Supplementary material

In our pre-registered proposal of the current study, we had indicated that we would submit our data to parametric tests (reported in the main manuscript). We had also planned to analyse the data using a drift diffusion model to gain deeper insights into how extended learning and memory differences affect the self-prioritisation effect. One of our goals was to test if the differences between associations and conditions can be attributed to perceptual or cognitive processes. Here, we report our attempts to implement such a model.

A drift diffusion model (DDM) makes superior use of the data available in an experiment where accuracy and reaction times are used as measures of performance in a 2AFC task (Voss, Nagler, & Lerche, 2013). Instead of analysing only means or medians of performance (RTs, accuracy), DDMs use the entire reaction time distributions and also take into account both correct and incorrect responses. A range of parameters can be estimated from these distributions that can potentially provide insight into which cognitive processes are influenced by the experimental manipulation.

We had planned to fit a ‘hierarchical drift diffusion’ model (HDDM; Wiecki, Sofer, & Frank, 2013) to our data. As the usage of HDDMs is relatively new in experiments investigating the effect of self-prioritisation, we wanted to follow 2 complementary approaches: a bottom-up data-driven exploratory approach and a top-down hypothesis-driven approach. For both approaches we used the HDDM toolbox (Wiecki et al., 2013) to estimate the drift rate v (average slope of the diffusion process), the threshold separation a, the starting point z and the duration of the non-decisional process t_0. Based on the design of our experiment and to be able to estimate a possible bias in the responses we utilised the ‘stimulus coding’ method. In this method, the HDDM fits the data using the two response decisions (match and non-match) as the opposing decision boundaries. Here, the interpretation of a shift in starting point (z, or bias) towards one of the boundaries (e.g. non-match) is more meaningful and has a more straightforward interpretation, than in the alternative ‘accuracy coding’ method in HDDM, where the opposing boundaries are correct or incorrect answers.

Top-down approach

For the top-down approach, we wanted to test two hypotheses. Firstly, our hypothesis that the differences in the response patterns are driven by memory suggests a difference in the threshold separations (*a*) between the different shape-identity associations. We expect the threshold separation in the self-association condition to be higher than for the other associations. Although this on its own would slow down responses, it would also make the responses less susceptible to noise during information accumulation leading to higher accuracy for the self-related than for other related associations (Sui, He, & Humphreys, 2012).A similar difference is also expected for the condition where memory is manipulated through exposure, with a larger threshold separation when exposure during learning was high, compared to medium and low. We would also expect a reduced difference between self- and other-related associations in the extended learning conditions. Secondly, to account for the reported differences in reaction times (and to overcome the effect of the proposed differences in threshold separation) the self-prioritisation effects could be due to a decisional bias (indicated by differences in the starting point *z*) or differences in (perceptual) information accumulation (indicated by differences in the drift rate *v)*. That is, RTs could be slower for the non-self related associations because of a starting point closer to the lower boundary (bias towards responding non-match) or due to a shallower drift rate (slower information accumulation). We additionally restricted the number of models to those that estimate *z* for each of the different associations (you, friend and stranger). This restriction is motivated by a finding of differences in bias across association types that was recently reported in a study that applied HDDM to a shape-detection-task under continuous flash suppression (Macrae, Visokomogilski, Golubickis, Cunningham, & Sahraie, in press). Based on these hypotheses, we limited the tested models to those that estimate the parameters *z, a* and *v* for each of the different shape-label conditions.

In such models, the tested variable (or variables) would be allowed to vary across the three identity conditions (self, friend, stranger), and a single value would be estimated across all three conditions for the other variables. That is, in a model that includes parameters *a* and *z* these two parameters would be estimated separately for the three identity conditions, but *v* and t_0 would be estimated over all identities. The model with the best fit to the data is then used to analyse differences between the different identities. The various models can be compared using the respective Deviance Information Criterion (DIC) values, which incorporates the models' goodness-of-fit to the data and the number of parameters included in the model. The better a model's fit to the data and the fewer the parameters it estimates, the lower its DIC value will be. A difference (reduction) in the DIC value of 2-6 points is positive evidence that one model is better than the other, 6-10 is strong evidence and more than 10 is very strong evidence (Kass & Raftery, 1995). We only considered models as superior when the DIC change exceeded 10 points. The tested models with their respective fits (DIC values) are shown in table 1.

Supplementary table 1: Tested models for each of the learning conditions following the hypothesis driven approach. Asterisks indicate non-convergence of the model based on R-hat convergence statistics.

Learning condition	standard		Shape-identity		Shape-label-identity		Shape-non-word	
Model (DIC)	*a, v, z**	(243)	*a, v, z*	(-7950)	*a, v, z*	(-8633)	*v, z*	(-3531)
	v, z	(895)	*v, z*	(-7698)	*v, z*	(-8526)	*a, v, z*	(-3466)
	a, z	(1473)	a, z	(-7591)	a, z	(-8475)	a, z	(-3232)
	z	(1484)	*z*	(-7413)	*z*	(-8211)	*z*	(-3187)

However, the model that yielded the best fit (*a, v,* z) in 3 of the four learning conditions did not converge for the standard learning condition, which we had planned to compare to the other conditions. Further, the second-best model (*v,* z) in the standard condition had a much weaker fit (DIC higher by more than 600 points). Moreover, it would have excluded one of the key parameters (*a*). Therefore, this model was not considered suitable to test our hypotheses.

Bottom-up approach

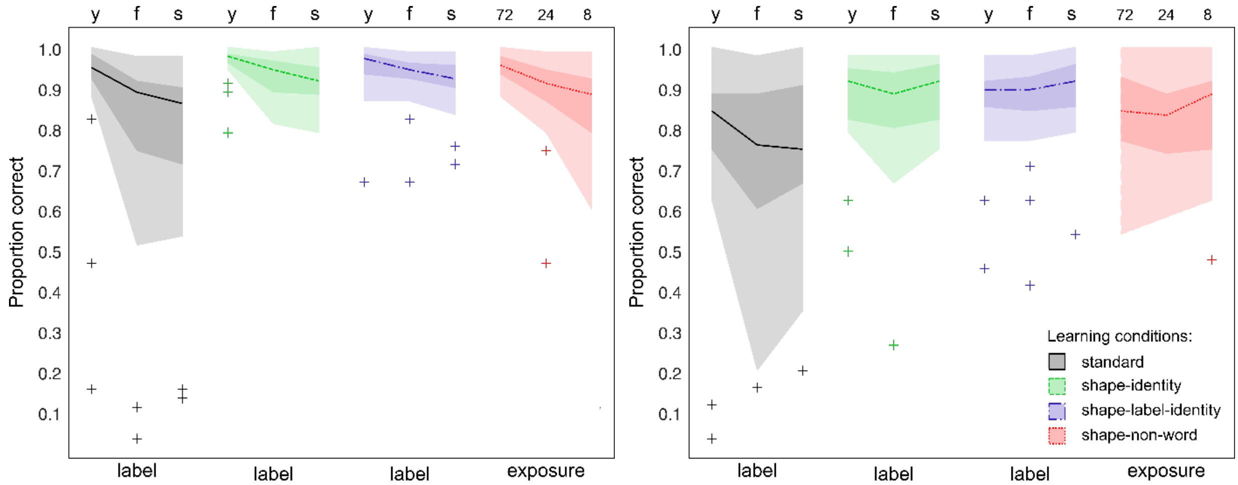
Here, the entire range of models that estimate distinct combinations of various parameters (*a, v, z* and t_0) was tested. Table 2 shows the DIC values for the 3 best fitting models per learning condition.

Supplementary table 2: Models with best fit for each of the learning conditions following the bottom-up approach. Asterisks indicate non-convergence of the model based on R-hat convergence statistics.

Learning condition	standard		Shape-identity		Shape-label-identity		Shape-non-word	
Model (DIC)	a, v, t_0*	(243)	a, v, z, t_0*	(-8258)	a, v, z, t_0	(-9201)	a, v, z, t_0*	(-3945)
	a, v, z, t_0	(287)	a, v, t_0	(-8147)	a, v, t_0	(-9334)	a, v, t_0	(-3780)
	v, z, t_0	(433)	*a, v, z*	(-7959)	a, z, t_0	(-8777)	v, z, t_0	(-3655)

To analyse the effect of the various learning conditions on the estimated parameters, we had planned to compare parameter estimates from the same model which was also among the three best fitting models in each of learning conditions (e.g. a, v, t_0 or a, v, z, t_0). However, we could not perform these comparisons, as for one or more of the learning conditions, these models did not converge ($\hat{R}$ convergence statistic) or did not include a parameter in one of the learning conditions (e.g., the model v, z, t_0 was not among the top 3 best fitting model for the two extended learning conditions). Note that convergence was tested for the shape-identity, shape-label-identity and the shape-non-word conditions only if the model converged in the standard condition, as this would be considered the baseline for testing our hypotheses.

One possible reason for a lack of convergence could be too few error trials. Accuracy in our experiment was generally high, and was even higher in the extended learning conditions compared to the standard learning conditions, with a number of participants performing error free for one or more of the associations. The accuracy data can be seen in supplementary figure 1.



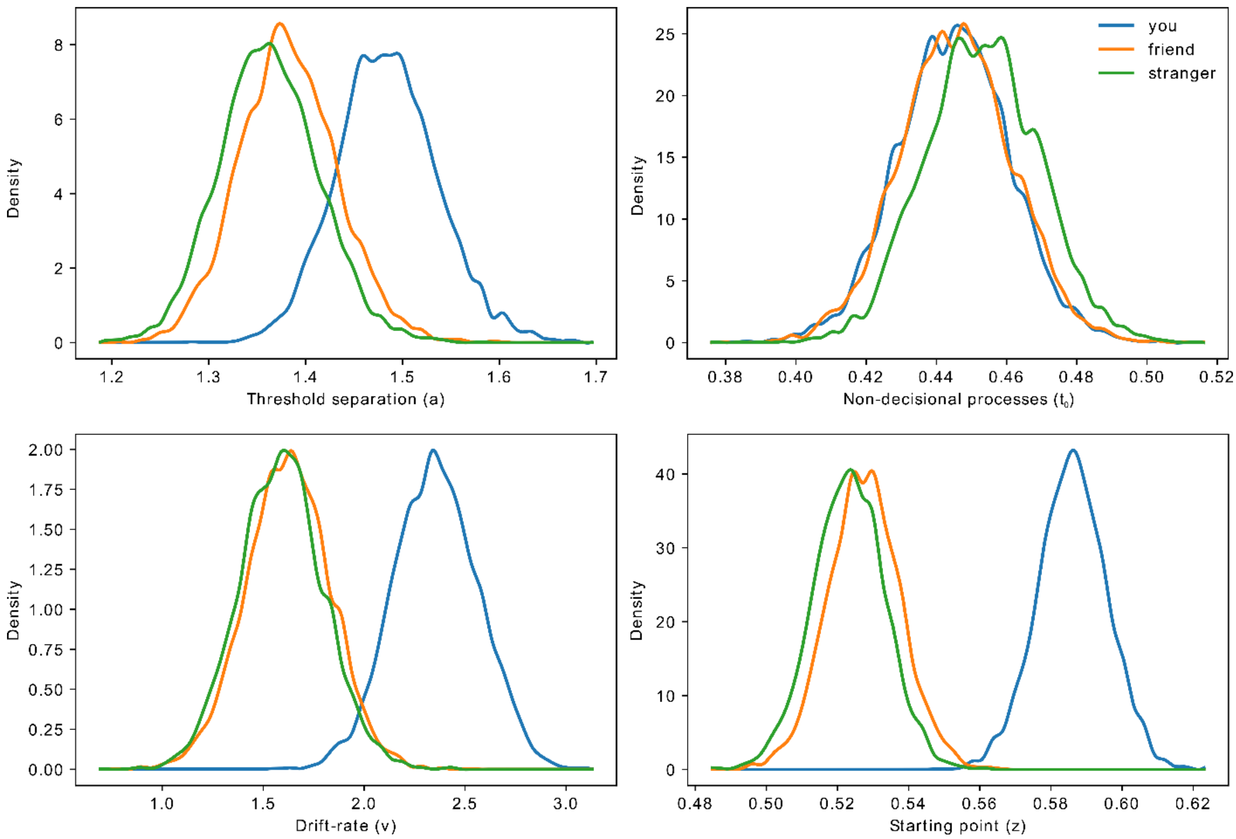
Supplementary figure 1: Accuracy data per learning condition and association* for the match-trials (left) and the non-match-trials (right). Lines show the median. Shaded areas indicate the full range (light) and the inter-quartile range (dark). Note that values that fall outside the range (1.5-times IQR beyond the 25th and 75th percentiles, respectively) are considered outliers. *y: you, f: friend and s: stranger; numbers give the repetition during learning for the different shape-non-word pairs.

Although, we were unable to draw any conclusions from fitting HDDMs to our data, we will briefly present the findings of the *a, v, z,* t_0 model for the standard learning condition below, as these might be of interest for further research.

During the experiment, all trials with reaction times shorter than 200ms or longer than 1100ms (time outs) were repeated to ensure sufficient amounts of data for modelling; 11.5% of trials were affected. The software was then set to remove 5% of the data as outliers. To model Bayesian posterior distributions, a Markov Chain Monte Carlo with 10,000 bootstraps was used. The first 1,000 bootstraps were discarded as burn in samples.

Supplementary table 2: Pairwise comparisons for the estimated parameters based on the different identities. Note that a value for p_{Bayes} of 0.5 would indicate full overlap of the estimated distributions of the two parameters and values of 0 or 1 indicate that the estimates do not overlap at all.

Parameter	**Comparison**	p_{Bayes}
Threshold separation *(a)*	you > friend	.927
	you > stranger	.956
	friend > stranger	.609
Non-decisional processes *(t_0)*	you > friend	.491
	you > stranger	.369
	friend > stranger	.372
Drift rate *(v)*	you > friend	.993
	you > stranger	.995
	friend > stranger	.547
Starting point *(z)*	you > friend	>.999
	you > stranger	>.999
	friend > stranger	.622



Supplementary figure 2: Posterior distributions for the standard learning condition. Parameters were estimated using HDDM where the parameters a, v, z, t_0 estimated seperatly for the three identities (you, friend and stranger)

The posterior distributions for all parameters are plotted in supplementary figure 2. Parameters were estimated separately for the different identities. Supplementary table 3 shows the pairwise comparisons between the posterior distributions for each of the parameters. These data show that associations related to the self are processed in distinct ways compared to associations not related to the self. The threshold separation a was larger for the self-related trials than for the other related trials. A larger threshold separation reduces the influence of noise and increases the probability of a correct response (Ratcliff & Tuerlinckx, 2002). For all identities, the starting point z was elevated suggesting a bias towards responding 'match'. This was more pronounced for the self-related trials than for the other related trials. A bias in this direction leads to shorter reaction times in the match- compared to the non-match trials. Additionally, the drift rate was higher for the self-related trials, than for the friend- and the stranger-related trials. A higher drift rate (steeper slope) again leads to shorter reaction times. Interestingly, there was no difference in non-decisional processes, such as motor response requirements. These data may explain the higher accuracy (larger threshold separation) and at the same time shorter reaction times (faster (perceptual) evidence accumulation), especially in the match-trials (starting point shifted towards 'match') for self-related association. However, as this is based on the model with the best fit that did converge, not on the model with the best fit overall, this interpretation should be considered with caution.

References

Kass, R. E., & Raftery, A. E. (1995). Bayes Factors. *Journal of the American Statistical Association*, *90*(430), 773–795. https://doi.org/10.2307/2291091

Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review*, *9*(3), 438–481.

Sui, J., He, X., & Humphreys, G. W. (2012). Perceptual effects of social salience: evidence from self-prioritization effects on perceptual matching. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(5), 1105.

Voss, A., Nagler, M., & Lerche, V. (2013). Diffusion models in experimental psychology. *Experimental Psychology*. Retrieved from http://econtent.hogrefe.com/doi/full/10.1027/1618-3169/a000218

Wiecki, T. V., Sofer, I., & Frank, M. J. (2013). HDDM: Hierarchical Bayesian estimation of the Drift-Diffusion Model in Python. *Frontiers in Neuroinformatics*, *7*. https://doi.org/10.3389/fninf.2013.00014