Supplement to Healthy Ageing and Visual Working Memory: The Effect of Mixing Feature and Conjunction Changes.

Stephen Rhodes, Mario A. Parra, Nelson Cowan, and Robert H. Logie

December 10, 2016

Contents

1 Analysis of Accuracy

1.1 Model and Approach

For raw trial level data, where responses are either correct or incorrect, we apply a hierarchical generalized linear model to estimate age differences in accuracy. This avoids a number of problems associated with analyzing categorical data with standard techniques (such as ANOVA), and most importantly, for our present purposes, avoids the potential for spurious evidence for interaction effects (Dixon, 2008). This analysis serves to estimate the *magnitude* of age differences in change detection accuracy and assess the extent to which this varies across trial types (color change, shape change, binding change) and block types (mixed, blocked).

In the analysis of raw accuracy the log odds of a correct response on a given trial was modelled as a linear combination of a grand mean parameter and deflections from the grand mean that represent main- and interaction-effects of our experimental factors (see Jaeger, 2008, for an introduction to logit models). These deflections were constrained to sum-to-zero via the use of effects coded variables (Ntzoufras, 2009). In effects coding, as with many other coding schemes, we are limited to $I-1$ indicator variables, where I is the number of levels in a given factor. One level is set to -1 for all indicator variables and acts as the reference level; the $I-1$ variables reflect the deflection from the mean attributable to each remaining level with that level coded 1 and the rest (except for the reference level) coded 0. The resulting coefficient associated with an indicator variable reflects the deflection from the mean associated with the positively coded factor level. These coefficients are constrained to sum-to-zero and the corresponding coefficient for the reference level is the negative sum of the $I-1$ coefficients associated with a given factor. Interaction variables are analysed similarly and reflect the product of these effects coded indicator variables (see, Ntzoufras, 2009). These effects coded variables for each trial in the data set were stored in the design matrix, $\mathbf{X}$. The notes accompanying the tables of results below detail the coding schemes used in this analysis.

Finally, as we had repeated measures from the same individuals across conditions, we modelled an additional effect of participant reflecting the fact that individuals will vary in their overall level of performance. Participant effects were assumed to be drawn from a normal distribution with a mean of zero and a standard deviation estimated from the data (as is typical in hierarchical modelling. For example see,

Gelman & Hill, 2007). Our model can be summarized as follows:

$$y_i \sim \text{Bernoulli}(\pi_i), \quad \text{for } i \text{ in } 1, \ldots, t$$
$$\text{logit}(\pi_i) = \log\left(\frac{\pi_i}{1-\pi_i}\right) = \beta_0 + \mathbf{X}_i\boldsymbol{\beta} + s_{j[i]}$$
$$s_j \sim \text{Normal}(0,\, \sigma_s), \quad \text{for } j \text{ in } 1, \ldots, n$$

where t is the number observations (or trials) and n is the number of participants. The first line gives the likelihood function; each trial is assumed to be a Bernoulli random variable with the underlying probability of success, π_i, determined by the second line. This second line models the log odds of the underlying success probability parameter as a linear combination of three components; (1) a grand mean parameter, β_0, (2) deflections from the grand mean represented by the parameter vector, $\boldsymbol{\beta}$, which is multiplied by the row in the matrix $\mathbf{X}$ containing the effects coded indicator variables for the corresponding trial, and (3) an additional participant level effect. The final line reflects the assumption that participant effects are normally distributed with a mean of 0 and standard deviation, σ_s.

As our estimation is Bayesian prior distributions must be placed on model parameters. For fixed effects we follow the suggestions of Gelman, Jakulin, Pittau, and Su (2008) and use a mildly informative Cauchy prior:

$$\beta_l \sim \text{Cauchy}(0,\, 2.5), \quad \text{for } l \text{ in } 1, \ldots, P$$

where P is the number of effects-coded variables in the design matrix. This mildly informative distribution reflects the belief that effects on the log odds scale will usually fall within a restricted range ($\pm$ 2.5) but, due to the Cauchy's heavy tails, does not rule out the possibility of larger effects. A grand mean of 0 in log odds space implies that average performance is at chance, therefore, to reflect our prior expectation that overall performance is likely to be above chance, our prior on β_0 was also a Cauchy distribution centered at 1 (corresponding to approximately 0.73 in proportion space) with scale of 2.5. Finally, as is common in Bayesian hierarchical modelling, we place a prior on the *precision* of the normally distributed participant level effect rather than the standard deviation. Precision is related to standard deviation via $\tau_s = 1/\sigma_s^2$ and we used the following vague prior:

$$\tau_s \sim \text{Gamma}(1.01005,\, 0.1005012)$$

which has a mode of 0.1 (low precision, high standard deviation) and standard deviation of 10 (see, Kruschke, 2015), thus is sufficiently broad on the log odds scale.

We took 50000 samples from the posterior distribution across 4 independent MCMC chains after a burn-in period of 5000 samples using JAGS (Just Another Gibbs Sampler, Plummer et al., 2003) via `rjags` (Plummer, 2015) in `R` (R Core Team, 2015). A multivariate BGR statistic of 1 was taken to indicate that the chains had converged on a stable distribution (Brooks & Gelman, 1998). It is common to thin MCMC chains (i.e. only retain every k^{th} sample) to reduce auto-correlation, however following the suggestions of Link and Eaton (2012) we do not do this and instead retain the whole large sample, which is more representative of the true posterior distribution than a smaller, thinned chain. Further, we did our best to ensure that the *effective sample size* (ESS, Kass, Carlin, Gelman, & Neal, 1998), the number of independent samples accounting for autocorrelation, was at least 10000 for the deflection parameters (as per the recommendations of Kruschke, 2015). The deflection parameters (contained in $\boldsymbol{\beta}$) are of primary interest and indicate the size of effects/ interactions in the data, thus we use the resulting posterior samples of these coefficients to construct specific contrasts that test hypotheses about patterns of performance. More detail on this approach and the JAGS code used to produce the model can be found at http://stephenrho.github.io/rjags-model.html.

1.2 Results

1.2.1 Experiment 1

The results of our model estimation for the blocked condition of Experiment 1 are presented in Table 1. To make sense of this complex pattern of results here we report a set of specific contrasts that test the key questions of the present research. These contrasts are presented with their highest density intervals (HDIs Kruschke, 2015) converted back to proportion scale (via the logistic transformation), whereas the coefficients presented in Tables are given on the log odds scale used in the modelling. In the blocked condition there was a clear effect of age with older adults' accuracy approximately 0.107 [0.063, 0.154] (95% HDI) lower than that of younger adults. There was no indication that this age difference varied across memory conditions (color: 0.124 [0.080, 0.170], shape: 0.080 [0.023, 0.137], binding: 0.100 [0.045, 0.157]). Indeed specific contrasts revealed that the accuracy difference between features (average of shape and color) and binding accuracy was roughly equivalent (0.002 [-0.039, 0.044]). This was also the case when comparing the feature condition with lowest accuracy, shape, to binding (-0.020 [-0.068, 0.029]) as is also common in studies like these (Brockmole, Parra, Della Sala, & Logie, 2008; Brown & Brockmole, 2010).

Table 1: Posterior quantities from logit model for the Blocked condition of Experiment 1

			95% HDI		
Parameter	Mean	Median	lower	upper	ESS
β_0	0.960	0.959	0.840	1.075	2780.328
$\boldsymbol{\beta}_1$: (1) Shape	-0.252	-0.252	-0.319	-0.186	19687.658
$\boldsymbol{\beta}_2$: (2) Binding	-0.194	-0.194	-0.262	-0.127	19149.694
$\boldsymbol{\beta}_3$: (3) SS6	-0.387	-0.387	-0.437	-0.338	25007.410
$\boldsymbol{\beta}_4$: (4) Older Group	-0.269	-0.269	-0.383	-0.154	2724.668
$\boldsymbol{\beta}_5$: (5) Change	0.038	0.038	-0.011	0.087	24298.394
$\boldsymbol{\beta}_6$: 1×3	0.092	0.092	0.027	0.159	20216.965
$\boldsymbol{\beta}_7$: 2×3	0.091	0.091	0.023	0.159	19863.237
$\boldsymbol{\beta}_8$: 1×4	0.086	0.086	0.021	0.154	19359.511
$\boldsymbol{\beta}_9$: 2×4	0.037	0.037	-0.029	0.106	19492.504
$\boldsymbol{\beta}_{10}$: 1×5	-0.198	-0.197	-0.265	-0.131	19549.995
$\boldsymbol{\beta}_{11}$: 2×5	0.021	0.021	-0.046	0.089	19457.065
$\boldsymbol{\beta}_{12}$: 3×4	0.095	0.095	0.046	0.145	25369.488
$\boldsymbol{\beta}_{13}$: 3×5	0.193	0.193	0.145	0.243	23725.779
$\boldsymbol{\beta}_{14}$: 4×5	-0.051	-0.051	-0.100	-0.002	24857.852
$\boldsymbol{\beta}_{15}$: $1 \times 3 \times 4$	-0.037	-0.037	-0.102	0.031	18689.917
$\boldsymbol{\beta}_{16}$: $2 \times 3 \times 4$	0.046	0.046	-0.021	0.113	18811.420
$\boldsymbol{\beta}_{17}$: $1 \times 3 \times 5$	-0.065	-0.065	-0.132	0.002	18661.653
$\boldsymbol{\beta}_{18}$: $2 \times 3 \times 5$	0.161	0.161	0.094	0.230	19234.499
$\boldsymbol{\beta}_{19}$: $1 \times 4 \times 5$	-0.042	-0.042	-0.110	0.023	19796.697
$\boldsymbol{\beta}_{20}$: $2 \times 4 \times 5$	-0.004	-0.004	-0.071	0.063	20054.320
$\boldsymbol{\beta}_{21}$: $3 \times 4 \times 5$	-0.012	-0.012	-0.061	0.037	25091.911
$\boldsymbol{\beta}_{22}$: $1 \times 3 \times 4 \times 5$	0.039	0.039	-0.026	0.107	19263.891
$\boldsymbol{\beta}_{23}$: $2 \times 3 \times 4 \times 5$	-0.050	-0.050	-0.118	0.017	18554.812
σ_s	0.366	0.362	0.276	0.462	11770.313

Note: The effects coded variables were as follows: (1) Shape = 1, Binding = 0, Colour = -1, (2) Shape = 0, Binding = 1, Colour = -1, (3) SS3 = -1, SS6 = 1, (4) Younger = -1, Older = 1, (5) No-Change = -1, Change = 1. Interaction contrasts were products of these effects coded variables.

The results from the mixed condition are presented in full in Table 2. Relative to the blocked condition there was a similar pattern of performance in the mixed condition with a large age-effect on accuracy (0.108 [0.079, 0.138]) with no clear modulation by trial type (color: 0.103 [0.073, 0.133], shape: 0.100 [0.050, 0.152], binding: 0.132 [0.081, 0.183], no-change: 0.072 [0.036, 0.108]). The contrast of accuracy for binding changes relative to the average of the two feature changes revealed a negligible difference in accuracy that was not credibly different from zero (-0.044 [-0.098, 0.009]) and this was also the case when just comparing shape and binding (-0.032 [-0.097, 0.033]).

Table 2: Posterior quantities from logit model for the Mixed condition of Experiment 1

			95% HDI		
Parameter	Mean	Median	lower	upper	ESS
β_0	1.167	1.167	1.082	1.251	6140.908
$\boldsymbol{\beta}_1$: (1) Shape	-0.416	-0.416	-0.516	-0.315	20718.657
$\boldsymbol{\beta}_2$: (2) Binding	-0.422	-0.422	-0.523	-0.325	21137.377
$\boldsymbol{\beta}_3$: (3) No-Change	-0.408	-0.408	-0.490	-0.330	10590.778
$\boldsymbol{\beta}_4$: (4) SS6	-0.343	-0.343	-0.408	-0.278	9492.274
$\boldsymbol{\beta}_5$: (5) Older Group	-0.300	-0.300	-0.383	-0.218	6656.232
$\boldsymbol{\beta}_6$: 1×4	0.305	0.305	0.208	0.406	19677.802
$\boldsymbol{\beta}_7$: 2×4	0.306	0.306	0.210	0.407	20769.862
$\boldsymbol{\beta}_8$: 3×4	-0.343	-0.343	-0.424	-0.263	10287.139
$\boldsymbol{\beta}_9$: 1×5	0.069	0.069	-0.031	0.167	20621.771
$\boldsymbol{\beta}_{10}$: 2×5	-0.005	-0.005	-0.103	0.095	18681.638
$\boldsymbol{\beta}_{11}$: 3×5	0.134	0.134	0.055	0.214	11079.891
$\boldsymbol{\beta}_{12}$: 4×5	0.020	0.020	-0.047	0.084	9411.368
$\boldsymbol{\beta}_{13}$: $1 \times 4 \times 5$	0.002	0.003	-0.095	0.106	20324.215
$\boldsymbol{\beta}_{14}$: $2 \times 4 \times 5$	-0.036	-0.036	-0.134	0.066	20673.431
$\boldsymbol{\beta}_{15}$: $3 \times 4 \times 5$	0.180	0.180	0.100	0.261	10208.954
σ_s	0.184	0.183	0.113	0.260	3565.263

Note: The effects coded variables were as follows: (1) Shape = 1, Binding = 0, No-Change = 0, Colour = -1, (2) Shape = 0, Binding = 1, No-Change = 0, Colour = -1, (3) Shape = 0, Binding = 0, No-Change = 1, Colour = -1, (4) SS3 = -1, SS6 = 1, (5) Younger = -1, Older = 1. Interaction contrasts were products of these effects coded variables.

To compare the two groups (mixed versus blocked presentation) directly we combined the data from change trials and estimated a separate logit model (see Table 3). This, again, revealed a large effect of age-group (0.120 [0.081, 0.159]) that was not modulated by the block type (blocked: 0.126 [0.067, 0.185], mixed: 0.113 [0.062, 0.165], difference: 0.013 [-0.066, 0.090]). Crucially, however, we wanted to examine whether the type of block interacted with age-group in determining difference between shape only and binding change detection accuracy. In this analysis the difference between shape and binding performance in the blocked condition was approximately 0.006 [-0.061, 0.074] smaller in the older group, whereas in the mixed condition this difference was -0.032 [-0.099, 0.031] larger in the older group. Thus contrasting the two conditions we find that the age difference in the binding cost is approximately 0.038 [-0.055, 0.132] larger in the mixed condition than in the blocked condition; the posterior mean is clearly negligible and the HDIs firmly include zero. It is worth emphasizing that the estimated difference between shape and binding in the older group of the mixed condition was 0.019 [-0.029, 0.067], whereas for the younger group this was -0.013 [-0.057, 0.031] (with negative values indicating higher binding accuracy). The negligible difference between feature and binding accuracy for older adults can hardly be considered indicative of a specific age-related binding deficit in the mixed condition. There was a clear overall effect of set size on accuracy (0.077 [0.057, 0.097]) but there was no clear evidence that increasing the number of to-be-remembered items affected the pattern of age differences in performance. Nor was there any indication of any of the other possible interactions including age-group (see Table 3).

Table 3: Posterior quantities from logit model comparing change trials in the mixed and blocked conditions of Experiment 1

			95% HDI		
Parameter	Mean	Median	lower	upper	ESS
β_0	1.179	1.180	1.070	1.288	3552.891
$\boldsymbol{\beta}_1$: (1) Shape	-0.510	-0.510	-0.579	-0.438	18647.387
$\boldsymbol{\beta}_2$: (2) Binding	-0.373	-0.373	-0.444	-0.301	18406.937
$\boldsymbol{\beta}_3$: (3) SS6	-0.215	-0.214	-0.271	-0.160	15768.569
$\boldsymbol{\beta}_4$: (4) Older Group	-0.336	-0.336	-0.441	-0.224	3574.116
$\boldsymbol{\beta}_5$: (5) Mixed	0.167	0.167	0.061	0.277	3579.921
$\boldsymbol{\beta}_6$: 1×3	0.110	0.110	0.039	0.182	17853.982
$\boldsymbol{\beta}_7$: 2×3	0.225	0.225	0.155	0.298	18700.585
$\boldsymbol{\beta}_8$: 1×4	0.080	0.080	0.009	0.152	17922.860
$\boldsymbol{\beta}_9$: 2×4	0.036	0.036	-0.036	0.108	19495.189
$\boldsymbol{\beta}_{10}$: 1×5	-0.055	-0.055	-0.123	0.018	17230.747
$\boldsymbol{\beta}_{11}$: 2×5	-0.199	-0.198	-0.270	-0.127	18597.355
$\boldsymbol{\beta}_{12}$: 3×4	0.021	0.021	-0.036	0.076	15911.808
$\boldsymbol{\beta}_{13}$: 3×5	-0.018	-0.018	-0.074	0.036	15803.075
$\boldsymbol{\beta}_{14}$: 4×5	-0.011	-0.011	-0.121	0.096	3583.536
$\boldsymbol{\beta}_{15}$: $1 \times 3 \times 4$	0.033	0.033	-0.035	0.106	18225.172
$\boldsymbol{\beta}_{16}$: $2 \times 3 \times 4$	0.011	0.011	-0.062	0.081	17924.071
$\boldsymbol{\beta}_{17}$: $1 \times 3 \times 5$	0.083	0.083	0.012	0.153	19307.092
$\boldsymbol{\beta}_{18}$: $2 \times 3 \times 5$	-0.030	-0.030	-0.101	0.041	17477.389
$\boldsymbol{\beta}_{19}$: $1 \times 4 \times 5$	0.034	0.034	-0.039	0.102	18294.783
$\boldsymbol{\beta}_{20}$: $2 \times 4 \times 5$	0.003	0.003	-0.071	0.073	17578.808
$\boldsymbol{\beta}_{21}$: $3 \times 4 \times 5$	-0.063	-0.063	-0.119	-0.007	16426.652
$\boldsymbol{\beta}_{22}$: $1 \times 3 \times 4 \times 5$	0.031	0.031	-0.042	0.097	18530.034
$\boldsymbol{\beta}_{23}$: $2 \times 3 \times 4 \times 5$	0.015	0.015	-0.057	0.088	17704.231
σ_s	0.463	0.461	0.375	0.555	10094.882

Note: The effects coded variables were as follows: (1) Shape = 1, Binding = 0, Colour = -1, (2) Shape = 0, Binding = 1, Colour = -1, (3) SS3 = -1, SS6 = 1, (4) Younger = -1, Older = 1, (5) Blocked = -1, Mixed = 1. Interaction contrasts were products of these effects coded variables.

1.2.2 Experiment 2

For the blocked condition of Experiment 2 the results of the logit model are given in Table 4. There was a clear age difference in overall accuracy (0.078 [0.046, 0.111]) and crucially this difference did not depend on memory condition (color: 0.060 [0.022, 0.098], location: 0.082 [0.046, 0.117], binding: 0.093 [0.050, 0.137]). Our specific contrasts between location only and binding showed that the age difference was not credibly different from zero (-0.012 [-0.051, 0.026]). The contrast with the color only condition suggested that the age difference was slightly smaller in the color condition relative to binding (-0.033 [-0.073, 0.007]), although once again this was not clearly different from zero.

Overall the pattern was similar—if not even more out of step with a disproportionate effect of age on conjunction change detection—in the mixed condition (see Table 5). The effect of age on change detection accuracy was clearly non-zero (0.087 [0.049, 0.127]) and there was no modulation by the type of change trial (color: 0.092 [0.044, 0.142], location: 0.102 [0.060, 0.145], binding: 0.099 [0.047, 0.153]), although there was no clear age-effect on no-change trials (0.029 [-0.021, 0.080]). Contrasting accuracy for location versus binding change (0.002 [-0.048, 0.054]) or for color versus binding change (-0.008 [-0.061, 0.045]) shows that any modulation of age difference by change-type is trivial.

Finally accuracy for change trials in the mixed and blocked conditions was combined to directly compare the two (Table 6). There was a clear age difference in the accuracy of change detection (0.120 [0.081, 0.159]) and the type of block did not appear to affect this (blocked: 0.073 [0.027, 0.120], mixed: 0.099 [0.051, 0.149], difference: -0.026 [-0.094, 0.042]). More importantly, the age difference in the binding condition was somewhat larger than that in the individual feature conditions (binding: 0.106 [0.061, 0.150]; color: 0.069 [0.030, 0.106], location: 0.082 [0.049, 0.116]). The contrast between color and binding accuracy was approximately -0.037 [-0.074, -0.001] larger for the older group than the younger group (for location this was -0.025 [-0.060, 0.010]). Importantly, it appeared that this age-group by change type interaction (color vs binding) was slightly larger in the blocked condition (-0.064 [-0.115, -0.011]) relative to mixed (-0.009 [-0.060, 0.043]). However, directly contrasting the two conditions revealed that, while suggestive of a larger age difference in the blocked condition, this difference was not credibly different from zero (-0.055 [-0.127, 0.019]; -0.047 [-0.119, 0.022] for location vs binding). Nevertheless this is in contrast to the findings of Cowan, Naveh-Benjamin, Kilb, and Saults (2006). It may be that younger and older participants adopted different response biases in the blocked condition, when the type of change possible on a given trial was known.

Table 4: Posterior quantities from logit model for the Blocked condition of Experiment 2

			95% HDI		
Parameter	Mean	Median	lower	upper	ESS
β_0	1.775	1.774	1.647	1.902	3436.526
$\boldsymbol{\beta}_1$: (1) Location	0.179	0.179	0.089	0.272	12392.361
$\boldsymbol{\beta}_2$: (2) Binding	-0.238	-0.238	-0.323	-0.152	12300.576
$\boldsymbol{\beta}_3$: (3) SS6	-0.608	-0.608	-0.672	-0.546	19626.783
$\boldsymbol{\beta}_4$: (4) Older Group	-0.315	-0.315	-0.443	-0.187	3523.703
$\boldsymbol{\beta}_5$: (5) Change	0.175	0.174	0.112	0.238	19083.220
$\boldsymbol{\beta}_6$: 1×3	0.168	0.168	0.077	0.256	13722.601
$\boldsymbol{\beta}_7$: 2×3	-0.011	-0.011	-0.096	0.074	12558.764
$\boldsymbol{\beta}_8$: 1×4	-0.057	-0.057	-0.150	0.033	12079.410
$\boldsymbol{\beta}_9$: 2×4	-0.005	-0.005	-0.092	0.081	12195.434
$\boldsymbol{\beta}_{10}$: 1×5	0.041	0.041	-0.049	0.132	13563.203
$\boldsymbol{\beta}_{11}$: 2×5	-0.070	-0.069	-0.157	0.015	12683.714
$\boldsymbol{\beta}_{12}$: 3×4	-0.005	-0.006	-0.069	0.056	20317.812
$\boldsymbol{\beta}_{13}$: 3×5	0.150	0.150	0.088	0.214	17520.950
$\boldsymbol{\beta}_{14}$: 4×5	-0.035	-0.035	-0.098	0.028	18471.498
$\boldsymbol{\beta}_{15}$: $1 \times 3 \times 4$	-0.102	-0.102	-0.193	-0.010	13239.213
$\boldsymbol{\beta}_{16}$: $2 \times 3 \times 4$	0.106	0.106	0.022	0.193	12631.022
$\boldsymbol{\beta}_{17}$: $1 \times 3 \times 5$	0.039	0.039	-0.056	0.126	13267.757
$\boldsymbol{\beta}_{18}$: $2 \times 3 \times 5$	0.099	0.098	0.014	0.186	12320.768
$\boldsymbol{\beta}_{19}$: $1 \times 4 \times 5$	0.038	0.038	-0.050	0.131	13393.196
$\boldsymbol{\beta}_{20}$: $2 \times 4 \times 5$	-0.071	-0.071	-0.154	0.018	12338.476
$\boldsymbol{\beta}_{21}$: $3 \times 4 \times 5$	-0.044	-0.044	-0.107	0.019	19844.658
$\boldsymbol{\beta}_{22}$: $1 \times 3 \times 4 \times 5$	-0.043	-0.043	-0.130	0.050	12542.040
$\boldsymbol{\beta}_{23}$: $2 \times 3 \times 4 \times 5$	0.027	0.027	-0.060	0.111	13016.536
σ_s	0.392	0.388	0.294	0.499	10589.907

Note: The effects coded variables were as follows: (1) Location = 1, Binding = 0, Colour = -1, (2) Location = 0, Binding = 1, Colour = -1, (3) SS3 = -1, SS6 = 1, (4) Younger = -1, Older = 1, (5) No-Change = -1, Change = 1. Interaction contrasts were products of these effects coded variables.

Table 5: Posterior quantities from logit model for the Mixed condition of Experiment 2

			95% HDI		
Parameter	Mean	Median	lower	upper	ESS
β_0	1.789	1.789	1.636	1.947	2295.380
$\boldsymbol{\beta}_1$: (1) Location	0.411	0.410	0.264	0.556	9169.901
$\boldsymbol{\beta}_2$: (2) Binding	-0.136	-0.137	-0.258	-0.009	12920.896
$\boldsymbol{\beta}_3$: (3) No-Change	-0.398	-0.398	-0.492	-0.308	9531.336
$\boldsymbol{\beta}_4$: (4) SS6	-0.543	-0.542	-0.618	-0.470	9191.074
$\boldsymbol{\beta}_5$: (5) Older Group	-0.354	-0.354	-0.507	-0.197	2153.631
$\boldsymbol{\beta}_6$: 1×4	0.196	0.196	0.048	0.341	8943.044
$\boldsymbol{\beta}_7$: 2×4	0.132	0.133	0.007	0.259	13329.995
$\boldsymbol{\beta}_8$: 3×4	-0.184	-0.184	-0.280	-0.095	8505.681
$\boldsymbol{\beta}_9$: 1×5	-0.200	-0.198	-0.345	-0.051	9114.740
$\boldsymbol{\beta}_{10}$: 2×5	-0.013	-0.013	-0.137	0.113	13122.054
$\boldsymbol{\beta}_{11}$: 3×5	0.264	0.264	0.173	0.359	9322.522
$\boldsymbol{\beta}_{12}$: 4×5	0.073	0.073	0.000	0.147	8898.226
$\boldsymbol{\beta}_{13}$: $1 \times 4 \times 5$	0.056	0.056	-0.092	0.203	8919.237
$\boldsymbol{\beta}_{14}$: $2 \times 4 \times 5$	0.002	0.002	-0.121	0.129	13160.935
$\boldsymbol{\beta}_{15}$: $3 \times 4 \times 5$	-0.026	-0.026	-0.117	0.067	9098.443
σ_s	0.473	0.468	0.360	0.594	13109.712

Note: The effects coded variables were as follows: (1) Location = 1, Binding = 0, No-Change = 0, Colour = -1, (2) Location = 0, Binding = 1, No-Change = 0, Colour = -1, (3) Location = 0, Binding = 0, No-Change = 1, Colour = -1, (4) SS3 = -1, SS6 = 1, (5) Younger = -1, Older = 1. Interaction contrasts were products of these effects coded variables.

The analyses of P_r and B_r reported in the main manuscript probe this issue further.

There was an interaction between age-group and set size in this combined data set such that age-differences in change detection accuracy were larger at set size 6 (0.115 [0.069, 0.162]) as compared to set size 3 (0.060 [0.034, 0.087]; contrast: -0.055 [-0.089, -0.022]). Interestingly there was some suggestion of a disproportionate age-effect for binding changes at set size 3 (contrast with color: -0.058 [-0.097, -0.017]; contrast with the average of feature conditions: -0.047 [-0.083, -0.009]) but not at set size 6 (contrast with color: 0.004 [-0.055, 0.061]; contrast with the average of feature conditions: 0.003 [-0.046, 0.053]). When performing specific contrasts to establish whether the age by change type interaction was *disproportionately* larger at set size 3 the results are inconclusive as zero is among the most credible values (binding versus color only: -0.062 [-0.131, 0.009], versus feature average: -0.050 [-0.111, 0.009]). Nevertheless it is difficult to think of an account of age-related binding deficits that would predict a larger deficit at a smaller set size. There was no clear evidence for any further interactions including age-group (see Table 6).

Table 6: Posterior quantities from logit model comparing change trials in the mixed and blocked conditions of Experiment 2

			95% HDI		
Parameter	Mean	Median	lower	upper	ESS
β_0	2.018	2.018	1.866	2.167	2729.411
$\boldsymbol{\beta}_1$: (1) Location	0.259	0.259	0.162	0.360	12559.754
$\boldsymbol{\beta}_2$: (2) Binding	-0.296	-0.297	-0.387	-0.208	13514.816
$\boldsymbol{\beta}_3$: (3) SS6	-0.484	-0.484	-0.552	-0.418	19886.761
$\boldsymbol{\beta}_4$: (4) Older Group	-0.408	-0.408	-0.562	-0.254	2736.175
$\boldsymbol{\beta}_5$: (5) Mixed	-0.006	-0.006	-0.157	0.147	2667.493
$\boldsymbol{\beta}_6$: 1×3	0.177	0.177	0.079	0.275	12019.178
$\boldsymbol{\beta}_7$: 2×3	0.080	0.080	-0.008	0.170	13270.580
$\boldsymbol{\beta}_8$: 1×4	-0.066	-0.066	-0.165	0.031	12959.336
$\boldsymbol{\beta}_9$: 2×4	-0.002	-0.002	-0.094	0.085	13387.035
$\boldsymbol{\beta}_{10}$: 1×5	0.030	0.030	-0.067	0.129	11666.722
$\boldsymbol{\beta}_{11}$: 2×5	0.020	0.020	-0.068	0.112	12580.812
$\boldsymbol{\beta}_{12}$: 3×4	0.014	0.014	-0.055	0.079	19355.098
$\boldsymbol{\beta}_{13}$: 3×5	-0.014	-0.014	-0.080	0.055	19059.012
$\boldsymbol{\beta}_{14}$: 4×5	-0.059	-0.059	-0.211	0.095	2566.470
$\boldsymbol{\beta}_{15}$: $1 \times 3 \times 4$	-0.050	-0.050	-0.147	0.048	13285.550
$\boldsymbol{\beta}_{16}$: $2 \times 3 \times 4$	0.066	0.066	-0.024	0.153	13288.505
$\boldsymbol{\beta}_{17}$: $1 \times 3 \times 5$	-0.036	-0.035	-0.133	0.062	13456.869
$\boldsymbol{\beta}_{18}$: $2 \times 3 \times 5$	-0.008	-0.008	-0.097	0.081	13824.915
$\boldsymbol{\beta}_{19}$: $1 \times 4 \times 5$	-0.049	-0.049	-0.147	0.051	11452.280
$\boldsymbol{\beta}_{20}$: $2 \times 4 \times 5$	0.079	0.079	-0.011	0.169	12187.668
$\boldsymbol{\beta}_{21}$: $3 \times 4 \times 5$	0.068	0.068	-0.001	0.134	19925.194
$\boldsymbol{\beta}_{22}$: $1 \times 3 \times 4 \times 5$	0.098	0.098	-0.001	0.197	12323.468
$\boldsymbol{\beta}_{23}$: $2 \times 3 \times 4 \times 5$	-0.071	-0.071	-0.160	0.020	12798.248
σ_s	0.669	0.666	0.554	0.795	11302.557

Note: The effects coded variables were as follows: (1) Location = 1, Binding = 0, Colour = -1, (2) Location = 0, Binding = 1, Colour = -1, (3) SS3 = -1, SS6 = 1, (4) Younger = -1, Older = 1, (5) Blocked = -1, Mixed = 1. Interaction contrasts were products of these effects coded variables.

2 Tables of Bayes Factor Output

2.1 Experiment 1

The full results of the default Bayes factor analysis of P_r and B_r for Experiment 1 are given in Tables 7 and 8, respectively. Log Bayes factors are given to reduce rounding error. A $\log(B) > 1.099$ or < -1.099 corresponds to a $B > 3$ or $< 1/3$, respectively.

Table 7: Bayes factors for analysis of P_r in Experiment 1

Omitted Component	$\log(B)$	% Error
AG × BT × MC × SS	1.383	2.944
AG × BT × SS	1.724	3.232
AG × BT × MC	1.569	4.215
AG × MC × SS	-1.179	2.465
BT × MC × SS	2.699	3.291
AG × BT	1.212	3.219
AG × SS	-3.527	4.088
AG × MC	2.741	4.077
BT × SS	-0.088	3.073
BT × MC	3.132	2.415
MC × SS	-3.357	3.917
AG	-15.669	3.213
BT	1.646	3.045
SS	-193.915	3.532
MC	-104.399	2.822

Note: AG = Age Group, BT = Block Type, MC = Memory Condition, SS = Set Size.

Table 8: Bayes factors for analysis of B_r in Experiment 1

Omitted Component	log(B)	% Error
AG × BT × MC × SS	1.647	3.570
AG × BT × SS	0.254	4.059
AG × BT × MC	1.656	3.940
AG × MC × SS	2.430	3.638
BT × MC × SS	-1.352	3.890
AG × BT	1.492	3.415
AG × SS	-4.568	3.123
AG × MC	2.448	4.461
BT × SS	1.865	3.144
BT × MC	-16.790	3.683
MC × SS	-16.612	4.600
AG	0.518	3.641
BT	-2.662	3.181
SS	-44.001	4.810
MC	-65.883	3.197

Note: AG = Age Group, BT = Block Type, MC = Memory Condition, SS = Set Size.

2.2 Experiment 2

Full results of the BANOVAs on P_r and B_r in Experiment 2 summarized in the manuscript are presented in Tables 9 and 10, respectively.

Table 9: Bayes factors for analysis of P_r in Experiment 2

Omitted Component	log(B)	% Error
AG × BT × MC × SS	0.891	4.307
AG × BT × SS	-0.010	3.878
AG × BT × MC	2.407	3.641
AG × MC × SS	1.115	3.396
BT × MC × SS	1.707	4.024
AG × BT	1.253	4.809
AG × SS	-1.210	3.373
AG × MC	2.792	3.480
BT × SS	1.725	3.356
BT × MC	0.883	3.561
MC × SS	-6.418	3.931
AG	-9.196	3.441
BT	0.799	3.097
SS	-235.006	3.876
MC	-14.647	3.812

Note: AG = Age Group, BT = Block Type, MC = Memory Condition, SS = Set Size.

Table 10: Bayes factors for analysis of B_r in Experiment 2

Omitted Component	log(B)	% Error
AG × BT × MC × SS	0.575	4.687
AG × BT × SS	0.122	4.745
AG × BT × MC	2.555	4.887
AG × MC × SS	2.191	4.020
BT × MC × SS	2.171	3.805
AG × BT	0.601	4.144
AG × SS	0.683	4.206
AG × MC	2.908	4.340
BT × SS	1.984	3.731
BT × MC	3.218	4.746
MC × SS	0.223	4.083
AG	-1.492	3.779
BT	0.452	3.804
SS	-6.973	4.446
MC	0.382	3.689

Note: AG = Age Group, BT = Block Type, MC = Memory Condition, SS = Set Size.

3 Results of Experiment 2b

As mentioned in the Experiment 2 discussion section of the main manuscript, we conducted an additional experiment in which trials containing changes to colour only or colour-location binding were mixed in the same trial-block. The primary aim of this experiment was to assess whether omitting trials testing memory for location only may induce a bias towards colour only at the expense of the binding between colour and location as this was closer to the design used in Cowan et al. (2006). This also served to replicate the findings of Experiment 2.

3.1 Participants

Twenty-four younger adults (mean age = 20.96, SD = 3.10) and 24 older adults (71.13, 4.13) were recruited from the same populations as Experiment 2. Thirteen of participants in the older group had taken part in Experiment 1 assessing colour-shape binding approximately a year previously. All older adults scored 27 or higher on the MMSE.

3.2 Design, Procedure, and Stimuli

All aspects of this experiment were identical to the mixed condition of Experiment 2 in the main manuscript with the only difference being the omission of trials on which only location changed resulting in a session made up of 12 practice trials and 128 experimental trials split across two blocks.

3.3 Results

3.3.1 Accuracy

Proportion correct (accuracy) across trial types (colour change, binding change, no-change), set sizes, and age-groups is presented in Table 11. The hierarchical model described above was fit to trial level data and a summary of the posterior distribution can be found in Table 12.

As we did in the main manuscript, the posterior chains can be used to set up specific hypotheses tests (see Kruschke, 2015). There was a large age-difference in overall accuracy (0.094 [0.055, 0.131]) and this did not greatly differ across the two change conditions (colour: 0.089 [0.051, 0.129], binding: 0.114 [0.064, 0.164]). Indeed the contrast of colour only and colour-location change detection between our two-groups clearly contained zero within its credible values (-0.025 [-0.074, 0.024]). The age-difference was somewhat reduced for no-change trials (0.057 [0.005, 0.107])

Table 11: Accuracy accross age-groups and experimental factors for Experiment 2b

		Younger	Older
No-Change	3	0.923 (0.014)	0.898 (0.017)
	6	0.665 (0.031)	0.576 (0.025)
Colour	3	0.971 (0.008)	0.919 (0.023)
	6	0.904 (0.016)	0.758 (0.034)
Binding	3	0.938 (0.015)	0.839 (0.026)
	6	0.849 (0.024)	0.727 (0.032)

Note: Standard errors are given in parenthesis.

and a specific contrast shows that the age-difference is approximately 0.044 [0.004, 0.087] larger for change trials relative to no-change (although note that the HDI encompasses a region many may consider practically equivalent to zero). Finally, as in Experiment 2, age-differences were particularly pronounced at set size 6 (0.137 [0.080, 0.195]) relative to set size 3 (0.058 [0.030, 0.086]; contrast: -0.079 [-0.127, -0.032]). Table 12 shows that higher order interaction coefficients all included zero within their (fairly wide) highest density intervals.

3.3.2 Discriminability and Bias

Figures 1 and 2 present estimates of discriminability (P_r) and bias (B_r) for Experiment 2b, respectively, and Tables 13 and 14 present the results of default Bayes factor analyses on these data. As can be clearly gleaned from these tables there was substantial evidence against an age by trial type interaction for both P_r ($B = 4.17$) and B_r ($B = 3.93$).

We conducted an additional analysis to directly compare Experiments 2 and 2b (combining the data sets and omitting the location only data from Experiment 2), the results of which are presented in Tables 15 and 16. This allowed us to better assess whether omitting trials on which location only could change had an effect on the performance of our younger and older groups. Importantly there was substantial evidence against the three way interaction between age, trial type, and experiment for both measures (P_r: $B = 4.8$, B_r: $B = 4.95$). Thus omitting location change trials, as done in Cowan et al. (2006), did not differentially affect older adults' ability to discriminate colour-location conjunction changes.

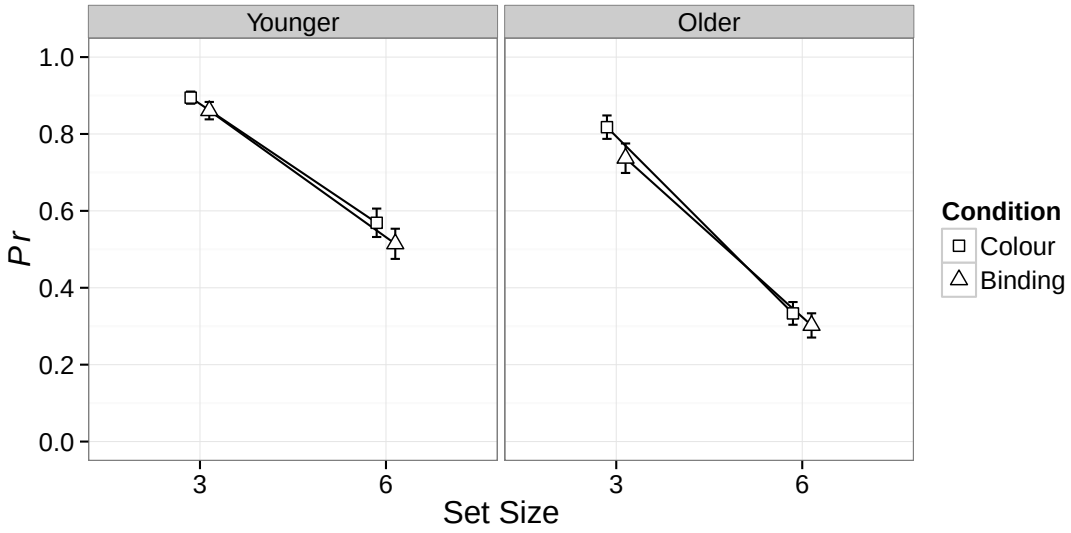

Figure 1: Mean (and ± standard error) estimates of discriminability (P_r) in Experiment 2b.

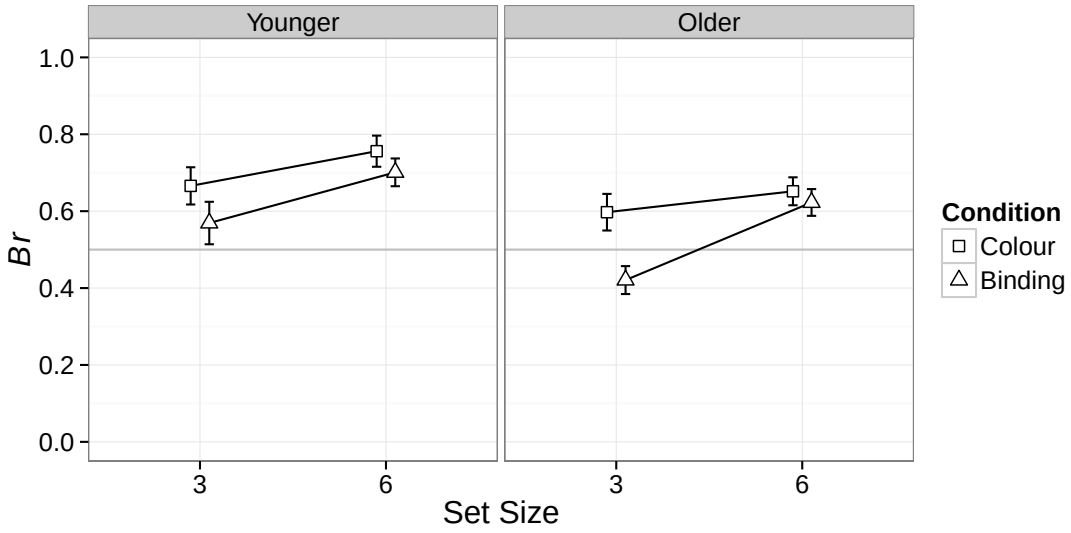

Figure 2: Mean (and ± standard error) estimates of bias (B_r) in Experiment 2b.

Table 12: Posterior quantities from logit model for Experiment 2b

			95% HDI		
Parameter	Mean	Median	lower	upper	ESS
β_0	1.902	1.901	1.745	2.062	3598.411
$\boldsymbol{\beta}_1$: (1) Binding	-0.073	-0.073	-0.203	0.054	15715.895
$\boldsymbol{\beta}_2$: (2) No-Change	-0.439	-0.438	-0.548	-0.327	9320.829
$\boldsymbol{\beta}_3$: (3) SS6	-0.676	-0.676	-0.768	-0.583	10195.538
$\boldsymbol{\beta}_4$: (4) Older Group	-0.411	-0.410	-0.570	-0.244	3357.847
$\boldsymbol{\beta}_5$: 1×3	0.253	0.252	0.123	0.382	16018.744
$\boldsymbol{\beta}_6$: 2×3	-0.268	-0.268	-0.380	-0.158	10051.656
$\boldsymbol{\beta}_7$: 1×4	-0.063	-0.063	-0.190	0.068	15641.582
$\boldsymbol{\beta}_8$: 2×4	0.224	0.224	0.114	0.334	9937.636
$\boldsymbol{\beta}_9$: 3×4	0.019	0.019	-0.072	0.113	10418.598
$\boldsymbol{\beta}_{10}$: $1 \times 3 \times 4$	0.059	0.059	-0.072	0.186	16339.511
$\boldsymbol{\beta}_{11}$: $2 \times 3 \times 4$	-0.038	-0.038	-0.149	0.072	9504.982
σ_s	0.449	0.444	0.318	0.591	7451.405

Note: The effects coded variables were as follows: (1) Binding = 1, No-Change = 0, Colour = -1, (2) Binding = 0, No-Change = 1, Colour = -1, (3) SS3 = -1, SS6 = 1, (4) Younger = -1, Older = 1. Interaction contrasts were products of these effects coded variables.

Table 13: Bayes factors for analysis of P_r in Experiment 2b

Omitted Component	$\log(B)$	% Error
AG × MC × SS	0.728	3.099
AG × MC	1.428	3.798
AG × SS	-5.143	3.141
MC × SS	1.486	3.405
AG	-6.432	2.928
MC	-2.766	3.979
SS	-121.954	3.134

Note: AG = Age Group, MC = Memory Condition, SS = Set Size.

Table 14: Bayes factors for analysis of B_r in Experiment 2b

Omitted Component	log(B)	% Error
AG × MC × SS	0.760	2.923
AG × MC	1.368	3.117
AG × SS	1.470	4.700
MC × SS	-0.091	2.717
AG	-0.872	2.923
MC	-4.304	3.039
SS	-8.992	3.579

Note: AG = Age Group, MC = Memory Condition, SS = Set Size.

Table 15: Bayes factors for analysis of P_r from Experiments 2 and 2b

Omitted Component	log(B)	% Error
AG × Ex × MC × SS	1.280	4.127
AG × Ex × MC	1.568	4.140
AG × Ex × SS	-0.505	3.569
Ex × MC × SS	1.319	3.476
AG × MC × SS	0.686	3.741
AG × Ex	0.947	4.024
Ex × MC	1.025	4.087
Ex × SS	0.628	3.875
AG × MC	1.840	4.367
AG × SS	-3.476	3.353
MC × SS	0.780	3.761
Ex	0.737	3.832
AG	-8.025	4.915
MC	-2.709	3.190
SS	-233.571	3.524

Note: AG = Age Group, Ex = Experiment (2 vs 2b), MC = Memory Condition, SS = Set Size.

Table 16: Bayes factors for analysis of B_r from Experiments 2 and 2b

Omitted Component	log(B)	% Error
AG × Ex × MC × SS	1.298	3.019
AG × Ex × MC	1.599	3.079
AG × Ex × SS	1.451	3.325
Ex × MC × SS	1.510	3.997
AG × MC × SS	0.639	4.941
AG × Ex	1.449	3.613
Ex × MC	0.221	3.509
Ex × SS	-0.657	2.944
AG × MC	1.407	3.716
AG × SS	1.867	3.128
MC × SS	-0.962	3.219
Ex	1.541	3.008
AG	-3.100	3.791
MC	-3.066	3.415
SS	-7.744	3.536

Note: AG = Age Group, Ex = Experiment (2 vs 2b), MC = Memory Condition, SS = Set Size.

References

Brockmole, J. R., Parra, M. A., Della Sala, S., & Logie, R. H. (2008). Do binding deficits account for age-related decline in visual working memory? *Psychonomic Bulletin and Review*, *15*(3), 543–547.

Brooks, S. P., & Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, *7*(4), 434–455.

Brown, L. A., & Brockmole, J. R. (2010). The role of attention in binding visual features in working memory: Evidence from cognitive ageing. *The Quarterly Journal of Experimental Psychology*, *63*(10), 2067–2079.

Cowan, N., Naveh-Benjamin, M., Kilb, A., & Saults, J. S. (2006). Life-span development of visual working-memory: When is feature binding difficult? *Developmental Psychology*, *42*(6), 1089–1102.

Dixon, P. (2008). Models of accuracy in repeated-measures designs. *Journal of Memory and Language*, *59*(4), 447–456.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and hierarchical/multilevel models*. Cambridge, UK: Cambridge University Press.

Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y.-S. (2008). A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics*, 1360–1383.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*(4), 434–446.

Kass, R. E., Carlin, B. P., Gelman, A., & Neal, R. M. (1998). Markov chain Monte Carlo in practice: a roundtable discussion. *The American Statistician*, *52*(2), 93–100.

Kruschke, J. K. (2015). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan*. Academic Press.

Link, W. A., & Eaton, M. J. (2012). On thinning of chains in MCMC. *Methods in Ecology and Evolution*, *3*(1), 112–115.

Ntzoufras, I. (2009). *Bayesian modeling using WinBUGS* (Vol. 1). John Wiley & Sons.

Plummer, M. (2015). rjags: Bayesian Graphical Models using MCMC [Computer software manual]. Retrieved from `http://CRAN.R-project.org/package=rjags` (R package version 3-15)

Plummer, M., et al. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In *Proceedings of the 3rd international workshop*

on distributed statistical computing (Vol. 124, p. 125).

R Core Team. (2015). R: A Language and Environment for Statistical Computing [Computer software manual]. Vienna, Austria. Retrieved from `http://www.R-project.org/`