Supplement to: 'The influence of long-term memory on working memory: Age-differences in proactive facilitation and interference'

Stephen Rhodes[1], Bradley R. Buchsbaum[1,2], & Lynn Hasher[1,2]

[1] Rotman Research Institute, Baycrest Hospital
[2] University of Toronto, Department of Psychology

Draft: 03 May 2021

Contents

Learning performance by working memory condition

Cued recall accuracy during the learning phase of the experiment is presented in Figure 1 separately for each experiment group. The groups are defined by manipulations

---

of the working memory task (retention interval, presence of a distracting task), not the learning phase of the experiment which was identical. Therefore, any differences in degree of learning we observe here reflect assignment of individuals to particular conditions.

Participants studied the 30 pairs in groups of 10 and continued to the final test of learning once accuracy was $\geq 80\%$ or once three loops had been completed. The plot also shows the number of participants who needed to loop through the pairs 2 or 3 times before moving on to the final test (see text next to the points).
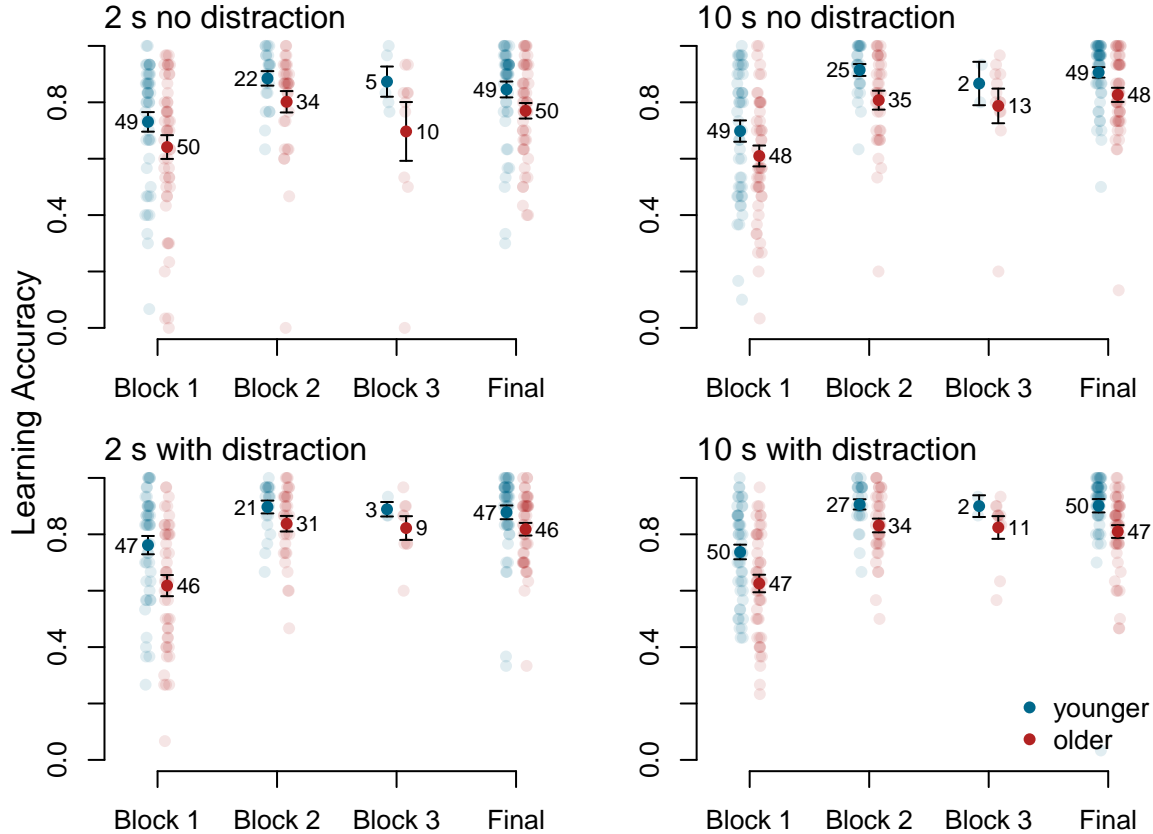

Figure 1. Accuracy during the learning phase split by experiment group. Error bars are within-subjects standard errors. The numbers next to the average points are the number of participants who completed each stage of the learning procedure. Participants continued studying pairs in blocks until they reached 80% or greater cued recall accuracy or three sets had been completed.

All participants completed the final test of learning in which all 30 pairs were tested in a random order by cuing word recall with each image. Results of the full Bayesian generalized mixed effects analysis of final learning accuracy are presented in Table 1. The coefficients for the effects of age-group and interval are credibly different from zero (i.e., zero is not in the 95% highest density interval or 95% HDI). Converting the model fitted values to probabilities, younger adults cued recall accuracy was approximately 0.928 [0.912, 0.942] (95% HDI) and for older adults 0.852 [0.824, 0.879] (difference: 0.076 [0.049, 0.104]). Final cued recall accuracy was also slightly higher for the group in the 10 s interval condition

Table 1

Results of generalized linear mixed effects model for final learning accuracy. Posterior mean, 95% highest density interval, percentage of posterior samples greater than zero, and Bayes factors in favor of the null and alternative.

	Mean	95% HDI Lower	95% HDI Upper	Perc > 0	B_{01}	B_{10}
Intercept	2.16	1.98	2.33	100.00	-	-
Distraction	-0.04	-0.17	0.09	26.78	6.24	0.16
Interval	0.16	0.03	0.29	99.13	0.46	2.19
Group	-0.40	-0.53	-0.27	0.00	<1e-4	>1e+4
Distraction $\times$ Interval	0.12	-0.01	0.25	96.07	1.59	0.63
Distraction $\times$ Group	0.02	-0.11	0.15	61.34	7.07	0.14
Interval $\times$ Group	-0.08	-0.21	0.05	12.51	3.79	0.26
Distraction $\times$ Interval $\times$ Group	0.00	-0.12	0.13	50.82	7.77	0.13

(0.910 [0.891, 0.928]) relative to those in the 2 s condition (0.880 [0.857, 0.903]; difference: 0.030 [0.005, 0.055]). The magnitude of the age-group difference in final learning accuracy was consistent across conditions, as shown in the age group interaction coefficients in Table 1.

Search performance

The first 4 trials of search alone were excluded as practice (leaving 16 observations) along with the 3 practice working memory trials. For the interval condition of 2 s there was one search trial per working memory trial (16 post-practice observations) and for the interval of 10 s there were 4 search trials during the working memory interval (64 post-practice observations). Figure 2 presents accuracy and reaction time in this task

Search accuracy

Search accuracy was analyzed via a generalized linear mixed effects model with the same priors as those used in the analyses in the main manuscript. Fixed effects of section (search alone = 1, search + working memory = -1), age group (older = 1, younger = -1), and interval (10 s = 1, 2 s = -1) were included along with their interactions and random participant intercepts and section effects. Estimates of fixed effects are presented in Table 2. There is a clear main effect of age group, with lower accuracy for older adults (see Figure 2), and an interaction between interval and section. This interaction is driven by better performance for search alone (0.799 [0.734, 0.859]) compared to with the working memory task (0.738 [0.665, 0.809]) in the 2 s condition (difference: 0.061 [0.003, 0.117]), whereas the opposite pattern was seen in the 10 s condition (alone: 0.732 [0.656, 0.804]; with working memory: 0.809 [0.751, 0.859]; difference: -0.076 [-0.127, -0.023]).
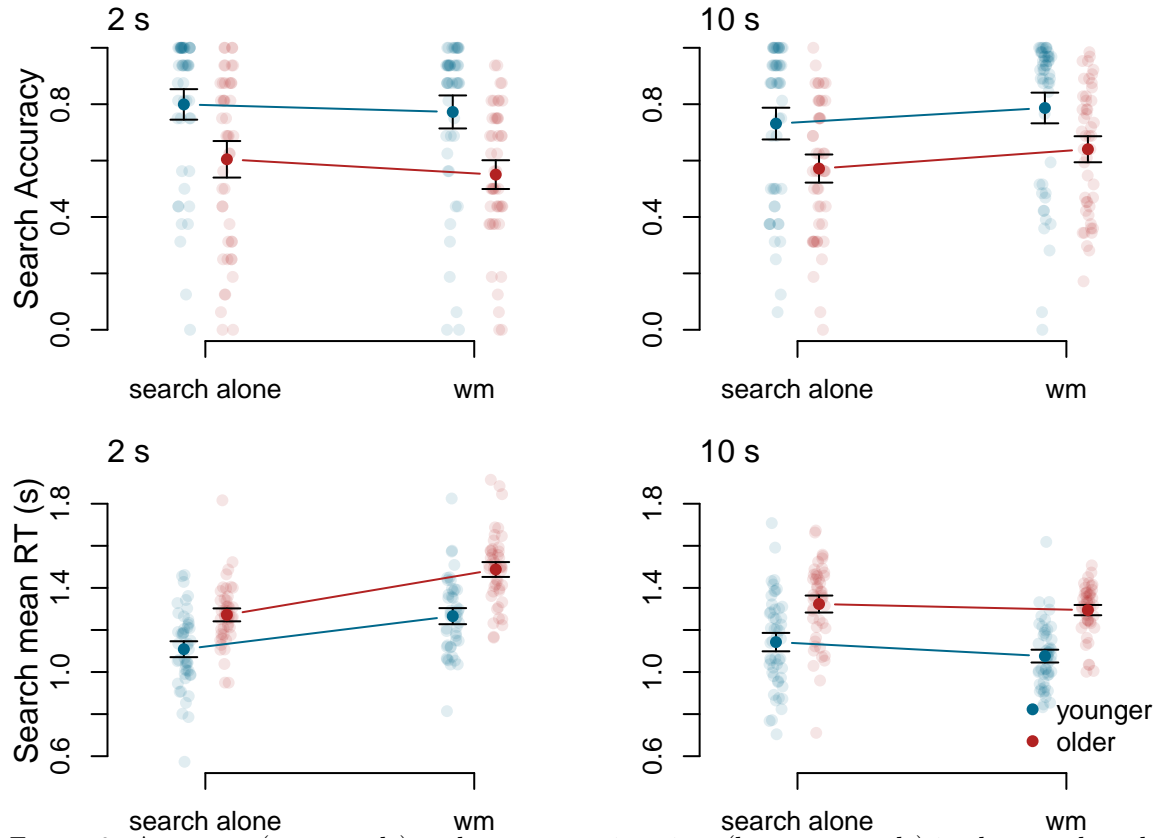
Figure 2. Accuracy (top panels) and mean reaction time (bottom panels) in the search task by interval length and age group. Participants did the search task by itself and during the working memory task interval.

Search reaction time

Reaction time in the search task was analyzed with a linear mixed effects model with the same fixed and random effects as the analysis of accuracy above. As we are modeling reaction time in seconds we used a Cauchy(1, 1) prior on the intercept term (suggesting we expect average RT to be around 1 s, with a high degree of uncertainty) and Cauchy(0, 1) priors on the other fixed effects. For the standard deviation of random effects we used a half-Cauchy(0, 1) prior and the same LKJ(1) prior on the correlation matrix. Estimates of fixed effects are presented in Table 3.

There is a main effect of interval, with faster response times in the 10 s condition relative to the 2 s condition. Older adults were also slower to respond than younger adults. There was also a main effect of section with faster responses for search alone versus search alongside the working memory task. Finally, interval and section interact as reaction times were shorter for search alone (1.182 [1.144, 1.218]) relative to search + working memory (1.361 [1.327, 1.395]) in the 2 s condition (difference: -0.179 [-0.222, -0.140]). In the 10 s condition reaction times did not significantly differ between search alone (1.229 [1.193, 1.267]) and search with the working memory task (1.184 [1.156, 1.213]; difference: -0.002

Table 2

Results of generalized linear mixed effects model for search accuracy. Posterior mean, 95% highest density interval, percentage of posterior samples greater than zero, and Bayes factors in favor of the null and alternative.

	Mean	95% HDI Lower	95% HDI Upper	Perc > 0	B_{01}	B_{10}
Intercept	1.23	0.98	1.49	100.00	-	-
Interval	0.01	-0.25	0.24	52.59	3.35	0.30
Group	-0.72	-0.97	-0.47	0.00	<1e-4	>1e+4
Section	-0.02	-0.13	0.09	34.87	6.87	0.15
Interval × Group	0.07	-0.17	0.33	71.39	2.86	0.35
Interval × Section	-0.20	-0.30	-0.09	0.03	0.01	72.93
Group × Section	0.02	-0.09	0.13	63.12	7.24	0.14
Interval × Group × Section	0.00	-0.10	0.11	53.37	7.77	0.13

Table 3

Results of linear mixed effects model for search reaction time (in seconds). Posterior mean, 95% highest density interval, percentage of posterior samples greater than zero, and Bayes factors in favor of the null and alternative.

	Mean	95% HDI Lower	95% HDI Upper	Perc > 0	B_{01}	B_{10}
Intercept	1.24	1.22	1.26	100.00	-	-
Interval	-0.03	-0.05	-0.01	0.07	0.73	1.38
Group	0.10	0.08	0.12	100.00	<1e-4	>1e+4
Section	-0.03	-0.05	-0.02	0.00	<1e-4	>1e+4
Interval × Group	0.01	-0.02	0.03	68.45	76.41	0.01
Interval × Section	0.06	0.04	0.07	100.00	<1e-4	>1e+4
Group × Section	-0.01	-0.02	0.00	7.64	46.78	0.02
Interval × Group × Section	0.00	-0.01	0.02	62.52	120.64	0.01

[-0.048, 0.047]).

The search data as a whole suggest that introducing the working memory task slowed participants down and reduced accuracy in the 2 s condition. In the 10 s condition participants did not slow down and improved in accuracy when the search task was performed alongside the working memory task. This is possibly due to differences in the amount of practice/experience with the search task between the two conditions, as participants in the 10 s condition completed 4 times as many search problems during the working memory task.

Low search performers and WM performance

The top panels of Figure 2 shows many participants with low accuracy in the search task. This potentially complicates our interpretation of the search task as being a distraction as low accuracy may reflect disengagement from this task in favor of focusing on maintaining information in working memory (Barrouillet, Bernardin, Portrat, Vergauwe, & Camos, 2007).

To assess possible biases associated with these low performing participants we conducted additional analyses of the working memory data from conditions with distraction. The same model as reported in the main manuscript (omitting fixed effects involving distraction) was fit to data from all participants and again to participants with 70% or higher accuracy in the search task when performed alongside working memory. This omitted 32/46 and 25/47 older participants and 14/47 and 14/50 younger participants in the 2 s and 10 s interval conditions, respectively. Figure 3 plots the fixed effects estimates from these two models. The two are clearly related and the only deviation (towards the top right of the figure) is the intercept term, as excluding low search performers leads to higher overall working memory accuracy. None of the remaining coefficients clearly differ, suggesting that low performers are not distorting our conclusions regarding working memory performance.
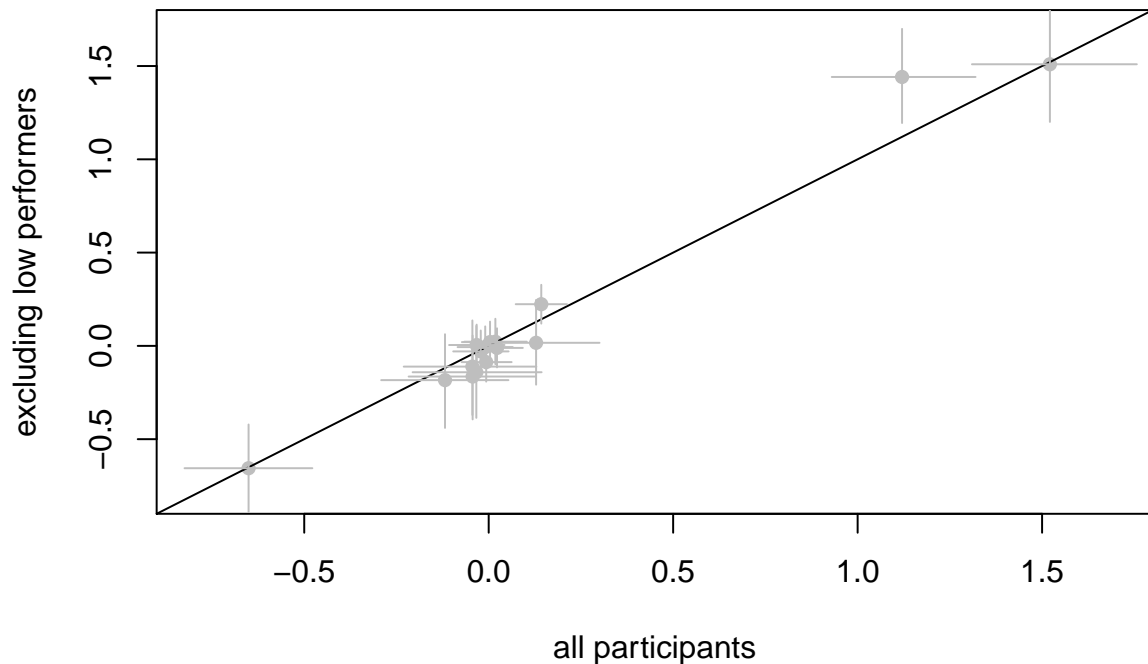

Figure 3. Fixed effects estimates and 95% credible intervals for analysis of working memory data included vs. excluding low search performers.

WM performance for new pairs by trial context

Bartsch and Shepherdson (2020) found that presenting pairs that matched previously learned pairs during a working memory task led to better retention of new pairs (as assessed by recognition performance) relative to when all pairs were new, whereas including rearranged pairs led to poorer performance for new pairs. Bartsch and Shepherdson (2020) suggest that when participants are able to rely on reliable information from long-term memory they are able to allocate more resources to storing new information in working memory.

Unfortunately our design had only two contexts—one where all pairs on that trial were new and the more frequent trial type with one match pair, two recombined pairs, and one new pair—in which new pairs appeared, which means we are unable to assess whether the same pattern of results holds for our cued recall task. Nevertheless, we assessed whether recall of new pairs was sensitive to the trial context. As younger participants only showed facilitation for pairs made up of previously learned elements, we may expect this group to better recall new pairs when the trial context allowed the use of LTM representations relative to when all pairs were new. For older adults, who exhibited proactive interference for recombined pairs and facilitation for match pairs, what to expect is less clear.
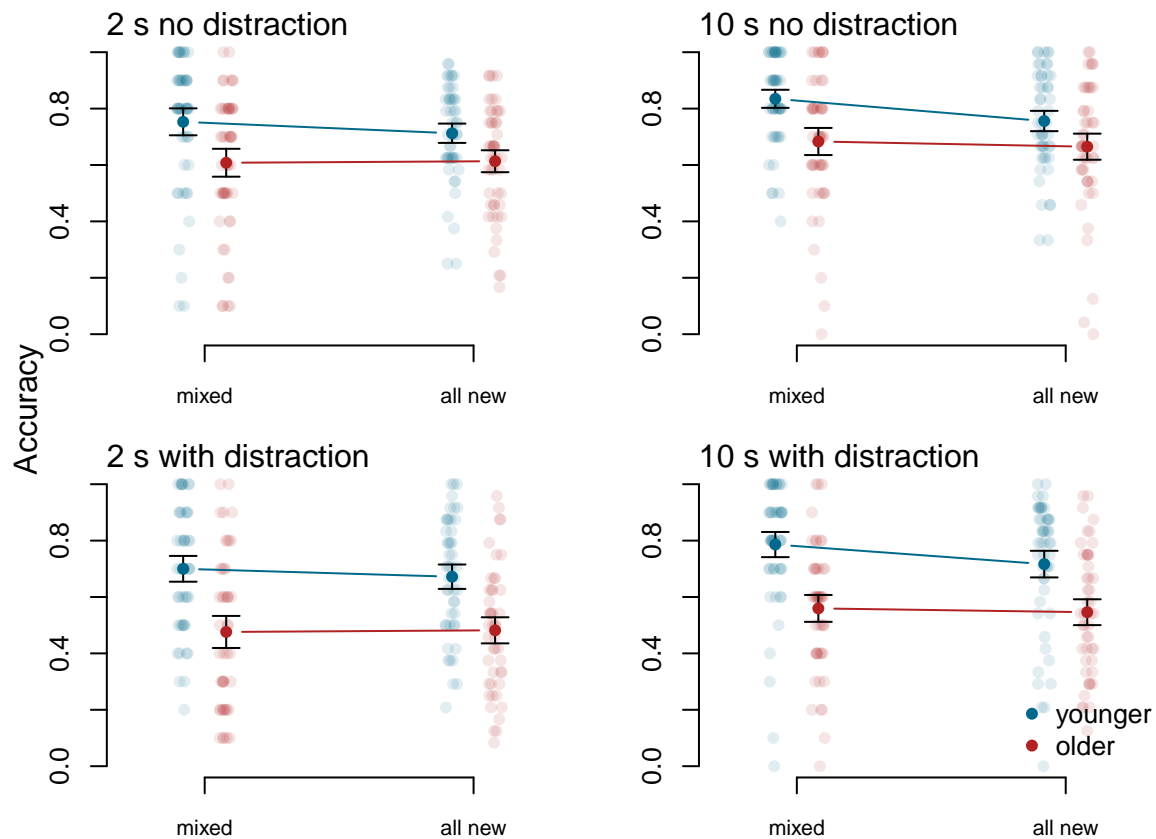

Figure 4. Accuracy in the working memory task for new pairs by condition and pair context. Error bars are within-subjects standard errors.

Taking recall accuracy for pairs that were new we categorized the trial context as

Table 4

Results of generalized linear mixed effects model on accuracy for new-new pairs in the working memory task taking into account trial context. Posterior mean, 95% highest density interval, percentage of posterior samples greater than zero, and Bayes factors in favor of the null and alternative.

| | | 95% HDI | | | | |
	Mean	Lower	Upper	Perc > 0	B_{01}	B_{10}
Intercept	0.87	0.73	1.01	100.00	-	-
All New	-0.11	-0.17	-0.06	0.01	<1e-4	>1e+4
Distraction	0.22	0.11	0.34	100.00	0.01	147.51
Interval	0.20	0.08	0.31	99.93	0.04	24.24
Group	-0.45	-0.57	-0.34	0.00	<1e-4	>1e+4
All New × Distraction	-0.02	-0.06	0.03	24.24	17.91	0.06
All New × Interval	-0.05	-0.10	-0.01	1.23	1.64	0.61
Distraction × Interval	0.01	-0.11	0.12	57.22	9.36	0.11
All New × Group	0.08	0.03	0.13	99.95	0.12	8.12
Distraction × Group	0.09	-0.03	0.20	93.36	3.26	0.31
Interval × Group	-0.02	-0.14	0.09	34.08	8.62	0.12
All New × Distraction × Interval	0.00	-0.05	0.05	48.91	23.45	0.04
All New × Distraction × Group	0.01	-0.04	0.05	63.16	22.01	0.05
All New × Interval × Group	0.03	-0.02	0.07	84.99	13.57	0.07
Distraction × Interval × Group	0.00	-0.12	0.11	48.10	9.55	0.10
All New × Distraction × Interval × Group	-0.01	-0.05	0.04	40.49	22.33	0.04

either all new, where all the pairs on that particular trial were new-new, or mixed (see Figure 4). We fit a generalized mixed effects model with the fixed effects of this trial context factor (all new coded 1, mixed coded -1), distraction, interval, and group (all coded as in the main analyses) as well as their interactions. In addition there was a random participant intercept and effect of trial context as well as a random item intercept. Posterior summaries are given in Table 4.

The trial context influences accuracy for new pairs, with higher accuracy on mixed trials relative to all new ($B_{10} = 14280$). There is an interaction between group and trial context ($B_{10} = 8.12$). As shown in Figure 4, the difference in accuracy between the mixed and all new trial types is more pronounced for the younger group. This is supported by Bayes factor contrasts that provide strong evidence for an effect in younger adults ($B_{10} = 3.66e+13$) but evidence in favor of the null for the older group ($B_{01} = 36.37$).

The 95% HDI for the interaction between trial context and interval excludes zero, however the Bayes factor does not provide compelling evidence for or against the null ($B_{01} = 1.64$). For the remaining interactions the null was favored by over 9-to-1.

Relationship between learning and WM performance

To assess the relationship between participants' final level of learning and performance for different item types in the working memory task we estimated a multivariate model with `brms`. In the main analyses, separate generalized mixed effects model were fit to accuracy from the two phases of the experiment with separate random participant effects. In the multivariate model these random participant effects are allowed to correlate across response variables. This involves no aggregation of the data; the data are modeled as they are (zeros and ones) and the correlations are estimated at the "latent" level (see Rouder & Haaf, 2019).

For the learning outcome the model was simplified to include a fixed effect of age group and a random participant intercept to capture between-participant variability in performance during part 1 of the experiment. The model for the working memory data included fixed effects of item type (coded as in the main analysis), age group, and their interaction plus a random intercept and random effect for item type. The random item type effect captures individual differences in proactive facilitation (through the contrast of match with new) and proactive interference (contrasts of old-new and mismatch with new).

Figure 5 presents estimates of the correlations and their uncertainty. The only correlations that do not include zero in their highest density intervals are between overall performance in the working memory task (WM Overall) and the final level of learning ($r = 0.417$ [0.314, 0.524]; $B_{10} > 10000$) as well as the correlation between proactive facilitation for match items in the working memory task (WM PF = match vs. other) with level of learning ($r = 0.388$ [0.187, 0.600]; $B_{10} = 58.64$), with higher levels of learning associated with greater facilitation. Estimates of the remaining correlations are associated with a high degree of uncertainty and Bayes factors that do not strongly favor the null or alternative (see Figure 5). However, in the case of the correlations between proactive interference (WM PI) and level of learning and between overall WM performance with estimates of proactive facilitation or interference (see bottom two rows in Figure 5) we can rule out very strong relationships between the other random effects.
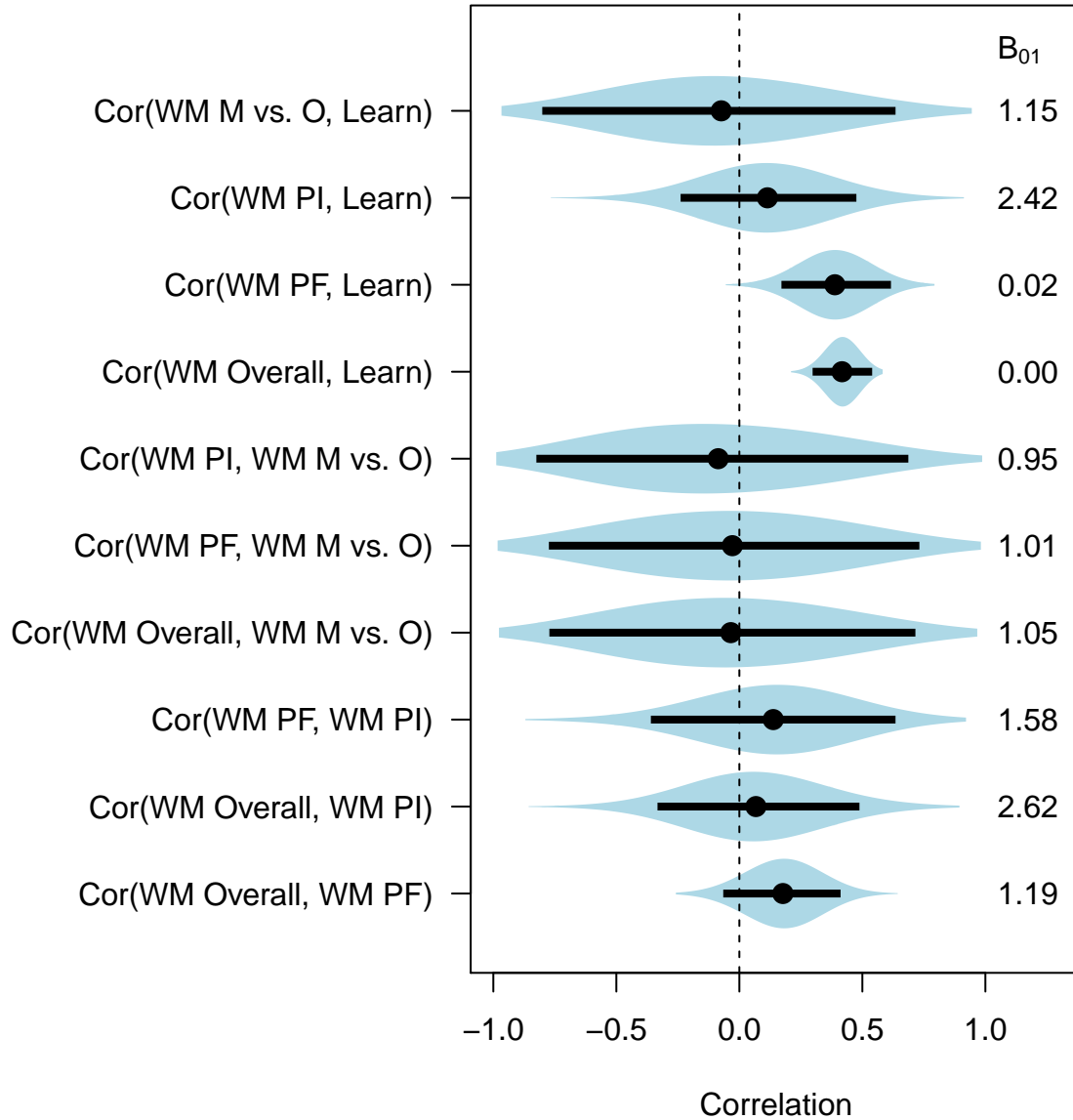
Figure 5. Correlations between final learning and WM performance estimated via a multivariate mixed effects model. Bayes factors for the null hypothesis are given next to the associated violin plot ($B_{01} < 1$ provides support for the alternative) Learn = final learning intercept, WM overall = Intercept, WM PF = match vs. other contrast, WM PI = new-new vs. mis-match/old-new contrast, WM M vs. O = mis-match vs. old-new contrast

Further categorizing responses

The main analysis just categorized responses as correct or incorrect. However, there are different, potentially informative, types of incorrect response. To explore this further recalled words in the working memory task were categorized as follows:

- `correct`: the recalled word was that presented with the cue image during the study phase of the trial
- `omission`: the participant did not attempt to recall a word
- `within`: the recalled word was presented with another image on this working memory trial (a within list error or transposition)
- `previous`: the recalled word was presented during the previous working memory trial
- `ltm`: the recalled word did not fall into any of the above categories and was one of the 30 words that this participant learned during the first phase
- `extra`: the recalled word does not fall into any of the other categories

This new 6 category outcome was analyzed with a categorical mixed effects model via `brms`. The model included fixed effects of interval, distraction, age group, and item type as well as random participant intercept and slope (item type) terms and a random effect of to-be-recalled word. In the categorical model the correct response category was selected as the reference category and coefficients for the remaining categories are given relative to the reference category:

$$\beta_k = \log\left(\frac{p(\text{response} = k)}{p(\text{response} = \text{correct})}\right).$$

In order to simplify the complex results of this model we focus on coefficients for which the HDI firmly excludes zero. Table 5 presents coefficients where the HDI excludes absolute parameter values of less than 0.1 (in the model's log ratio units). To summarize, as older adults make more errors, there are age differences for all response types. However, there are no interactions between age and item type, with the exception of the Group × New-new vs. mis/old interaction for the within response category [explain...]. Further, the factors of interval and distraction appear to play very little role in modulating response categories.

Figure 6 presents the estimated response probabilities by age group and item type. In Figure 7 we have rescaled the probabilities to account for the fact that older adults make more errors. In this figure the values reflect the estimated proportion of errors that fall into each response category. Crucially, it is clear that, once we account for differences in error rates, older adults do not exhibit higher rates of ltm intrusions.

This is also shown in Figure 8 which plots contrasts of different pair types to the new-new baseline for this rescaled measure (p(response) as a proportion of errors). Focusing on the differences between younger and older participants (the gray densities), perhaps the clearest difference in response types is in the within category. Specifically, it appears that, relative to new-new, younger adults were less likely to make within list errors (i.e., transpositions) for mis-match and old-new pairs. This is not seen in the older group.

Table 5

Results of categorical model on working memory responses. Posterior mean, 95% highest density interval, and percentage of posterior samples greater than zero. Note we have only selected coefficients where the HDI excludes ± 0.1.

|	Mean	95% HDI Lower	95% HDI Upper	Perc > 0
extra: Intercept	-2.87	-3.03	-2.70	0.00
ltm: Intercept	-4.47	-4.79	-4.18	0.00
omission: Intercept	-3.66	-3.95	-3.40	0.00
previous: Intercept	-5.08	-5.53	-4.69	0.00
within: Intercept	-2.46	-2.62	-2.31	0.00
extra: Group	0.33	0.19	0.45	100.00
extra: Match vs. other	-1.12	-1.36	-0.90	0.00
extra: New-new vs. mis/old	0.27	0.16	0.37	100.00
ltm: Group	0.49	0.29	0.73	100.00
ltm: Match vs. other	-1.58	-2.24	-1.01	0.00
ltm: New-new vs. mis/old	-1.23	-1.50	-0.94	0.00
omission: Group	0.87	0.65	1.08	100.00
omission: Match vs. other	-1.50	-1.92	-1.10	0.00
previous: Group	0.80	0.48	1.18	100.00
previous: Match vs. other	-2.06	-3.18	-1.09	0.00
within: Interval	-0.34	-0.48	-0.22	0.00
within: Group	0.63	0.50	0.76	100.00
within: Match vs. other	-1.07	-1.30	-0.85	0.00
within: Group × New-new vs. mis/old	-0.27	-0.35	-0.19	0.00
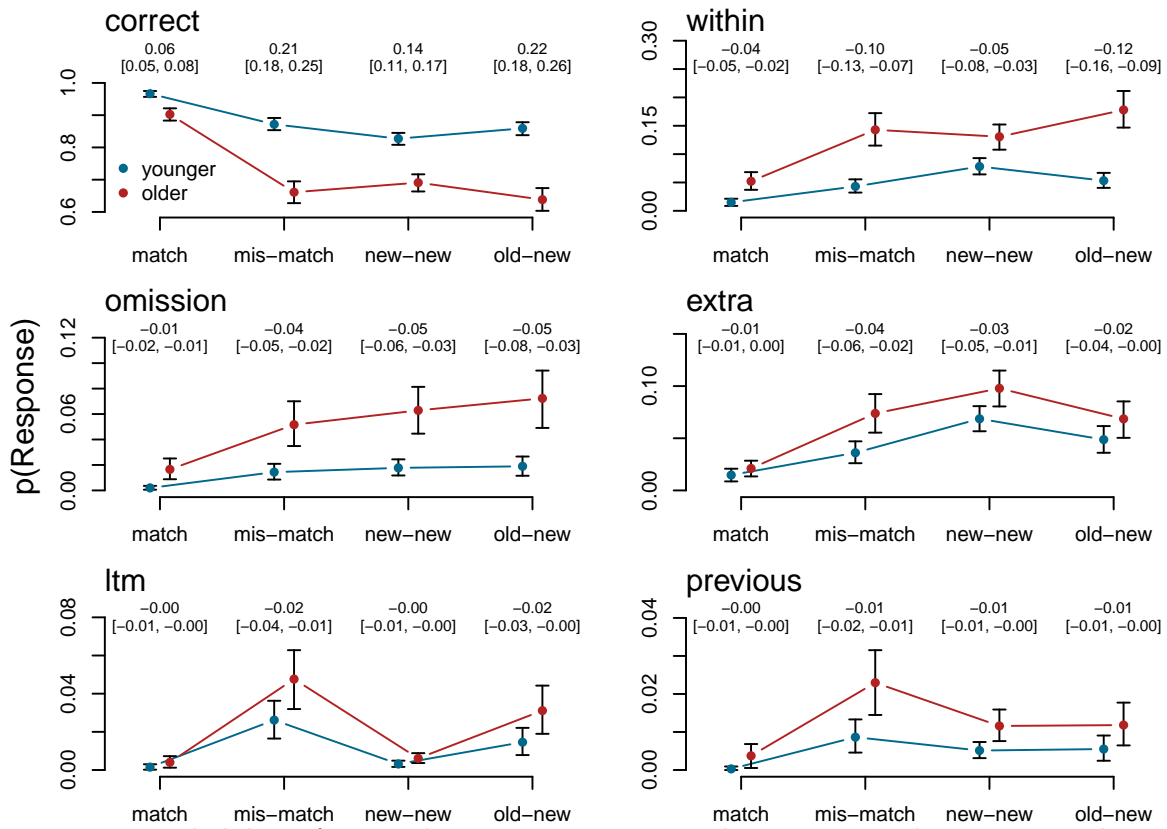
Figure 6. Probability of particular response categories by item type, distraction, and age group. Posterior mean and 95% highest density interval. Text above points gives posterior mean and HDI of group difference for that item type.
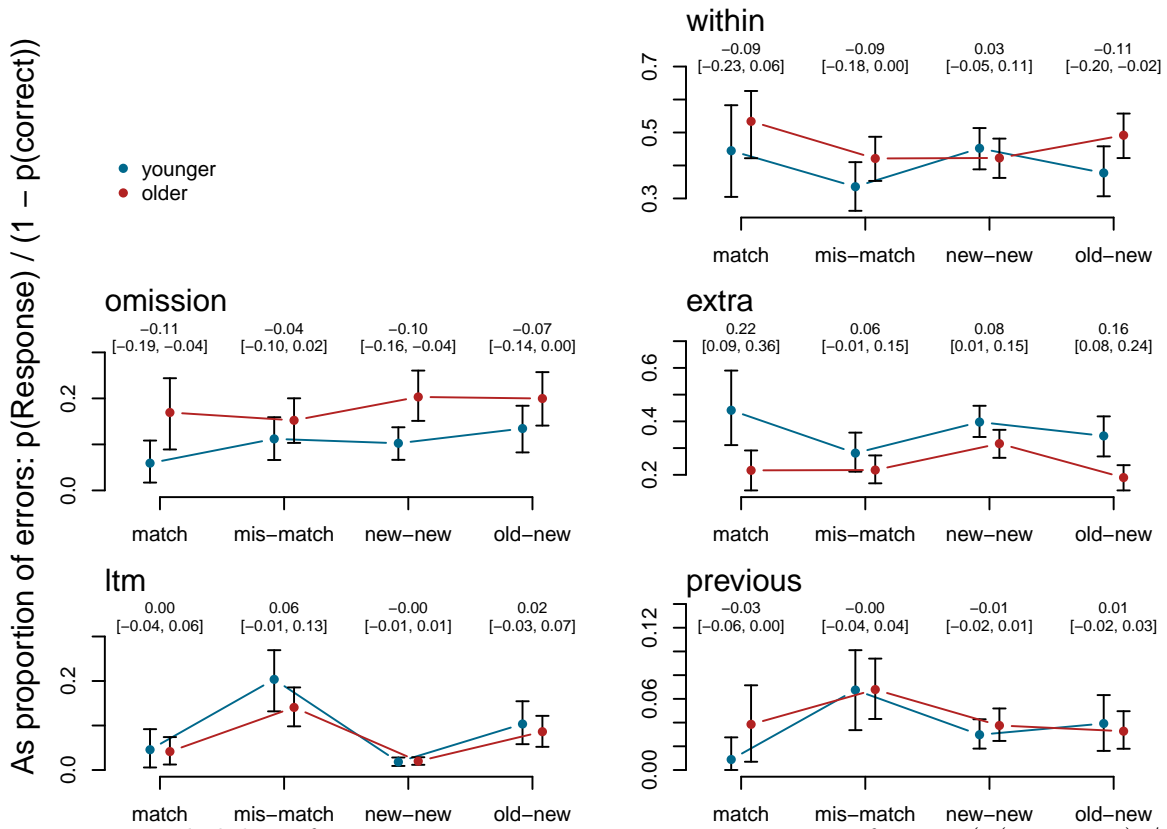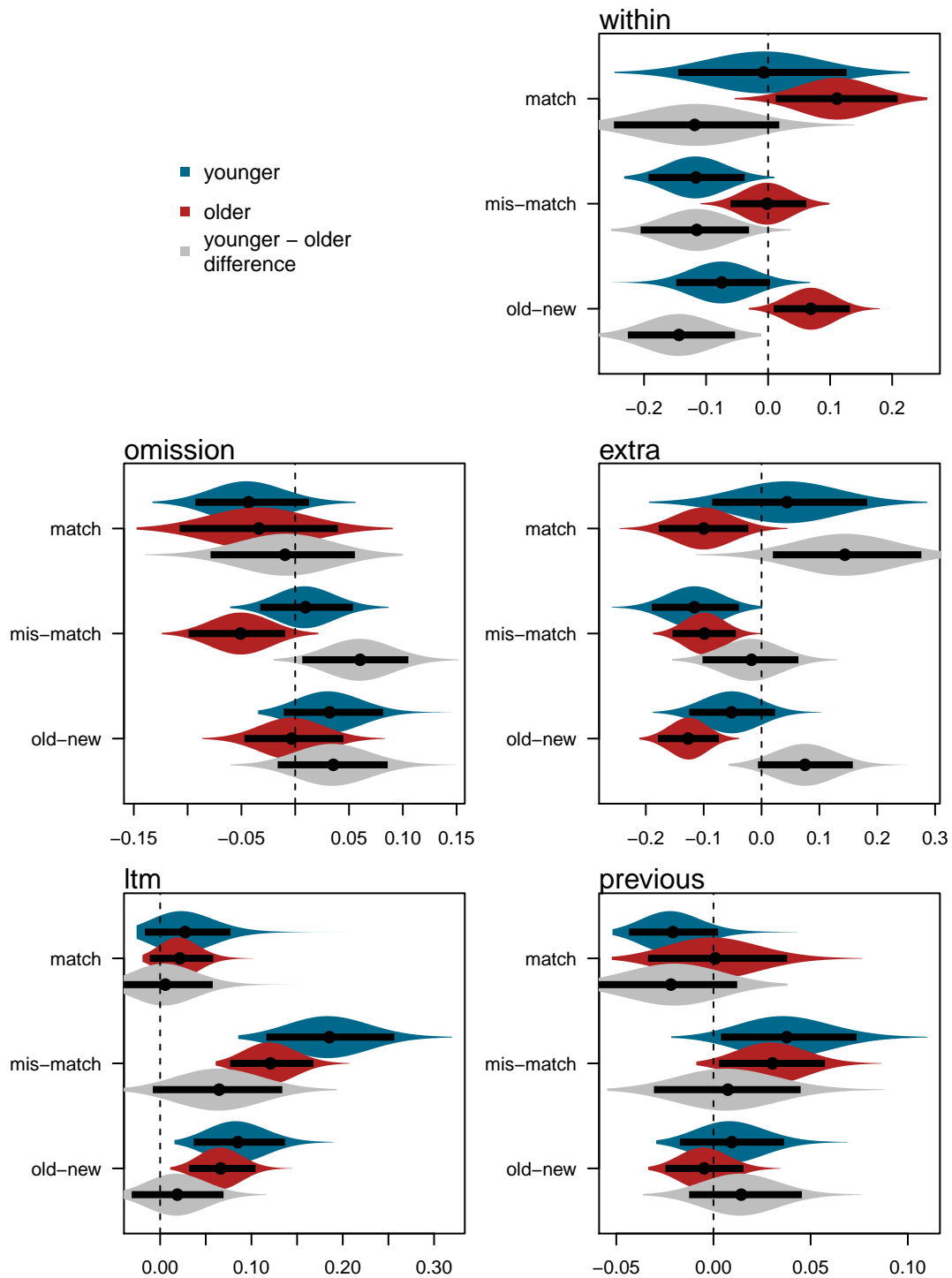
Figure 7. Probability of error response categories as a proportion of errors (p(Response) / [1 - p(correct)]). Posterior mean and 95% highest density interval. Text above points gives posterior mean and HDI of group difference for that item type.

Figure 8. Posterior contrasts (density, posterior mean and 95% highest density intervals) for the different working memory response categories (expressed as a proportion of errors)

References

Aust, F., & Barth, M. (2020). *papaja: Create APA manuscripts with R Markdown.* Retrieved from https://github.com/crsh/papaja

Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., & Camos, V. (2007). Time and cognitive load in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(3), 570–585.

Bartsch, L. M., & Shepherdson, P. (2020). Freeing capacity in wm through the use of ltm representations. https://doi.org/10.31234/osf.io/hj4rv

Rouder, J. N., & Haaf, J. M. (2019). A psychometrics of individual differences in experimental tasks. *Psychonomic Bulletin & Review*, *26*(2), 452–467.