>
<td>.82 (.11)</td>
<td>1.36 (.86)</td>
</tr>
<tr>
<td>Likely new cue</td>
<td>.85 (.11)</td>
<td>.61 (.20)</td>
<td>1.52 (.85)</td>
</tr>
<tr>
<td>70% cues</td>
<td>Cued</td>
<td>.77 (.10)</td>
<td>.77 (.10)</td>
</tr>
<tr>
<td>Likely old cue</td>
<td>.58 (.26)</td>
<td>.84 (.11)</td>
<td>1.38 (.82)</td>
</tr>
<tr>
<td>Likely new cue</td>
<td>.86 (.09)</td>
<td>.61 (.19)</td>
<td>1.55 (.88)</td>
</tr>
<tr>
<td>80% cues</td>
<td>Cued</td>
<td>.80 (.09)</td>
<td>.81 (.10)</td>
</tr>
<tr>
<td>Likely old cue</td>
<td>.52 (.27)</td>
<td>.87 (.12)</td>
<td>1.37 (.78)</td>
</tr>
<tr>
<td>Likely new cue</td>
<td>.87 (.08)</td>
<td>.55 (.19)</td>
<td>1.45 (.78)</td>
</tr>
<tr>
<td>Older</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uncued/baseline</td>
<td>.80 (.13)</td>
<td>.67 (.22)</td>
<td>1.49 (.79)</td>
</tr>
<tr>
<td>60% cues</td>
<td>Cued</td>
<td>.76 (.11)</td>
<td>.71 (.13)</td>
</tr>
<tr>
<td>Likely old cue</td>
<td>.61 (.24)</td>
<td>.82 (.15)</td>
<td>1.40 (.75)</td>
</tr>
<tr>
<td>Likely new cue</td>
<td>.87 (.10)</td>
<td>.56 (.22)</td>
<td>1.38 (.67)</td>
</tr>
<tr>
<td>70% cues</td>
<td>Cued</td>
<td>.79 (.10)</td>
<td>.73 (.12)</td>
</tr>
<tr>
<td>Likely old cue</td>
<td>.55 (.28)</td>
<td>.82 (.14)</td>
<td>1.23 (.69)</td>
</tr>
<tr>
<td>Likely new cue</td>
<td>.90 (.10)</td>
<td>.52 (.23)</td>
<td>1.53 (.74)</td>
</tr>
<tr>
<td>80% cues</td>
<td>Cued</td>
<td>.83 (.08)</td>
<td>.79 (.10)</td>
</tr>
<tr>
<td>Likely old cue</td>
<td>.51 (.26)</td>
<td>.87 (.12)</td>
<td>1.42 (.67)</td>
</tr>
<tr>
<td>Likely new cue</td>
<td>.91 (.07)</td>
<td>.50 (.24)</td>
<td>1.41 (.71)</td>
</tr>
</tbody>
</table>

Note. CR = correct rejection rate; HR = hit rate.

Somers D older adult M = .26, younger adult M = .38, t(66) = 1.88, p = .06.

To simplify analyses and increase reliability we collapsed across cue reliability and calculated a single cued performance score for each individual; however, the main results replicated when assessing each level of cue reliability in separate regressions. Cued and uncued accuracy was measured using the net percent correct under these conditions. As in Experiment 1 we followed the hierarchical regression analysis reported in Selmeczy and Dobbins (2013). In Step 1, uncued performance was a significant predictor of overall cued performance, b = 0.46, t(66) = 10.07, p < .001. Critically, adding gamma in Step 2 significantly increased fit, explaining an additional 5.0% of the variance, F(1, 65) = 9.42, p = .003; increase in R squared using Somers D = 9.9%, suggesting that metacognitive monitoring was a significant predictor of cueing benefit. Furthermore, in Step 3 adding the main effect of age group did not significantly increase fit (t < 1), and in Step 4 age group also did not interact with any other variables (t < 1). Thus, once again metacognitive monitoring was a significant predictor of cueing benefit and additionally the magnitude of this relationship was similar across age groups. Further analyses of confidence demonstrate largely similar patterns of mean confidence across age groups as well, and are described in the Supplemental Material.

Summary

The results converged with those of Experiment 1. Despite the greater complexity of the cueing, older and younger adults performed similarly with respect to conditional and net accuracy. Thus in terms of accuracy, the data demonstrate that the older adults incorporate dynamic cues into their recognition judgments as effectively as the young. However, the groups markedly diverged when examining RTs and older adults demonstrate a drastic and disproportionate slowing during correct recognition judgments when they are invalidly cued. Indeed, in conditions in which the young fail to demonstrate any reliable slowing, the older adults often slowed by up to 1 s. As with Experiment 1, the older adults also show a very prominent decline in working memory scores on the OSPAN demonstrating that their preserved accuracy in incorporating the cues was not an artifact of unrepresentative sparing of complex working memory span in this sample. Finally, once again metacognitive monitoring was a significant predictor of cueing benefit and this effect was similar in magnitude for both younger and older adults.

General Discussion

In contrast with our predictions, older adults were comparable to young adults in terms of using external cues to appropriately bias recognition judgments and hence elevate accuracy. Indeed, if one were to focus solely on the accuracy data, then it would be fair to characterize their performance as largely indistinguishable from that of young adults. This is positive news given how often research finds age-related cognitive deficits in both memory and executive domains (Park et al., 2001; Rhodes & Kelley, 2005) and it suggests that adaptive biasing of judgments may be a skill that is well preserved in aging. This would be a particularly important ability in memory domains where older adults evidence greater declines in accuracy, such as source and associative memory (Balota, Dolan & Duchek, 2000). It is important to emphasize that the older adults in the current study demonstrated markedly reduced scores on the OSPAN (as is typical) and thus, if the cognitive control processes captured by this ubiquitous working memory measure were necessary for dynamically altering recognition criteria in a trial-wise fashion, then older adults should have demonstrated poorer cue integration abilities. Conversely, the fact they did not suggests that the processes critical for OSPAN per-
formance are not essential for dynamic integration of recognition cues. This finding converges with very recent work by Starns and Olchowski (2014) that suggests that contrary to numerous prior characterizations, repeatedly shifting the decision criterion may not be particularly effortful or cognitively demanding in the same sense as implied by measures of cognitive control such as the OSPAN. Specifically, Starns and Olchowski (2014) demonstrated that participants can easily and repeatedly shift decision criterion when forced to explicitly process color cues signaling likely recognition strength. It is important to emphasize that dynamic regulation of the decision criterion has typically been assumed to represent a fairly effortful phenomenon. For example, criterion adjustment has been described as “effortful” (Bruno, Highman, & Perfect, 2009, p. 810) requiring “mental energy” (Stretch & Wixted, 1998, p. 1390) and as a “controlled executive processes” (Dobbins & Kroll, 2005, p. 1186). These characterizations of criterion regulation as an effortful and cognitively demanding process have primarily arisen from studies focusing solely on response accuracy and bias, and RT data have often not been considered (but see Curran, DeBuse, and Leynes, 2007). Although the current data demonstrate similar patterns of accuracy and conditional response probabilities across the age groups, it is possible that similar response rates may arise from different strategies. In the current data, older and younger adults are quite easily distinguished with respect to RT data and we next discuss how these relative slowing differences might arise.

As shown in Figures 2 and 5, the older adults much more markedly slowed correct responding when invalidly cued whereas the RT costs of young adults in this situation were fairly minimal. To further appreciate this one can consider the six invalid cueing conditions illustrated in Figure 5 (gray points). In only one of the six did the young adults reliably differ from their baseline RT, which can be confirmed by examining the plotted intervals with respect to zero. In contrast, every analogous interval for the elderly

Figure 5. Conditional speeding and slowing across the groups and cue conditions for correct responses. Black points reflect valid cue conditions and gray points indicate invalid cue conditions. New items represent correct rejections and old items represent hits. The error bars reflect ±1.2 standard errors of the cell mean. The condition labels are appended with “(R)” to note that these are median reaction times (RTs) adjusted relative to uncued performance. Each row corresponds to a different level of cue reliability, which is indicated in the right margin of the plot.
Table 4
Average Median Reaction Times With Standard Deviations in Parentheses

<table>
<thead>
<tr>
<th>CR</th>
<th>Hit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Younger Uncued/baseline</td>
<td>2.14 (.78)</td>
</tr>
<tr>
<td>Likely old cue 60% cues</td>
<td>2.18 (.93)</td>
</tr>
<tr>
<td>Likely new cue 70% cues</td>
<td>1.83 (.64)</td>
</tr>
<tr>
<td>Likely old cue 80% cues</td>
<td>2.28 (1.29)</td>
</tr>
<tr>
<td>Likely new cue 80% cues</td>
<td>1.76 (.59)</td>
</tr>
<tr>
<td>Likely old cue</td>
<td>2.47 (1.36)</td>
</tr>
<tr>
<td>Likely new cue</td>
<td>1.75 (.53)</td>
</tr>
</tbody>
</table>

Older Uncued/baseline | 3.15 (1.02) | 3.17 (1.32) |
| Likely old cue 60% cues | 4.08 (1.68) | 2.83 (1.11) |
| Likely new cue 70% cues | 2.66 (.84) | 3.75 (1.67) |
| Likely old cue 80% cues | 4.41 (2.23) | 2.77 (1.01) |
| Likely new cue 80% cues | 2.62 (.91) | 4.21 (2.55) |
| Likely old cue | 4.27 (2.01) | 2.80 (1.12) |
| Likely new cue | 2.59 (.89) | 3.94 (2.55) |

Note. CR = correct rejection rate.

participants falls well above the zero point, demonstrating they were prominently slowing in a situation in which the young were only marginally affected and this pattern unsurprisingly yielded reliable statistical differences in slowing for both difference and proportional (see Footnotes 1 and 2) RT measures.

Looking at the data in its entirety, it seems quite unlikely that the slowing phenomenon reflects a simple strategic difference in the reliance upon cues across the two groups because the change in conditional response rates is similar for both groups during both valid and invalid cueing. In contrast, if the older adults had been considerably more cautious in their cue use than the young then one would have expected generally smaller effects in the response proportions (Figures 1 and 3). Second, the older adults sped responding similarly to the young during valid cueing relative to baseline (Figures 2 and 5). If they were indeed suspicious of the cues and cautious in their use, one would expect both less speeding for valid cues and more slowing for invalid cues, whereas only the latter occurs and clearly to a disproportionate extent.

In contrast to a general or strategic difference, the slowing phenomenon of the older adults was very specific and restricted to conditions in which their memorial expectations were in conflict with the encountered recognition evidence. Importantly, they were able to overcome this discrepancy as well as the young because the losses, in terms of response accuracy, were comparable across the groups. Nonetheless, they slowed by as much as a second during this specific situation that occurs on the small minority of trials during which the generally useful cues are incorrect.

As a working hypothesis, the current data suggest a generally preserved ability to dynamically incorporate cues with perhaps an emerging deficit in behavioral inhibition. The notion that memory and other cognitive impairments seen in aging may reflect diminished inhibitory capacity is well established (e.g., Castel, Bartola, Hutcherson, Logan, & Yap, 2007; Hasher & Zacks, 1988; Lustig, Hasher, & Zacks, 2007; Rabbitt, 1965; Spieler, Balota, & Faust, 1996). However, several lines of evidence supporting this as a general impairment have not held up, or produced mixed results. For example, the widely used Stroop interference task was held to suggest an inhibitory deficit in older adults, but a meta-analysis that incorporated potential corrections for general slowing did not support this conclusion and instead suggested that the slowness of the elderly during Stroop interference was consistent with their already considerably slowed baseline performance (Verhaeghen & De Meersman, 1998). These types of findings have led to the suggestion that inhibitory deficits may be domain specific with some tasks or processes being more sensitive to decline than others (e.g., Bugg et al., 2007; Kramer et al., 1994; Rush, Barch, & Braver, 2006).

In the current study, the behavioral inhibition requirement is presumably linked to using recognition memory signals to override cue-induced expectations about one’s own subsequent memory experiences and the demand is fairly rare given the generally high reliability of the cues. This is a situation that is fairly different from the type of interference examined in the Stroop because participants are not provided with a strong prior belief of the target color or word that will be subsequently viewed and hence cannot prepare a judgment in advance. In contrast, in the current explicit memory cueing paradigm the likely status of the upcoming memory probe is forecast in advance which may be more analogous to a go–no go situation in which the judgment or response is prepared prior to the arrival of the probe. Regardless, frameworks that focus on inhibitory deficits in healthy aging (e.g., Hasher, & Zacks, 1988) anticipate that elderly participants might have difficulties in such circumstances because they are less able to rapidly use the accumulating memory evidence to suppress the inappropriate judgment bias instilled by the cues. If this is correct, we anticipate that older adults would also demonstrate an impairment in accuracy if they were pressured to respond more quickly and this impairment would be first evident in a tendency to falsely confirm invalidly cued memory expectations. Speeding manipulations would also be useful in distinguishing behavioral inhibition problems from the possibility that older adults simply repeatedly reconfirm their judgments when overtly disagreeing with environmental recommendations they know to be generally correct. Under this account, the disproportionate slowing is essentially postdecisional and would not result in markedly increased errors during speeded judgments. From this perspective then, the slowing represents a strategy specific to situations in which older adults must overtly disagree with an external information source deemed generally valid and given this, the characteristic slowing in Figures 2 and 5 would also be seen when cueing perceptual, semantic, and other judgments. Thus future work may benefit from exploring several issues using cued decision paradigms, namely, whether the requirement to speed recognition decisions reveals age-related declines in cueing accuracy benefits, whether this (or the disproportionate slowing phenomenon during self-paced judgment) is linked to independent measures of inhibitory control such as Stroop or Go–No Go, and whether these phenomena are specific to episodic judgments or extend to perception and other domains. With respect to the last consideration, it may also be informative to compare domains where older adults may have strong a priori beliefs of
age-related impairment to those in which assume their performance is preserved.

Although we have used the basic signal detection model to characterize the cues as yielding a repeatedly shifting criterion or decision bias it is important to note that neither the signal detection nor more complex decision models, such as the drift diffusion model, actually formally incorporate notions of cognitive effort or behavioral inhibition. Thus, if the disproportionate slowing phenomenon identified here is linked to behavioral inhibition processes, these processes will need to be considered separately from, or added to, current popular decision recognition decision models (see Voss, Voss, & Klauer, 2010 for related issues in drift diffusion models). In addition, it is important to note that the observation of similar adaptive biasing (in terms of accuracy) for the young and older adults here does not mean that differences may not emerge in other criterion shifting paradigms. Whereas the EMC paradigm used here provides an explicit recommendation on each trial, typical studies of base-rate manipulations instead instruct the observers (correctly or incorrectly) to expect large relative frequency differences between old and new items in the upcoming test. Under these conditions, observers must hold onto such information throughout testing, with no environmental cue on each trial highlighting the appropriate strategy, leading to the possibility that some criterion shifting studies may be more sensitive to goal neglect or distraction considerations (see Kane & Engle, 2003, for similar ideas in Stroop interference), which, like considerations of cognitive effort or behavioral inhibition, are well outside current recognition decision models. For example, future work could assess whether older adults would show greater deficits in criterion shifting compared to young adults under conditions that more heavily tax cognitive resources such as divided attention situations.

Finally, the current study found that in addition to largely intact cue integration, older adults have similar metacognitive monitoring abilities during recognition as younger adults and the relationship between metacognition monitoring and cue integration is similar across age groups. Thus, consistent with Dodson, Bawa, and Krueger (2007), our study also suggests that older adults’ monitoring (based on retrospective confidence ratings) during item recognition is largely intact. Frameworks of metacognition generally distinguish between metacognitive monitoring (i.e., reflecting on mental processes) and metacognitive control (i.e., regulating behavior based on metacognitive monitoring output; Nelson & Narens, 1990). Our current work demonstrates that older adults were able to use their monitoring in order to successfully regulate their behavior in the form of increased cue integration. However, in the current experiments the cues were always on screen making them available for every decision. Future work could more specifically examine metacognitive control and aging by requiring participants to actively ask for cues (e.g., by pressing the space bar) on a trial-by-trial basis to determine whether robust cueing benefits would still occur under such situations. In addition, whereas in the current design there was no age-related decline in metacognitive monitoring, impairments in monitoring are found in older adults when using cued recall or source memory tasks (Dodson, Bawa, & Krueger, 2007; Kelley & Sahakyan, 2003; Rhodes & Kelley, 2005). Thus, future work could assess whether the relationship between metacognitive monitoring and cueing benefit would be impaired in older adults if source judgments were cued (i.e., Likely Source A vs. Likely Source B) as opposed to simple item recognition judgments.

**Conclusion**

Older and younger adults are equally accurate at incorporating external recommendations into their recognition judgments and this finding does not support the assumption that active, trial-wise regulation of recognition decision criteria is cognitively demanding, insofar that older adults commonly have reduced cognitive control as measured by tasks such as the OSPAN. However, older adults demonstrate a remarkable slowing when invalidly cued about their upcoming memories. Because the key difference between the groups occurs specifically on trials in which demands for behavioral inhibition are high, it might be profitable to also test the idea that a decline in some form of behavioral inhibition underlies the group difference. Furthermore, because age-related declines in memory are not prominent in basic recognition but are in free-recall (Crain & McDowd, 1987), it may be the case that targeting a specific kind of inhibitory demand, namely the ability to use emergent, bottom-up memory signals to override interfering context information (e.g., a just recalled item that should not be rereported, or an expectation about an upcoming memory signal) may be a fruitful avenue to explaining the disproportionate slowing phenomenon documented here. Under this approach, older individuals who demonstrate more prominent slowing in the current paradigm would be expected to demonstrate greater proactive and output interference in recall-based paradigms.

3 Under the drift diffusion model, base rate effects are often modeled via a drift in the accumulation start point, which is an alternative way of formalizing a decision bias (see Ratcliff & Smith, 2004).

**References**


Received June 23, 2015
Revision received August 26, 2015
Accepted September 17, 2015