### Model fitting procedure cut outs

The pattern of results described above vary when split by model type. Cases where the LTM model were the best fit have the strongest evidence against the 2nd fit best model compared to the cases where the other types of models were best fit (mean difference of 5.64, SEM=0.555). The explicit bias and RL models have weak evidence against the second best-fit model at mean BIC difference of 1.58 (SEM= 0.404) and 1.91 (SEM=0.94) respectively. The mean difference in BIC for participants that fit the meta-learning integrated model best was 3.30 (SEM =0.752) (Figure 10). These results suggest that the LTM model overall has the highest likelihood of capturing learning behavior in the RLWM task, even while it is not the model with the least number of parameters.

### Capturing Individual differences cut out

The results of this model fitting procedure yielded somewhat different results than the original study. As discussed above, we did not find that one model outperformed the others reliably. Rather, (Figure 11). This was true even when, as in the case of integrated models, they effectively included the basic models as particular cases. In principle, this could be because the BIC procedure penalized more complex models for their increased numbers of parameters.

### Chart, box and whisker chart Description automatically generated

Figure 10: BIC difference for the first and second fit models grouped by model that was the best fit. Red lines show the 6 to10 (strong evidence), 2 to 6 (positive evidence) and 0 to 2 weak evidence ranges. The LTM model has the highest probability of being the preferred model over other models in cases where LTM was the best fit model.

### Parameter analysis

Assessing the parameters’ effects on learning outcomes is difficult as the relationships among parameters is complex and determined by the specific formulae in ACT-R. However, some parameters might have linear and interacting effects on learning outcomes. To reiterate, the model learning was influenced by a maximum of 6 parameters ranging on 5 values (Table 1). The pure RL model had 2 parameters, learning rate (α) and selection noise (T); the pure LTM model had 3 parameters, activation noise (s), memory decay rate (d) and attentional spreading activation weight (W). The two integrated models, meta-learning and explicit biased model utilized the 5 parameters above with an additional bias parameter (β)in the explicit biased model that defined the proportion of RL vs LTM to use throughout the learning and test phases. Bias was measured for the meta-learning model by determining the proportion of RL used by the meta-learner throughout the learning phase.

A wide range of parameter values were explored to capture as much individual variability as possible within the computational constraints of running simulations of all possible parameter combinations. After model fitting, a total of 35 parameter-value sets across all four models, out of 15,865 possible sets, described all 83 participants. We expected, for most parameters, that the fit values would fall normally around the recommended ACT-R values. But, we found, regardless of model type, that the distributions of fit parameter values were on either a positive or negative skew across the range of sampled parameters.

Timeline

Description automatically generated with medium confidenceAcross all subjects that fit LTM strategy containing models, on the memory decay parameter had a negative skew:75% fell on the highest level in the range explored and the next 20% were characterized by the second highest value of decay rate. There is not sufficient variability in these data to estimate the isolated effects of the decay rate parameter on learning outcomes, but it is weakly correlated with learning rate, accuracy at end of learning, accuracy at testing and the difference in learning between the set-sizes (Figure 13). It is however, only moderately negatively correlated with rate of forgetting for both set-size 3 (Spearman’s r = -0.3) and set-size 6 (spearman’s -0.26). Subject fit values for the retrieval noise parameter on the other hand, were, relatively, more uniformly sampled from the range of values explored with a less pronounced negative skew (only 35% of fit subjects had the highest decay rate). Higher retrieval noise was related to higher accuracy during test for set-size 6 (spearman’s r = 0.41) and, to a lesser degree to accuracy at test for set-size 3 (spearman’s r = 0.28) which was unexpected. The parameter did not account for differences in learning rate and forgetting after break but was related to learning differences between the set-sizes (spearman’s r = 0.32). This is perhaps the first bit of evidence that suggests that single parameters in isolation may not have large driving effects on learning outcomes; learning behavior might be better explained by the combined effect of all or a majority of the parameters. For instance, in integrated models, higher levels of noise in the LTM portion of the model shifts preference to the RL portion of the model when lower levels of noise and higher levels of RL learning rate occur. Relationships like these are difficult to capture in linear models. The bias in the meta-learning model is explored in detail in the next section.

Figure 13: Spearman correlations of individual parameters and learning features divided by set-size. Column s3 for set-size 3, column s3s6 for the difference between the set-sizes and s6 for set-size 6.The parameters compared are **alpha**:RL learning rate, **egs**: RL selection noise, **bll**: memory decay rate, **imag:** attentional spreading activation weight, **ans:** activation noise and **bias:** proportion of RL used by integrated models.

The last of the LTM parameters, the spreading-activation parameter attempts to capture individual differences in attention and working memory (Lovett et al., 2000). The default level of 1 (in ACT-R 7x) and above injected too much spreading activation to capture individual differences so a range between 0.1 and 0.5 (Table 1) was used. Subject-fit values were positively skewed, where 50% of subjects fell onto the lowest value 0.1, followed closely by 27.1% for the second lowest parameter value. Increasing values of spreading activation trended in the same direction with increasing rates of forgetting (learn-test difference) during test (spearman’s r= 0.41, set-size 3 and 0.37, set-size 6), test accuracy (spearman’s r = 0.43 and.41, set-size 3 and 6 respectively), and the sets-size 3 vs 6 learning difference (spearman’s r = 0.31). It is not related with learning rate for both set-sizes but related to learning accuracy only in set-size 6 (spearman’s r= 0.28, 0.05 for set-size 3).

Regarding the two RL parameters RL noise and learning rate alpha, both parameters are slightly skewed but in opposing directions and also seem to be more strongly predictive of learning outcomes than the LTM parameters. It should be noted that the interpretation of these correlations is limited due to the small size of the samples in all but the LTM only model. The noise parameter is positively skewed (40% of fit subjects falling on the lowest values of 0.1) and trends in the opposite direction to learning accuracy for both set-sizes (spearman -0.4, -0.25 set-size 6) but is related to learning rate and test accuracy only in set-size 6 (spearman -0.25, -0.28). Lastly, the learning rate alpha parameter trends strongly in the same direction with set-size 6 learning accuracy (spearman 0.72) and testing accuracy (spearman’s r = 0.7) and also trends positively with the learning difference between the set-sizes (r = 0.48). This suggests that perhaps learning rate alpha could account for better learning in both set-sizes, is the strongest predictor among the parameters, and affects both similarly. Learning rate alpha is relatively weakly related to set-size 3 learning accuracy (0.24) and learning rate (0.19). It is not related to testing accuracy in set-size 3 (0.05) and in test forgetting for both set-sizes (-0.06 and -0.02 set-size 6).

**Bias parameter**

The 6th parameter, learning bias, was pre-defined in the explicit bias integrated model (figure 5) at 20%, 40%, 60% and 80% probability of RL use but it was measured in the meta-learning model. We found that 9 out of 10 subjects that fit this model used RL only 20% of the time. The image is slightly more complicated when considering the meta-learning model. Recall that the meta-learning model used the RL-based, production utility learning to select either the RL or LTM learning models. This RL meta-learner selected different proportions of RL and LTM for the two set-sizes. This, as predicted was influenced by the relative success of the sub-models as determined by the current value of their parameters (Figure 14). For example, looking at the model data only, significantly higher levels of RL were selected to perform the task as the value of the learning rate alpha increased; however, this increase was associated with set-size 6 blocks only and a decrease in the proportion of RL used for the set-size 3 block was observed. Interestingly, the changes in the three LTM parameters exhibited similar trends in the proportion of RL used in both set-size blocks. An increase in the noise and memory decay parameters resulted in more use of the RL subsystem but this shift was more pronounced in the set-size 6 block. Similarly, an increase in the spreading-activation parameter, which favors the LTM model, led to a related decrease in RL use, again, at a more pronounced level for the set-size 6 block. Only 9 of our 83 participants fit this model best Chart, calendar

Description automatically generatedbut we observed that the estimated bias towards RL in the set-size 6 increased significantly, to around 50% of the time, compared to set-size 3 (set-size 3: M = 0.3; SEM= 0.05; set-size 6:M = 0.47, SEM = 0.15).

Figure 13: Model data only showing the effect of changes of parameter values on the proportion of RL used for the set-size 3 and set-size 6 blocks on the proportion of RL sub-model used in the meta-learning RL integrated model. Parameters: **alpha**:RL learning rate, **egs**: RL selection noise, **bll**: memory decay rate, **imag:** attentional spreading activation weight, **ans:** activation noise.