Healthy Aging and Visual Working Memory: The Effect of Mixing Feature and Conjunction Changes
Stephen Rhodes, Mario A. Parra, Nelson Cowan, and Robert H. Logie

CITATION
Healthy Aging and Visual Working Memory: The Effect of Mixing Feature and Conjunction Changes

Stephen Rhodes  
University of Edinburgh

Mario A. Parra  
University of Edinburgh and Heriot-Watt University

Nelson Cowan  
University of Missouri

Robert H. Logie  
University of Edinburgh

It has been suggested that an age-related decrease in the ability to bind and retain conjunctions of features may account for some of the pronounced decline of visual working memory (VWM) across the adult life span. So far the evidence for this proposal has been mixed with some suggesting a specific deficit in binding to location, while the retention of surface feature conjunctions (e.g., color-shape) appears to remain largely intact. The present experiments follow up on the results of an earlier study, which found that older adults were specifically poor at detecting conjunction changes when they were mixed with trials containing changes to individual features, relative to when these trials were blocked (Cowan, Naveh-Benjamin, Kilb, & Saults, 2006). Using stimuli defined by conjunctions of color and shape (Experiment 1), and color and location (Experiment 2) we find no evidence that older adults are less accurate at detecting binding changes when trial types are mixed. Further, analysis of estimates of discriminability provides substantial-to-strong evidence against this suggestion. We discuss these findings in relation to previous studies addressing the same question and suggest that much of the evidence for specific age-related VWM binding deficits is not as strong as it first appears.

Keywords: visual working memory, change detection, cognitive aging, feature binding

Supplemental materials: http://dx.doi.org/10.1037/pag0000152.supp

In the literature on source memory deficits in long-term memory an influential suggestion has been that older adults specifically struggle to bind distinct elements of an episode together (Naveh-Benjamin, 2017). Indeed studies using a variety of stimuli, from pairs of unrelated words to pictures of faces and scenes, have convincingly shown that age differences in recognition performance are disproportionately large for pairs of items relative to the items individually (see Old & Naveh-Benjamin, 2008, for a review and meta-analysis). The success of this account in the long-term memory realm has led to the suggestion that binding deficits may account for age-related change to working memory. In particular, researchers (Brockmole, Parra, Della Sala, & Logie, 2008; Cowan, Naveh-Benjamin, Kilb, & Saults, 2006; Sander, Lindenberger, & Werkle-Bergner, 2012) have become interested in whether a deficit in binding together the basic features of objects (e.g., color, shape, location) can account for some of the pronounced decline of visual working memory (VWM) observed across the life span (e.g., Brockmole & Logie, 2013; Johnson, Logie, & Brockmole, 2010). However, in contrast to the pronounced between-item associative deficit observed in long-term memory—and more recently in working memory (Chen & Naveh-Benjamin, 2012; Peterson, Schmidt, & Naveh-Benjamin, 2017)—the evidence concerning this within-object “level of binding” (Zimmer, Mecklinger, & Lindenberger, 2006) has been less clear (however, see Peterson & Naveh-Benjamin, 2016).

Evaluating the Evidence for Age-Related VWM Binding Deficits

Mitchell and colleagues were the first to assess the effect of age on the retention of feature bindings in working memory in a series...
of experiments (Mitchell, Johnson, Raye, Mather, & D’Esposito, 2000; Mitchell, Johnson, Raye, & D’Esposito, 2000). In these experiments they presented younger and older participants with a sequence of three nameable clip-art-like objects (for 1 second each) on a 3 × 3 grid and then, following an 8.5-second delay, probed memory for the objects, locations, or object–location pairings in separate blocks. They found significant age differences in performance when the task required maintenance of object–location conjunctions, whereas there was no significant age difference when assessing memory for the individual components. These findings have been taken to indicate that older adults have a specific difficulty in maintaining “what was where” in VWM.

However, upon further consideration, these findings are not as convincing as they first appear. Namely, reporting a significant age difference in one condition and not in another is not evidence for a disproportionate effect of age in that condition; this requires a test of the age group-by-condition interaction (see Gelman & Stern, 2006; Nieuwenhuis, Forstmann, & Wagenmakers, 2011). Crucially, in the experimental experiments of Mitchell, Johnson, Raye, Mather et al. (2000) this interaction was not significant by conventional standards (see p. 199), so it is unlikely that the interaction was present in this case. Thus it appears that in these studies there was actually no clear evidence that age disproportionately affects binding to location in VWM. Evidence for crucial interactions is also missing from more recent studies that have been cited in support of VWM binding deficits (Borg, Leroy, Favre, Laurent, & Thomas-Antérion, 2011; Fandakova, Sander, Werkle-Bergner, & Shing, 2014).

Other recent studies have been more equivocal on this issue. In two experiments, Read, Rogers, and Wilson (2016) found that the effect of age on change detection was not clearly larger for discriminating feature (color or shape) to location binding changes relative to changes to these features individually. On the other hand, also in two experiments, Peterson and Naveh-Benjamin (2016) presented colored shapes in different locations and found a slightly larger age effect on change-detection performance for the binding of object and location relative to object or location alone.

Using a recall task, Peich, Husain, and Bays (2013) found that older adults were more likely than younger adults to relocate a previously seen feature (color or orientation) to a location that was occupied by a different feature; that is, older adults were more likely to commit so-called “mis-binding” errors. However, following up on this, Pertsov, Heider, Liang, and Husain (2015) have suggested that once age differences in the recall of the features themselves are accounted for, older adults do not commit any more mis-binding errors than younger adults. In summary, the literature following the early findings of Mitchell and colleagues has been somewhat mixed but in general suggests little to no additional age-related binding deficit when location is concerned.

Studies assessing the binding between surface features (e.g., shape and color), as opposed to the binding of features to location, have more clearly suggested that there is no disproportionate VWM deficit for conjunctions of features over and above the general drop in performance observed for the component features (provided that the influence of potential verbal strategies is minimized, either via the use of difficult to name stimuli or concurrent suppression, Brockmole et al., 2008; Brown, Niven, Logie, Rhodes, & Allen, 2017; Isella, Molteni, Mapelli, & Ferrarrese, 2015; Parra, Abrahams, Logie, & Della Sala, 2009; Peterson & Naveh-Benjamin, 2016; Rhodes, Parra, & Logie, 2016). While some have reported age-by-condition interactions these either go against the expectation of an age-related binding deficit (Brown et al., 2017, Experiment 3) or have failed to replicate (Brown & Brockmole, 2010; Brown et al., 2017; Rhodes et al., 2016). Thus the weight of evidence, so far, appears to support the absence of an age-related binding deficit for surface features.

In light of the pronounced associative deficit in long-term memory (Naveh-Benjamin, 2000; Old & Naveh-Benjamin, 2008) and working memory (Chen & Naveh-Benjamin, 2012; Peterson et al., 2017), the lack of a binding deficit for within-object features in VWM is perhaps surprising. However there are reasons to predict differential effects of age on these theoretically different levels of binding (see Zimmer et al., 2006, for discussion). First between-item, or relational, binding appears to be heavily dependent on the hippocampus in both working memory (Parra et al., 2015; Piekema, Rijpkema, Fernández, & Kessels, 2010) and long-term memory (Staresina & Davachi, 2010), and the effects of healthy aging in terms of volumetric changes and white matter hyperintensities are especially pronounced in this region (Raz & Rodrigue, 2006; Yang, Goh, Chen, & Qiu, 2013). By contrast retaining the binding of color and shape does not appear to rely on the hippocampal formation (Parra, Della Sala, Logie, & Morcom, 2014; Parra et al., 2015). Second the binding of feature conjunctions in VWM appears to proceed relatively automatically (e.g., Allen, Baddeley, & Hitch, 2006), whereas the association of contextual, or extrinsic, attributes does not appear to proceed as obligatorily in working memory tasks (Ecker, Maybery, & Zimmer, 2013). Thus, as the effects of age tend to be reduced for relatively automatic processes (see Craik & Bialystok, 2006, for a review), we may predict the relational and conjunctive binding mechanisms to dissociate with age.

The apparent absence of a specific age-related deficit for conjunctions of color and shape in change detection tasks becomes even more intriguing when paired with the large deficit observed in the early stages of Alzheimer’s disease (e.g., Parra, Abrahams, Logie, & Della Sala, 2010). Consequently a great deal of effort has been devoted to identifying potential boundary conditions under which healthy older adults may struggle to retain conjunctions of surface features. Age differences in change detection performance have been assessed under conditions of increased attentional load (Brown & Brockmole, 2010), extended encoding duration (Brown et al., 2017; Rhodes et al., 2016), or with potentially interfering information in the retention interval (Brown et al., 2017) and none of these manipulations have been found to specifically impair older adults’ ability to detect conjunction changes. However, the findings of one study in the literature (Cowan et al., 2006) point to another potential boundary condition, under which an age-related binding deficit may be observed, that has yet to be addressed in this context.
Motivation for the Present Experiments

Some of the strongest evidence that older adults may struggle to detect changes to conjunctions over features alone so far appears to come from the findings of Cowan et al. (2006). Using a change detection paradigm, Cowan et al. (2006) found that, when color–location conjunction change trials were mixed with changes to color only, there was a disproportionate effect of age on the detection of conjunction changes. Older adults showed an increased tendency to miss the conjunction changes. When these different trial types were presented in separate blocks, however, there was no clear age-related binding deficit. Thus mixing feature and conjunction changes may reveal specific age-related problems in maintaining and utilizing bound representations in VWM. Cowan et al. (2006) suggested that older adults were more likely to perform the change detection discrimination on the basis of probe familiarity which, in the mixed condition, would be sufficient to detect salient feature changes but not the less salient swaps of color and location. Relying only on familiarity in the blocked condition, on the other hand, would not support the detection of any changes, thus the older adults in this condition may have adopted a different criterion, or encoding/retrieval strategies that improved their performance in the binding condition.

Nevertheless, there are some potential issues with the paradigm used by Cowan et al. (2006) that warrant detailed discussion. In their study, memory was probed by a single circled item in the test array with unprobed items also present, unchanged. Additionally, as the largest array contained 10 items, colors were selected from the set of seven with replacement, leading to the presence of duplicates in the arrays. Due to this a color–location binding change in these experiments always introduced a duplicate into the test array, whereas a color change always introduced a unique color. When conditions were blocked this could have given the answer away to participants; therefore, Cowan et al. (2006) ensured that on no-change trials in the binding condition the probe was always a duplicate and in the color-only condition the unchanged probe was unique (see Cowan, Saults, & Blume, 2014, p. 1826 for more detail). Consequently in the binding condition it was sufficient to focus on duplicates and in the color condition to focus on unique colors. Age differences in apprehending this intricate aspect of the task might have contributed to age differences in performance. Further, it seems likely that noting which colors are duplicated in an array is a simpler task than noting which colors are unique. This would introduce a benefit for binding trials specifically when trials were blocked as opposed to mixed. This is evident in the Cowan et al. (2006) data (see Table 2, p. 1095), as there was a clear difference in terms of sensitivity ($d'$) between the color and binding conditions in Experiment 1A in which these trials were mixed. In Experiment 2A, where trials were blocked, there was very little difference between the two conditions and, in fact, for younger adults sensitivity to binding changes was slightly better than that for color changes. So it may be that participants in the studies reported by Cowan et al. (2006) were able to use additional cues to guide the change detection decision that may have differentially benefited features and conjunctions, and have had a differential effect on older compared with younger participants.

In addition, the findings of Chen and Naveh-Benjamin (2012) cast doubt on a role of mixing versus blocking trials in the emergence, or exacerbation, of an age-related binding deficit. They used a continuous recognition paradigm in which participants studied a stream of face–scene pairs with interspersed tests of memory for items or associations following varying delays. The commonly observed associative deficit was no larger when item and associative trials were mixed together relative to when they were encountered in separate trial blocks. Given that, to our knowledge, only these two studies have directly addressed this issue, the present experiments aimed to follow up on the findings of Cowan et al. (2006) by directly comparing, in younger and older healthy adults, mixed and blocked trial lists for stimuli defined by conjunctions of color and shape (Experiment 1), and color and location (Experiment 2).

Experiment 1

In Experiment 1 we assessed the effect of mixing switches of color–shape conjunction with trials in which a brand new feature (color or shape) is introduced at test on older adults’ change detection performance. There were a number of procedural changes between this study and the experiments of Cowan et al. (2006). Most notably during the test phase of our change detection task we only presented a single item with the unprobed items absent. We also only assessed two set sizes (3 and 6), meaning that we could sample features for each array without replacement. The absence of duplicates in the studied arrays and the absence of unprobed items in the probe displays avoids the potential inadvertent cues that were present in the paradigm of Cowan et al. (2006), which could have created an effect in that study due to strategic differences rather than binding differences. Of course another major difference between this experiment and that of Cowan et al. is that they assessed color-location binding, whereas we looked at the binding of color and shape, a form that has appeared to be largely age-invariant in studies using a blocked design, given our interest in potential boundary conditions on this. This is addressed in Experiment 2.

Method

Participants. Forty eight younger adults from the student population of the University of Edinburgh took part in Experiment 1 in return for £5 for the 45-min session. Forty nine healthy older adults from the University of Edinburgh Psychology volunteer panel of members of the general public also took part and were offered £5 in return for participation. These groups were split between two conditions; one in which color, shape, and color–shape conjunction changes were mixed together and another in which they were presented in separate trial blocks. Table 1 provides participants’ demographic information. All older adults scored 27 or above on the Mini Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975).

Stimuli and apparatus. Stimulus features were drawn from sets of eight difficult-to-name colors and eight abstract polygons taken from Brockmole et al. (2008). Items in the memory array were constructed by selecting colors and shapes from these sets without replacement. Stimuli were presented on a gray background in eight locations surrounding the center of the screen following an invisible circle (radius = 2.6°, at an approximate viewing distance of 50 cm). Objects measured approximately 1° visual angle and
were separated center-to-center by at least 2°. The experiment was programmed using PsychoPy (Peirce, 2007) and presented over a 18” E96f + SB ViewSonic monitor with a resolution of 1,024 × 768 and refresh rate of 100 Hz.

**Design and procedure.** Prior to the main change-detection task both groups completed the National Adult Reading Test (Nelson, 1982) to obtain an estimate of verbal IQ (see Table 1) and a test of color vision (Dvorine, 1963). The older group also completed the MMSE prior to completing the main part of the experiment.

The general trial sequence of the change detection task is presented in Figure 1A along with examples of the kinds of trials presented to participants (Figure 1B). Participants initiated each trial by pressing the space-bar and following a 1,000-ms fixation cross the memory array appeared for 900 ms. This was followed by a 1,000-ms blank retention interval and then a single central probe item which remained present until a response was made. Finally, in line with the procedure of Cowan et al. (2006), participants were presented with feedback for 1,000 ms in the form of a fixation cross that was colored green for a correct response or red otherwise.

Half of the trials presented to participants involved no change as the probe was selected at random from the three or six objects presented. The remaining half of trials were split between color change, shape change, and binding change types (either blocked or mixed). A color change involved filling a previously seen shape with a color from outside the original memory set and a shape change involved presenting a new shape in a previously seen color. A binding change involved presenting combination of a color and shape from separate memory objects as the probe item (see Figure 1B). As described above, participants in the blocked condition saw these changes in separate blocks with their own change trials, whereas participants in the mixed condition saw the three kinds of change trial interspersed with no-change trials.

### Table 1

**Participant Characteristics for Experiment 1**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Blocked Younger</th>
<th>Older</th>
<th>Mixed Younger</th>
<th>Older</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>25</td>
</tr>
<tr>
<td>N Female</td>
<td>17</td>
<td>17</td>
<td>24</td>
<td>18</td>
</tr>
<tr>
<td>Mean age (SD)</td>
<td>20.708 (2.896)</td>
<td>70.958 (5.614)</td>
<td>21.125 (2.659)</td>
<td>70.280 (4.430)</td>
</tr>
<tr>
<td>Years of education</td>
<td>16.104 (2.613)</td>
<td>17.062 (2.898)</td>
<td>16.521 (2.164)</td>
<td>15.960 (2.428)</td>
</tr>
<tr>
<td>NART Verbal IQ</td>
<td>111.915 (5.564)</td>
<td>121.225 (4.264)</td>
<td>111.298 (7.142)</td>
<td>118.690 (5.756)</td>
</tr>
<tr>
<td>MMSE</td>
<td>—</td>
<td>29.333 (.761)</td>
<td>—</td>
<td>29.600 (.816)</td>
</tr>
</tbody>
</table>

**Note.** NART = National Adult Reading Test; MMSE = Mini Mental State Examination.
The main experiment was split into three blocks with 32 change and 32 no-change trials each distributed evenly across the different set sizes. For the blocked condition all change trials in a given block were of a single type and for the mixed condition a change was equally likely to occur for color, shape, and binding. Participants in the blocked condition were given six practice trials looking for a particular kind of change before the corresponding block, whereas participants in the mixed condition were given 18 practice trials before the first block with all three kinds present. In the blocked condition the order of the three memory conditions (color, shape, binding) was fully counterbalanced. All experiments reported here were approved by the University of Edinburgh research ethics committee.

**Analysis.** In their main analysis, Cowan et al. (2006) aimed to separate out the contribution of sensitivity and response bias to change detection accuracy, and we aim to do the same. While they reported analyses of the (normal equal variance) signal detection theory measures, d' and c, in the present analysis we opt for P*, or “corrected recognition” and its associated bias measure, B, (Snodgrass & Corwin, 1988). Selective influence studies have suggested that change detection with highly distinguishable, or suprathreshold, stimuli is mediated by discrete, all-or-nothing states rather than a graded continuous decision variable (Donkin, Tran, & Nosofsky, 2014; Rouder et al., 2008; although this may not characterize performance for more subtle changes, Keshvari, van den Berg, & Ma, 2013). Consequently, the high-threshold measure used here should better allow us to separate discriminability from bias. Nevertheless, in addition we also ran analysis with d' as the outcome measure; the P*, and d' analyses were in agreement.

Many previous reports of no additional binding cost with healthy aging (e.g., Brockmole et al., 2008; Brown & Brockmole, 2010; Isella et al., 2015) have been hampered by the fact that failure to reject the null, in and of itself, is not sufficient to argue against the presence of an Age × Condition interaction. Therefore in the present analysis of discrimination and bias metrics we opt for a model comparison approach (see also, Brown et al., 2017; Rhodes et al., 2016). To obtain Bayes factors for the comparison of models defined by the presence or absence of specific main- or interaction-effects we use the default family of priors outlined by Rouder, Morey, Speckman, and Province (2012).

These data were analyzed using the BayesFactor package in R (Morey & Rouder, 2015; R Core Team, 2015). The analysis proceeded by comparing a full model, with all possible main effects and interactions, to reduced models omitting a single component at a time.¹ The resulting Bayes factors (B) reflect the weight of evidence in favor of omitting the component in question from the full model. Consequently a B < 1 implies evidence in favor of the effect, whereas a B > 1 denotes evidence against inclusion of the effect in the full model. Bayes factors have a simple interpretation (i.e., the reduced model is B times as likely as the full model given the data and priors) without necessary recourse to thresholds; nevertheless a B > 3 or < 1/3 would usually be considered “substantial”.

**Results**

Table 2 presents accuracy by age group and block type for each memory condition and set size split by whether or not a change had actually occurred. Analyzing raw accuracy (i.e., correct/incorrect) responses with standard techniques (such as ANOVA) can result in evidence for spurious interaction effects (Dixon, 2008). Therefore we applied a hierarchical generalized linear mixed effects model using a logit link function (see Jaeger, 2008). The details of this modeling are necessarily involved and, given that our main interest is in estimates of discriminability and bias, we present the full results of analyses of accuracy in the online supplementary material. In summary, however, there was no indication that age differences in change detection accuracy were larger for conjunction changes, nor was there any indication that mixing versus blocking different type of trial changed this.

Corrected recognition for Experiment 1 is presented in Figure 2A. Unsurprisingly the default Bayes factor analysis revealed overwhelming support for main effects of memory condition, set size, and age group (all Bs < 1/100). By contrast the weight of evidence was against an overall effect of block type (B = 5.19, see Figure 2). Turning to interactions, specifically those including age-group, there was strong evidence for an interaction between age and set size (B = 0.0294) which was qualified by substantial evidence for a three-way interaction with memory condition (B = 0.308). As can be seen in Figure 2 the drop in P* associated with increasing the number of to-be-remembered items from three to six was less pronounced for the older group especially in the shape and binding conditions. Indeed when the analysis is conducted

¹This was achieved using the anovaBF function from the BayesFactor package. The default settings were used with the exception that “which-Models” was set to “top”, which compares a full model to reduced models omitting a specific component, and the number of MCMC iterations was set to 50,000. An additional 10,000 samples were taken until the proportional error for all comparisons was below 5%.

<table>
<thead>
<tr>
<th>Trial Type</th>
<th>Block Type</th>
<th>Younger</th>
<th>Older</th>
<th>Younger</th>
<th>Older</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change Color</td>
<td>Blocked</td>
<td>3</td>
<td>.924 (.026)</td>
<td>.828 (.027)</td>
<td>.966 (.008)</td>
</tr>
<tr>
<td></td>
<td>Mixed</td>
<td>6</td>
<td>.805 (.024)</td>
<td>.693 (.026)</td>
<td>.917 (.018)</td>
</tr>
<tr>
<td>Shape</td>
<td>Blocked</td>
<td>3</td>
<td>.740 (.032)</td>
<td>.586 (.039)</td>
<td>.737 (.023)</td>
</tr>
<tr>
<td></td>
<td>Mixed</td>
<td>6</td>
<td>.635 (.034)</td>
<td>.547 (.035)</td>
<td>.714 (.030)</td>
</tr>
<tr>
<td>Binding</td>
<td>Blocked</td>
<td>3</td>
<td>.750 (.026)</td>
<td>.596 (.035)</td>
<td>.742 (.033)</td>
</tr>
<tr>
<td></td>
<td>Mixed</td>
<td>6</td>
<td>.742 (.031)</td>
<td>.659 (.044)</td>
<td>.734 (.028)</td>
</tr>
</tbody>
</table>

**Note.** Standard errors are given in parenthesis. The mixed condition results in one accuracy estimate per set size for no change trials.
omitting the color condition the weight of evidence is against the three-way interaction ($B = 4.05$).

Crucially there was no suggestion in this data set that older adults were disproportionately less sensitive to binding changes; the reduced model omitting the age group-by-condition interaction was favored over the full model by approximately 15 to 1 ($B = 15.51$). Further there was substantial evidence against modulation of this by block type ($B = 4.8$). For the remaining interactions including age group, Bayes factors favored their omission from the full model (all $Bs > 3$).

Cowan et al. (2006) also found modulation of response bias, with older adults adopting a more conservative criterion for binding changes, leading to a greater miss rate. However, the same trends did not appear in the present data set (see Figure 2B). The data did not clearly favor the presence of an overall effect of age bias over the null ($B = 1.68$) but there was good evidence against interactions between age-group and condition ($B = 11.56$), and the three-way interaction with block type ($B = 5.24$). Tables of the full output of these, and subsequent, default Bayes factor analyses can be found in the online supplementary material.

**Discussion**

In the literature on feature binding in VWM there has been little evidence to suggest that healthy older adults disproportionately struggle to detect changes to color–shape conjunctions relative to changes to these individual features (e.g., Brown et al., 2017; Rhodes et al., 2016). However, in these studies changes to individual features are presented in separate trial blocks to changes to feature binding. The findings of Cowan et al. (2006) suggest that mixing these trial types may in fact reveal a specific difficulty for older adults in discriminating conjunction changes, perhaps linked to a greater reliance on overall probe familiarity at test. Experiment I found no indication that mixing trial types specifically affected older adults’ accuracy for detecting binding changes; in fact, it provided fairly strong evidence against this interaction.

There are a number of differences between our experiment and those of Cowan et al. (2006). For one, we assessed the binding between surface features, whereas they assessed the binding of color to location. Given previous suggestions that binding of what to where in VWM is a specific problem for older adults (Mitchell, Johnson, Raye, Mather et al., 2000; Mitchell, Johnson, Raye, & D’Esposito, 2000; Borg et al., 2011), whereas surface feature binding is largely unaffected (Brockmole et al., 2008; Brown et al., 2017), we may expect the two to be differentially affected by experimental manipulation. However, as reviewed above, much of the early evidence for location binding deficits is questionable and subsequent research has been less clear (Read et al., 2016; Peterson & Naveh-Benjamin, 2016).

What may prove more important are the procedural differences between our experiment and those of Cowan and colleagues. In fact there are good reasons to think that the paradigm used here is an improvement. First, like Cowan et al. (2006) we had participants make a judgment on a single item; however, unlike their study we did not present unprobed items. Further we selected the colors for each memory array from a set of eight without replacement. The presence of duplicated colors and unprobed items in the test array in the study of Cowan et al. (2006) may have acted as an additional cue as to whether a change had occurred and what kind of change could have occurred (see Cowan et al., 2014). For color change trials in their study the probe was always unique, whereas for binding change trials the probe was always a duplicate. The present study, by avoiding duplicates and presenting a single item in the probe array, circumvents these potential confounds.
Second, in the present experiment, in addition to color–shape change trials, we included trials on which color or shape changed individually. This was the case in the experiments of Chen and Naveh-Benjamin (2012), who also found no effect of mixing item and binding changes. On the other hand, location was not constrained in the experiments of Cowan et al. and was never probed as an isolated feature, only in the combination of color and location. Including trials probing memory for both of the individual component features of a conjunction may be important as it ensures that participants are motivated to pay attention to each component feature equally (as a change is just as likely to occur for one feature as it is for the other). There is some evidence that attentional biases such as these to one component feature may be particularly harmful to older adults (Benjamin, Diaz, Matzen, & Johnson, 2012).

Therefore, in a second experiment we assessed color–location binding using the same single probe paradigm deployed in Experiment 1, but with change trials for both component features. If mixing trial types specifically affects older adults’ ability to detect color–location conjunction changes we should be able to demonstrate this, despite procedural modifications.

Experiment 2

This second experiment assessed color–location binding, as done in Cowan et al. (2006), in a situation where cues at test were minimized and incentive to pay attention to each component feature was equal, therefore resulting in a stronger test of an age-related binding deficit.

Method

Participants. Forty eight younger adults were recruited from the student population of the University of Edinburgh and 49 healthy older adults were recruited from the Psychology research volunteer panel. None of these individuals had participated in Experiment 1. Participants were offered £5 in return for participation for the 45-min session. Table 3 presents the demographic information of the participants split between the mixed and blocked conditions. Once again age groups were well matched for years of education and the healthy older adults received higher estimates of verbal IQ from the National Adult Reading Test (NART). All older adults scored 27 or above on the MMSE.

Stimuli and apparatus. In Experiment 2 stimuli were defined by both color and location. Both colors and locations were selected from the sets described in Experiment 1. The locations were constrained to surround the center of the screen following an invisible circle (radius = 2.6°, see Figure 1C). As shape was irrelevant, colors were presented in circles with a radius of approximately 0.5° of visual angle and separated center-to-center by at least 2°. Figure 1C gives examples of memory and test arrays in this experiment. In the probe display, the probe object could appear in a color not found in the array (color-change trials), a color found in the array could appear in a location not found in the array (location-change trials), a color previously found in the array could move to a location previously used for a different color (binding change), or an object could appear exactly as before (no change).

Design and procedure. The general design and trial procedure was identical to Experiment 1. Participants either saw different kinds of change trial (see Figure 1C) mixed together in the same block of trials or separately in their own trial block.

Results

Accuracy across the experimental conditions of Experiment 2 is presented in Table 4. Once again the hierarchical logit model gave no strong indication of a disproportionate effect of age on the detection of conjunction changes. Further, there was no evidence that mixing different trials led older adults to miss conjunction changes, as found by Cowan et al. (2006; see online supplementary material for full results).

Turning to our main outcome of interest, corrected recognition (Pr) for Experiment 2 is presented in Figure 3A. The Bayesian ANOVA, once again, revealed overwhelming evidence for effects of memory condition, set size and age (all Bs < 1/100). Also in line with Experiment 1 the weight of evidence was against an overall effect of block type (B = 2.22). For interactions including group, there was substantial evidence of an interaction between age and set size (B = 0.2981). In this experiment the overall age difference was larger at set size 6 relative to set size 3 and there was no suggestion that this varied across memory conditions (B = 3.051). Most importantly omitting the age group-by-memory condition interaction resulted in a model that was favored by more than 16 to 1 over the full model (B = 16.32) and there was strong evidence against the three-way interaction with probe type (B = 11.1).

The BANOVA marginally favored the exclusion of the three-way interaction between age, set size and memory condition (B = 3.05) and when contrasting color and binding conditions only the data favored neither the inclusion or exclusion of this component (B = 1.28). Assessing Figure 3 it appears that older adults were

Table 3

Participant Characteristics for Experiment 2

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Blocked</th>
<th></th>
<th>Mixed</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Younger</td>
<td>Older</td>
<td>Younger</td>
<td>Older</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>24</td>
<td>25</td>
<td>24</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>N Female</td>
<td>18</td>
<td>16</td>
<td>17</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>Mean age (SD)</td>
<td>20.708 (2.528)</td>
<td>70.000 (4.770)</td>
<td>21.125 (1.727)</td>
<td>71.417 (4.671)</td>
<td></td>
</tr>
<tr>
<td>Years of education</td>
<td>16.021 (1.992)</td>
<td>16.420 (2.929)</td>
<td>16.625 (1.583)</td>
<td>16.583 (3.847)</td>
<td></td>
</tr>
<tr>
<td>NART Verbal IQ</td>
<td>108.448 (4.059)</td>
<td>119.420 (3.590)</td>
<td>108.875 (5.707)</td>
<td>119.325 (5.293)</td>
<td></td>
</tr>
<tr>
<td>MMSE</td>
<td>—</td>
<td>29.320 (.852)</td>
<td>—</td>
<td>29.500 (.933)</td>
<td></td>
</tr>
</tbody>
</table>

Note. NART = National Adult Reading Test; MMSE = Mini Mental State Examination.
slightly less sensitive to binding changes at set size 3. Of course the evidence is indecisive in this respect but this general pattern is certainly not expected under an age-related binding deficit which would surely be compounded by increasing encoding and retention demands.

For the four-way interaction between all variables, the data were not particularly convincing either way ($B = 2.44$). As can be seen in Figure 3 this ambiguity appears to be driven by the performance of younger adults in the blocked location condition. There is a simple explanation as to why our younger group displayed a smaller set size effect in the location condition. As stimulus locations were selected from a restricted set, and participants in the blocked condition were aware of what kind of change was possible on a given trial, an appropriate strategy in the location-only condition was to note the empty locations in the memory array; if, when the probe appeared, the locations were still empty then there had been no change, whereas if one were occupied then a location change must have occurred. Using this strategy would make larger arrays easier as there are fewer empty locations to monitor. It is interesting to note that, overall, younger adults were more likely to note this aspect of the task as evidenced by their pattern of performance and post-experiment discussion with the researcher. Assessing such strategy difference between age groups on working memory tasks will prove important in future work to separate out basic effects of healthy aging on binding mechanisms from difference in strategy use (see Peterson & Naveh-Benjamin, 2016, for this approach). In support of this proposal when location is omitted from the analysis, and only the color and binding conditions are contrasted, the evidence against the four-way interaction is slightly more convincing ($B = 3.03$).

Once again we assessed response bias ($B_r$) as Cowan et al. found more conservative responding for conjunction trials in older adults. This data is depicted in Figure 3B. There was evidence of an effect of age on bias ($B = 0.22$), with older adults exhibiting a more neutral response bias than younger adults (see Figure 3B). Further there was strong evidence against the age by condition ($B = 18.33$).

Table 4

<table>
<thead>
<tr>
<th>Trial Type</th>
<th>Blocked</th>
<th>Mixed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Younger</td>
<td>Older</td>
</tr>
<tr>
<td>Change Color</td>
<td>3</td>
<td>.948 (.013)</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>.820 (.022)</td>
</tr>
<tr>
<td>Location</td>
<td>3</td>
<td>.924 (.031)</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>.917 (.027)</td>
</tr>
<tr>
<td>Binding</td>
<td>3</td>
<td>.919 (.017)</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>.826 (.018)</td>
</tr>
<tr>
<td>No change Color</td>
<td>3</td>
<td>.930 (.014)</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>.740 (.027)</td>
</tr>
<tr>
<td>Location</td>
<td>3</td>
<td>.932 (.013)</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>.810 (.024)</td>
</tr>
<tr>
<td>Binding</td>
<td>3</td>
<td>.927 (.018)</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>.654 (.027)</td>
</tr>
</tbody>
</table>

Note. Standard errors are given in parenthesis. The mixed condition results in one accuracy estimate per set size for no change trials.

This document is copyrighted by the American Psychological Association or one of its allied publishers. This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.
and age by condition by block type ($B = 12.87$) interactions (see online supplementary material).

**Discussion**

Experiment 1 examined whether mixing color–shape conjunction change trials with changes to the individual components specifically affected older adults’ ability to discriminate between old and new feature bindings. The results went against this suggestion. As the experiment of Cowan et al. (2006) assessed the binding of color and location, Experiment 2 used stimuli defined by these features instead. Again the results suggest that older adults have no specific problem in determining a change of color–location binding and that mixing these trials with changes to either color or location individually does not modulate this.

Thus rather than the type of binding being the crucial factor determining the differing findings of the present experiments and those of Cowan et al., we suspect that another factor was at play in their finding of a specific deficit. As previously outlined, one potentially crucial factor may have been the omission of trials on which only location changed (which were included here) that may have incentivized participants to prioritize color. To assess this we ran an additional study with 24 younger and 24 older adults that was identical to Experiment 2 but with location change trials omitted. The results, described fully in the online supplementary material, replicated the key findings of Experiment 2. Indeed the omission of the age-by-trial type (color change, binding change, no change) interaction for $P_r$ was favored in this data set by a factor of over 4 to 1. Thus it appears that the omission of location change trials, by itself, is not enough to reproduce the findings of Cowan et al. (2006).

We tentatively suggest that the findings of Cowan et al. (2006) can largely be put down to the presence of duplicates in the memory array and unprobed items in the probe array. In the Introduction we noted that in that study, the probed item in the color condition always was a unique color (whether or not it was changed from the studied array), whereas in the binding condition it was always a duplicate, whether or not there was a change (see, Cowan et al., 2014). In the blocked condition participants knew exactly which kind of change to expect and thus may have appreciated that in the binding condition; for example, it was sufficient to note only the repeated colors. Finally, an in-press correction of Cowan et al. (2006) reports that, due to an error, the binding deficit in their mixed-block experiment was smaller than was reported in the article.

It is clear that the use of a single probe without the presence of unprobed items is a better way of addressing the question of the efficacy of feature binding in healthy aging (see also, Read et al., 2016, for similar findings). If there were a true effect of mixing different trial types on the ability of older adults to discriminate binding changes the present paradigm should have shown this. On the contrary, three experiments (including an auxiliary experiment discussed above) have demonstrated no effect of mixing versus blocking trials.

**General Discussion**

Whether or not healthy aging affects the ability to temporarily maintain feature conjunctions in VWM is of both theoretical and practical importance. The magnitude of age differences increase greatly when participants are required to actively maintain visual information relative to verbal (see, e.g., Johnson et al., 2010) and it has been suggested that reduced feature integration across the adult life span may account for some of this (Brockmole et al., 2008; Cowan et al., 2006; Sander et al., 2012). Further, there is growing evidence of a large feature binding deficit for surface features (e.g., shape and color) in early Alzheimer’s disease, which may help distinguish this from healthy aging (e.g., Parra et al., 2010). Many studies, so far, suggest that there is no differential effect of healthy aging on the temporary retention of color–shape conjunctions (Brockmole et al., 2008; Brown et al., 2017; Isella et al., 2015; Rhodes et al., 2016). Also, while it has previously been proposed that binding to location may be a specific difficulty for older adults, a critical examination of the literature reveals that the initial evidence for this suggestion is underwhelming (Borg et al., 2011; Mitchell, Johnson, Raye, Mather et al., 2000; Mitchell, Johnson, Raye, & D’Esposito, 2000), and recent studies have been more equivocal on this issue (Pertzov et al., 2015; Read et al., 2016; Peterson & Naveh-Benjamin, 2016).

Cowan et al. (2006) reported that healthy older adults were much less able to detect changes to color–location binding when these trials were mixed with trials containing a change to color only relative to when these trial types were blocked. If genuine, this effect may reveal differences between younger and older adults in the use of familiarity-based recognition in the change-detection task or strategic differences when the type of possible change to prepare for is known. Experiment 1 assessed the effect of mixing versus blocking feature and conjunction changes for stimuli defined by color and shape. Overall accuracy gave no indication that older adults specifically struggled in detecting binding changes and block type did not modulate this. Further, analysis of corrected recognition as a measure of discriminability (Snodgrass & Corwin, 1988) provided strong evidence against a role of mixing trial types. Given the number of methodological changes between Experiment 1 and the experiments of Cowan et al. (2006), in Experiment 2 we went on to assess whether we would find a role for mixing versus blocking with stimuli defined by color and location. Again, we found evidence against an overall age-related binding deficit and against any modulation by block type. As argued above, it seems likely that procedural aspects, namely the inclusion of unprobed items and the use of duplicated colors, led to the appearance of an age-related binding deficit that does not generalize. A follow-up experiment suggested that this discrepancy cannot be accounted for by the fact that in our experiments we included change trials for each of the component features alone and also served to replicate our main findings. Unlike previous investigations in this area, our measure of performance (corrected recognition) was chosen to provide a more accurate characterization of the discrimination underlying change detection with large changes between study and test (see Donkin et al., 2014; Rouder et al., 2008). This choice of measure, justified on the basis of previous research, and the converging analysis of raw accuracy and $d’$, increases our confidence in the absence of the crucial interaction in these experiments.

Our findings regarding color–shape binding are in line with a growing number of studies that either fail to find a specific age-related binding deficit in VWM or provide evidence against one (Brockmole et al., 2008; Brown et al., 2017; Isella et al., 2015;
Peterson & Naveh-Benjamin, 2016; Rhodes et al., 2016) and argue against the mixing of conjunction with feature changes as a boundary condition on this. In relation to the literature on binding to location in VWM, the results of Experiment 2 add to a mixed pattern of findings. As outlined in the Introduction, initial studies (Mitchell, Johnson, Raye, Mather et al., 2000; Mitchell, Johnson, Raye, & D’Esposito, 2000) claimed to show that older adults struggle to detect changes to object locations in a 3 × 3 grid (see also Borg et al., 2011; Fandakova et al., 2014); however, the evidence to support this claim was either insufficient or not provided. What has followed has been a more varied pattern of findings with some failing to demonstrate a disproportionate effect of healthy aging on binding what was where in VWM (Pertzov et al., 2015; Read et al., 2016) and others finding such an effect (Peterson & Naveh-Benjamin, 2016). Given the considerable methodological overlap between these studies and the present work the precise origin of this discrepancy is unclear; however, the extant literature is inconsistent with a large deficit where location is concerned (Peterson & Naveh-Benjamin, 2016).

Future work would benefit from a more nuanced treatment of location as a feature. For example, there is evidence to separate the categorical representation of location (i.e., that the red circle was located toward the bottom right) from more fine-grained coordinate-level representations (Postma, Kessels, & van Asselen, 2008). In addition there is the well known distinction between allocentric, that is viewpoint-independent, and egocentric, viewpoint-dependent, spatial representation (O’Keeffe & Nadel, 1978). It has been suggested that remembering where items are in allocentric space may be a more relational form of binding, where identity is bound to an external frame of reference (see Baddeley, Jarrod, & Varga-Khadem, 2011, for discussion), in which case we may expect to observe an age-related deficit given older adults’ well established problems with relational binding in both long-term memory (Naveh-Benjamin, 2000; Old & Naveh-Benjamin, 2008) and working memory (Chen & Naveh-Benjamin, 2012; Peterson et al., 2017). The simultaneous presentation of a memory array used here may foster a more egocentric representation, placing less demand on relational mechanisms. Further presenting all items at once may result in higher order structural representations of the overall properties or configuration of an array (Brady, Konkle, & Alvarez, 2011), which may be relatively preserved in normal aging (Olson et al., 2004). Of course we did not set out to manipulate the reliance on different frames of reference but this seems a fruitful area for further investigation.

Another area in need of further investigation is the distinction between relational and conjunctive forms of binding. In the literature as a whole there is generally a confound between the type of binding required and the type of material used. Studies assessing relational, or between-item, binding tend to use complex, meaningful material such as images of faces or scenes (e.g., Chen & Naveh-Benjamin, 2012), whereas studies assessing conjunctive, or within-item, binding use basic features like color and shape (e.g., Brown et al., 2017). Recently, van Geldorp, Parra, and Kessels (2015) addressed this directly, using stimuli defined by color and shape, by either presenting these features as spatially separate items to be related or as unitized objects with the color filling the shape. In general older adults were poorer than younger and middle-aged adults at recalling the pairings of features and crucially this did not appear to depend on the mode of presentation. This initial evidence suggests that maybe stimulus complexity is an important factor in some of the discrepant findings in the literature. However, it is clear that more research, in particular obtaining independent estimates of feature and binding memory, is needed on this topic. Simultaneous assessments of the different theoretical “levels of binding” (Zimmer et al., 2006) such as these are required to better establish when older adults do and do not struggle to bind information in working memory.

Our findings and review of the literature suggest that conjunctive binding mechanisms hold little promise in understanding the decline of working memory with age. So what potential mechanisms may be evoked to explain the pronounced overall effect of age observed on tasks like the one used here? Recent work has suggested that a deficit in top-down control of attention leads to older adults encoding more irrelevant information into working memory (see Sander et al., 2012, for a recent review). For example, Jost, Bryck, Vogel, and Mayr (2011) provide electrophysiological evidence (using the contralateral delay activity) that during the initial encoding of items in a change-detection task, older adults were less able than younger adults to avoid attending to task-irrelevant stimuli. In our experiment we did not manipulate the presence of distractor stimuli; however, it is interesting to note that in the individual feature conditions when trial types were blocked it was beneficial to ignore the task-irrelevant feature in favor of the task-relevant one. The fact that we observed no modulation of the overall age effect on change-detection performance by block type (mixed vs. blocked) suggests that older adults did not struggle to do this kind of filtering. That being said, another type of interference may be responsible for much of the observed age effect; namely, proactive interference (PI). In change-detection tasks groups of younger adults have been found to make more false alarms to lure items taken from memory arrays one or two trials previous to the current one, suggesting failure to update the contents of working memory (Hartshorne, 2008; Makovski & Jiang, 2008). Indeed age differences in the ability to counter PI do appear to make a large contribution to age differences on working memory tasks; however, age differences remain even in conditions of low PI (Bowles & Salthouse, 2003; Emery, Hale, & Myerson, 2008). The extent to which failure to update the contents of working memory on a trial-to-trial basis contributes to age differences in change detection as an interesting avenue for future work.

Recent work employing paradigms that obtain more fine-grained information from participant responses suggests that older adults’ working memory representations are less precise. Older adults appear to require larger magnitude changes in order to accurately perform change-detection tasks (Noack, Lovden, & Lindenberger, 2012) and tend to exhibit greater variability in recalling features following short delays (Peich et al., 2013; Pertzov et al., 2015). The stimuli typically used in change detection tasks are categorically distinct and for younger adults appear to be sufficient to support an all-or-nothing discrimination process (Donkin et al., 2014; Rouder et al., 2008). Nevertheless, it is not clear that this is also the case for older adults and future work should aim to address this. Thus the lower overall performance observed here and elsewhere (e.g., Brockmole et al., 2008; Brown et al., 2017; Cowan et al., 2006) may reflect difficulty in comparing the probe item to representations distorted by noise (particularly in the shape condition of Experiment 1; see Figure 2A). In relation to the discussion of relational binding above, it is inter-
detecting changes to feature conjunctions when these trials were al. (2006), who found that older adults were less accurate at increasing the amount of time given to study objects may allow (Hardman & Cowan, 2015). Finally, it has been suggested that adding additional feature load increases the binding cost performance (Luck & Vogel, 1997), recent work clearly shows of to-be-remembered features per object had no effect on magnitude of the binding cost. Likely candidates include the opportunity for verbal rehearsal of the to-be-remembered features, which may disproportionately benefit performance for individual features (Peterson & Naveh-Benjamin, 2016). In fact, more detailed modeling of responses suggests that, actually, capacity for bindings is much smaller than capacity for individual features (Bays, Wu, & Husain, 2011; Cowan, Blume, & Saults, 2013). However performance is characterized, it is clear that there are factors mediating the magnitude of the binding cost. Likely candidates include the Probing why some functions appear more susceptible to age than others holds great promise for our understanding of cognitive aging and cognition more generally (see Logie, Horne, & Petit, 2015). Finally there is inconsistency in the working-memory literature regarding the efficacy of feature binding in general. In Experiment 1, relative to shape only, which was the most difficult individual feature condition, we found no performance cost to binding (see also Brown et al., 2017); however, this is not always the case, as sometimes a clear binding cost is found (e.g., Allen et al., 2006; Peterson & Naveh-Benjamin, 2016). In fact, more detailed modeling of responses suggests that, actually, capacity for bindings is much smaller than capacity for individual features (Bays, Wu, & Husain, 2011; Cowan, Blume, & Saults, 2013). However performance is characterized, it is clear that there are factors mediating the magnitude of the binding cost. Likely candidates include the opportunity for verbal rehearsal of the to-be-remembered features, which may disproportionately benefit performance for individual features (Peterson & Naveh-Benjamin, 2016). In the present study we used stimuli deliberately constructed to minimize the influence of verbal recoding (Brockmole et al., 2008), arguably providing a better assessment of visual feature binding. The number of to-be-remembered items appears to be another important factor in the magnitude of the binding cost. Increasing the array size appears to increase the likelihood that memory representations will contain partial information (e.g., knowledge of color only and not shape, Cowan et al., 2013). Further, while early work suggested that the number of to-be-remembered features per object had no effect on performance (Luck & Vogel, 1997), recent work clearly shows that adding additional feature load increases the binding cost (Hardman & Cowan, 2015). Finally, it has been suggested that increasing the amount of time given to study objects may allow individuals to engage elaborative encoding strategies that strengthen knowledge of feature combinations (Allen et al., 2006), although the evidence bearing on this is far from compelling (see Rhodes et al., 2016).

Conclusions

The present work followed up on the experiments of Cowan et al. (2006), who found that older adults were less accurate at detecting changes to feature conjunctions when these trials were mixed with trials on which individual features only could have changed, relative to when these trials were presented in separate blocks. Using stimuli defined by color–shape (Experiment 1) and color–location (Experiment 2) conjunction, and improving methods of probing VWM wherever possible, we found this not to be the case. The present data suggest that discrimination of binding changes is no more affected by healthy aging than the discrimination of changes to individual features, and that this is not modulated by block type. In addition, our review of the literature suggests that the claims of early studies, that older adults have a specific feature binding deficit in working memory, lacked the appropriate evidence. While inconsistent results still remain to be explained, especially with regards to location binding, the present findings help to rule out a potential mediator of age-related change to visual working memory.

References


Received September 13, 2016
Revision received December 12, 2016
Accepted December 21, 2016