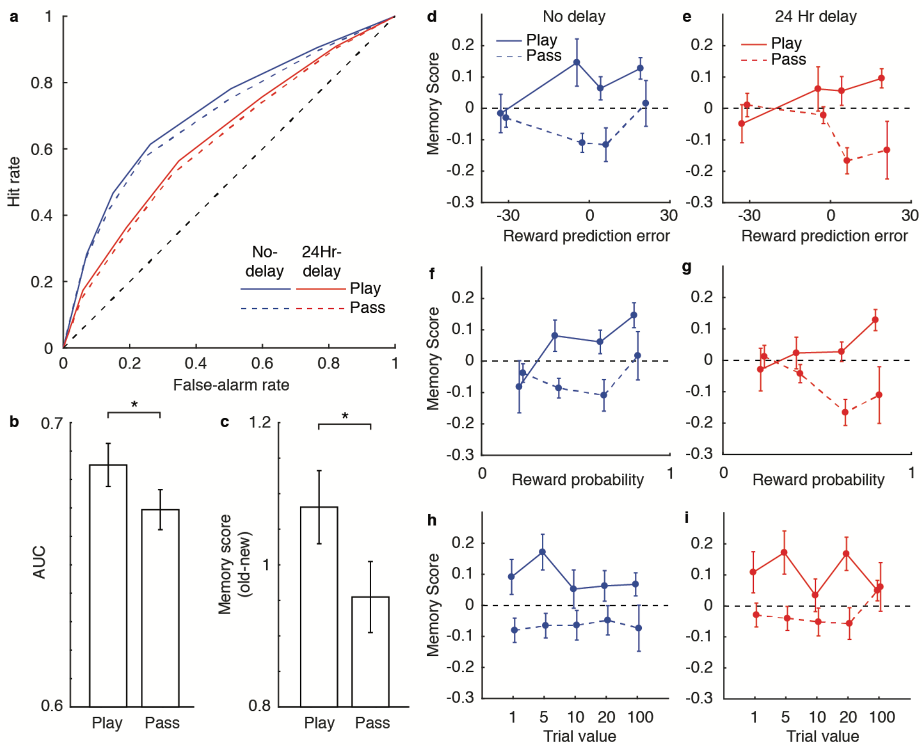
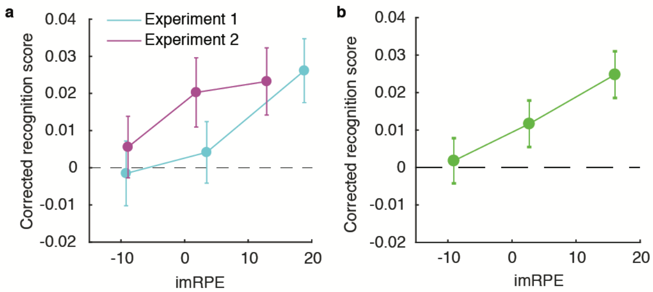
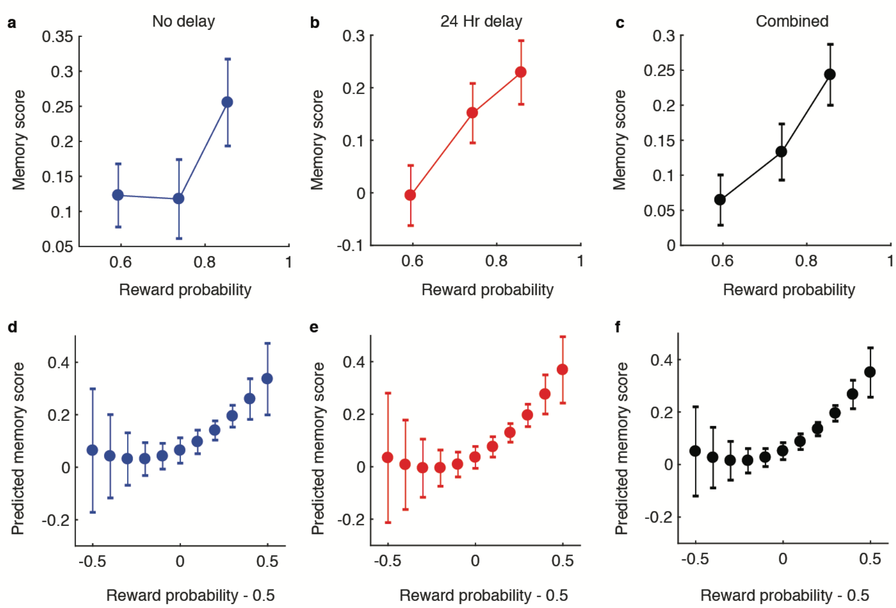
Supplementary figures



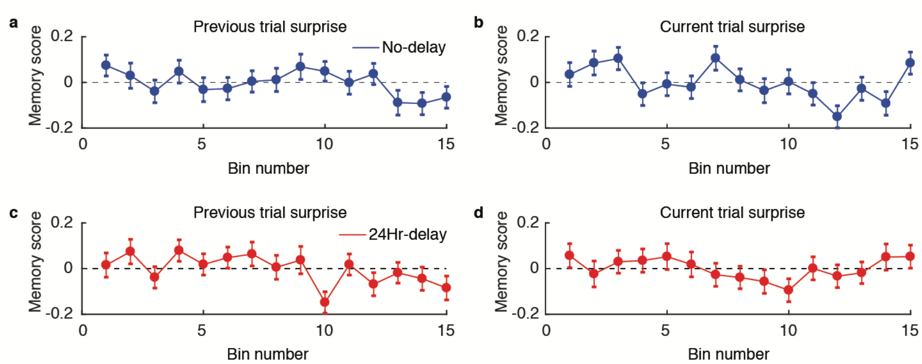
Supplementary Figure 1. Replication of main effects after omitting all trials in which subsequent memory had low confidence (“guess”). To rule out the possibility that low confidence trials are driving our results or adding unnecessary noise, we repeated the main findings of the study after omitting all trials in which the confidence score was 1 (both target and foil). **a**, The ROC curves for play/pass trials. **b**, Area under the ROC curve was greater for play versus pass (t(199) = 2.77, P = 0.0061, d = 0.20, 95% CI: 0.0045–0.027; group difference, t(198) = -0.060, P = 0.95, d = -0.0085, 95% CI: -0.023–0.022). **c**, Mean pairwise difference in memory score between the “old” images and their semantically-matched foil images was greater for play versus pass (t(198) = 3.48, P < 0.001, d = 0.25, 95% CI: 0.055–0.20; group difference, t(197) = 0.92, P = 0.36, d = 0.13, 95% CI: -0.076–0.21). **d**,**e**, Positive relationship between image RPE and subsequent memory (t(199) = 2.61, P = 0.0096, d = 0.18, 95% CI: 0.0008–0.0054; group difference, t(198) = -0.49, P = 0.63), d = -0.069, 95% CI: -0.0058–0.0035. **f**,**g**, Positive relationship between reward probability and subsequent memory (t(199) = 2.95, P = 0.0036, d = 0.21, 95% CI: 0.093–0.47, group difference, t(198)=0.14, P = 0.89, d = 0.020, 95% CI: -0.35–0.40). **h**,**i**, No relation between reward value and subsequent memory (t(199) = -1.68, P = 0.094, d = -0.12, 95% CI: -0.056–0.0044, group difference, t(198) = 0.76, P = 0.45, d = 0.11, 95% CI: -0.037–0.084).



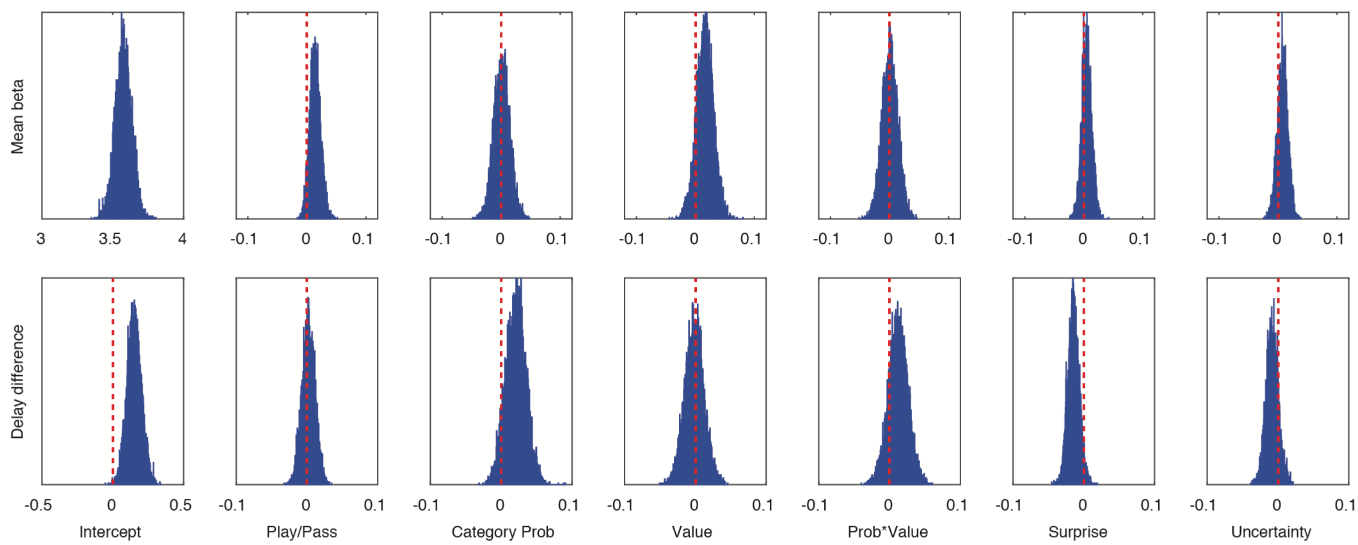
Supplementary Figure 2. The effect of image RPE on response bias-adjusted recognition memory. It is possible that our finding of image RPE's effect on memory may be entirely driven by a shift in the response threshold, rather than a stimulus-specific increase in recognition memory. To test for this, we re-analyzed our data using the corrected recognition score, computed as hit rate minus false alarm rate on a trial-by-trial basis, for the play trials. **a**, Effect of image RPE on corrected recognition for each experiment. We find that our main effect of image RPE vs. recognition memory holds for Experiment 1 ($t(199) = 2.55$, $P = 0.012$, $d = 0.18$, 95% CI: 0.0003–0.0022), and shows a non-significant, but positive trend for Experiment 2 ($t(173) = 1.25$, $P = 0.21$, $d = 0.095$, 95% CI: -0.0005–0.0022). **b**, Effect of image RPE on corrected recognition after combining data from both experiments. With the combined data, we find a significant effect of image RPE on memory ($t(373) = 2.57$, $P = 0.011$, $d = 0.13$, 95% CI: 0.0002–0.0019).



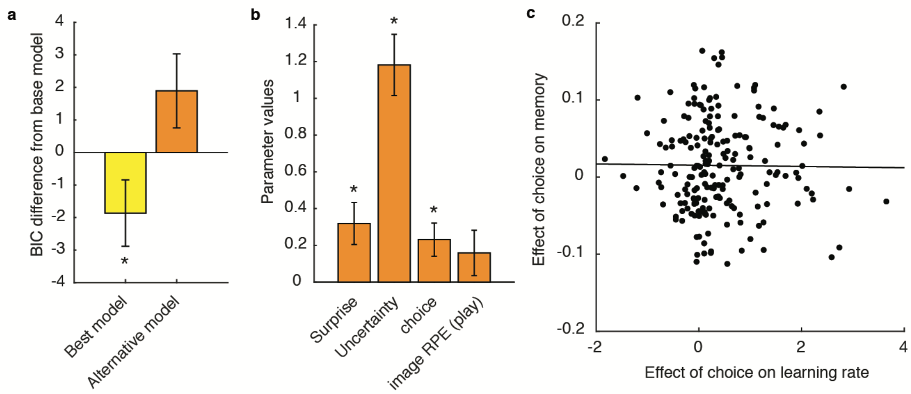
Supplementary Figure 3. Testing the effect of reward uncertainty on recognition memory. **a-c**, Effect of reward probability on memory score, for probabilities greater than 0.5. If our results are driven by reward uncertainty, we should see a negative effect of reward probability on memory if we only consider trials in which the probability ranges from 0.5 to 1. We show this relationship for the no delay (**a**), 24 hour delay (**b**) and combined group (**c**) in Experiment 1. To test for the effect of reward probability on memory, we pooled all trials with a subjective reward probability greater than 0.5, then fit a mixed-effects regression model in which the dependent variable was memory score, and the independent variables included an intercept term, reward probability, a binary variable for whether the lure image was seen before the target image during the memory task, the order in which each image was presented during the learning task, and dummy variables associated with each participant and each unique image. The coefficient for reward probability was significantly greater than zero across participants (no delay: t(5730) = 2.28, P = 0.023, 95% CI beta: 0.081–1.08; 24hr delay: t(4716) = 2.21, P = 0.027, 95% CI beta: 0.066–1.10; combined: t(10608) = 3.03, P = 0.0024, 95% CI beta: 0.20–0.91), indicating that reward probability was positively associated with recognition memory strength. **d-f**, We also tested for the presence of a quadratic relationship between reward probability and memory strength, where the reward probabilities are shifted by 0.5, so that the edges correspond to lower reward uncertainty. Therefore, a significant quadratic relationship would indicate an effect of reward uncertainty on memory strength. We fit a linear regression model of the form $y = \beta_0 + \beta_1 * x + \beta_2 * x^2$, where y is memory score and x is reward probability shifted by 0.5, to each participant's data. We then performed a t-test on the β_2 coefficients, and found no significant difference from zero (no delay (**d**): t(108) = 0.69, P = 0.49, d = 0.066, 95% CI: -1.02–2.11; 24 hour delay (**e**): t(90) = 0.90, P = 0.37, d = 0.094, 95% CI: -0.81–2.13; combined (**f**): t(199) = 1.10, P = 0.27, d = 0.078, 95% CI: -0.48–1.67). The figures show the average predicted curve from the regression fits across both experiments, which if anything shows a downward hump. Error bars indicate standard error across participants. Colors indicate time between encoding and memory testing; blue = no delay, red = 24 hour delay, black = combined.



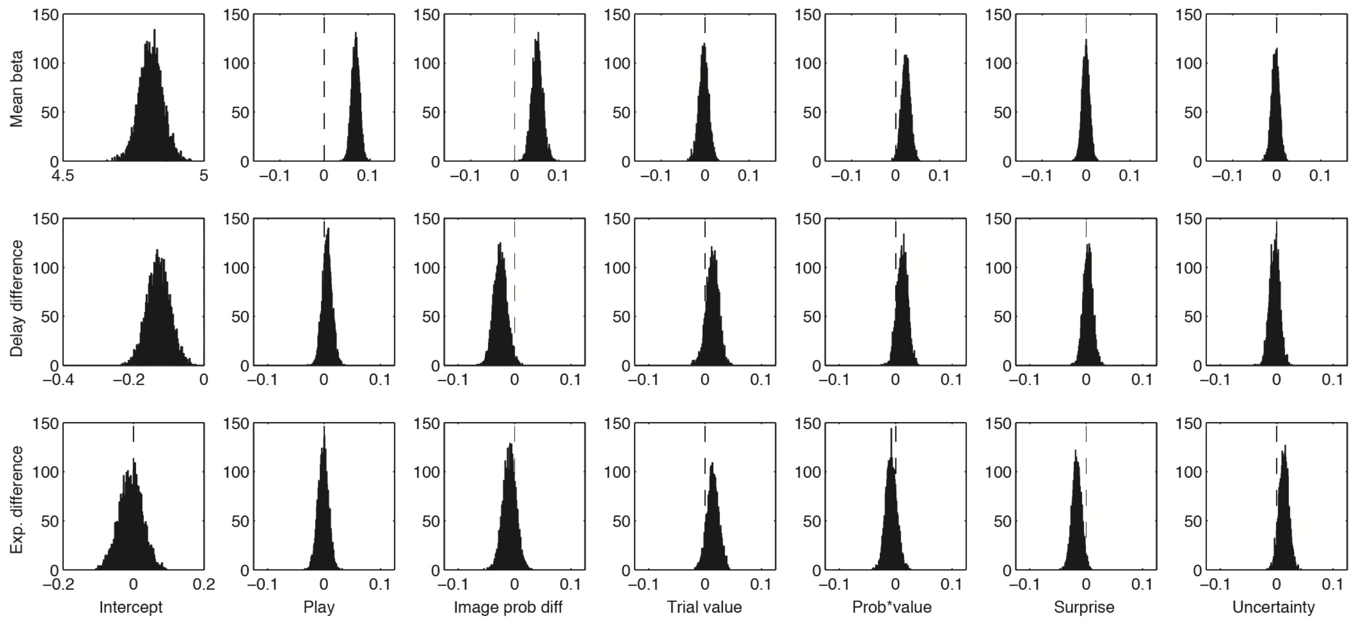
Supplementary Figure 4. Effect of strong versus weak surprise on memory strength. We tested whether extreme values of surprise had differential effects on memory. To do so, we analyzed the effect of surprise on memory after binning surprise (current & previous trial) into 10 equally-sized bins per participant. Plotting the across-participant mean memory score for each bin, we found no clear positive or negative trend, and no indication that extreme values of surprise lead to different effects on memory.



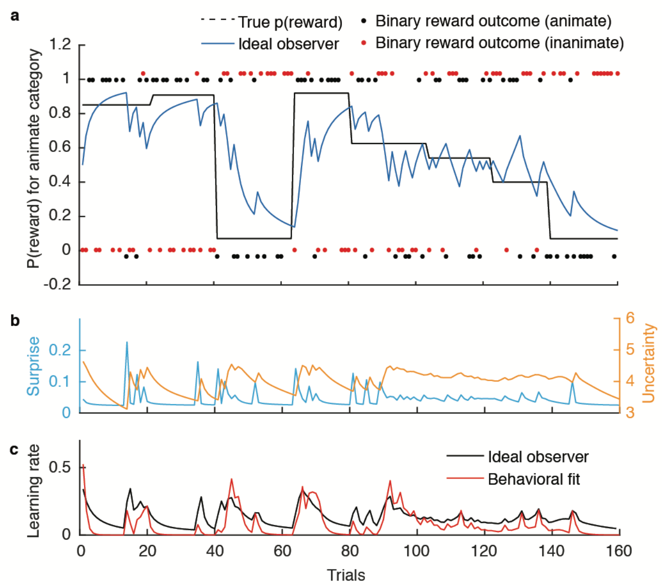
Supplementary Figure 5. Hierarchical regression estimating memory scores for foil items. Posterior probability densities for mean predictor coefficients (μ_x; top row) and delay condition parameter difference (D_x; bottom row) estimated through MCMC sampling over the graphical model described in Fig. 6a informed by data for foil items semantically matched to the images presented in the decision making task. Unlike fits to the target items (Fig. 6b), coefficients for task predictors related to reward probability and play/pass did not deviate appreciably from zero. However, delay condition difference coefficients were positive for the intercept term, indicating that participants in the 24 hour delay condition tended to report higher memory scores for foils, consistent with the poorer discriminability in the 24 hour delay condition (see Fig. 3).



Supplementary Figure 6. RL model fitting to test for the effect of anticipatory attention on learning and memory. **a**, Comparison of the Bayesian information criterion (BIC) of the best fitting model (yellow bar, corresponding to the yellow bar in Fig. 2c) and a model including both choice and image RPE as additional parameters. The difference between the BIC of the two models versus the BIC of the base model, which includes no additional parameters to explain learning rate, is plotted here (see "RL model fitting" under Methods). This BIC difference was significantly less than zero for the best fitting model (yellow bar; $t(199) = -2.39$, $P = 0.018$, $d = -0.17$, 95% CI: -4.53 to -0.43), indicating that the model better explains participant choice behavior, but not for the alternative model (orange bar; $t(199) = 1.66$, $P = 0.098$, $d = 0.12$, 95% CI: -0.35 to 4.14). **b**, Maximum likelihood estimates of surprise ($t(199) = 2.78$, $P = 0.0060$, $d = 0.20$, 95% CI: 0.093–0.55), uncertainty ($t(199) = 7.10$, $P < 0.001$, $d = 0.50$, 95% CI: 0.85–1.51), choice ($t(199) = 2.57$, $P = 0.011$, $d = 0.18$, 95% CI: 0.054–0.41), and image RPE_{play} ($t(199) = 1.29$, $P = 0.20$, $d = 0.091$, 95% CI: -0.084–0.40) parameters of the alternative model (orange bar in **a**). For **a**&**b**, error bars indicate standard error across participants, and asterisks indicate statistical significance. **c:** Effect of playing (gambling) on learning rate vs. effect of playing on recognition memory. The degree to which participants (black points) adjusted learning rate according to play/pass behavior (abscissa) showed no association with the degree to which they enhanced memory of items presented on play trials (ordinate; spearman $r(198) = 0.047$, $P = 0.51$).



Supplementary Figure 7. Coefficient estimates for hierarchical model fit to all participants across both experiments. **Top:** Posterior probability density over mean coefficient estimates at the population level for each parameter in the hierarchical model fit to participant memory scores. Leftmost column reflects the intercept indicating overall memory scores for old items, whereas all other columns reflect the degree to which learning task-related factors affected subsequent memory. **Middle:** Posterior probability density over delay difference estimates for each parameter in the hierarchical model. **Bottom:** Posterior probability density over experiment difference estimates quantifying the difference in coefficient values across the two experiments for each parameter in the hierarchical model.



Supplementary Figure 8. Example task structure (**a**), with normative measures of surprise (**b**), uncertainty (**b**), and learning rate (**c**) across time. Learning rate from the Bayesian ideal observer model (**c**, black line) was computed on a trial-by-trial basis (see Methods: Bayesian ideal observer model). Learning rate from the RL model (**c**, red line) was computed by fitting the behavior of an example participant with an RL model that included both surprise and uncertainty as modulators of learning rate (see Fig. 2c; yellow bar).

Supplementary Tables:

Supplementary Table 1. Mean and 95% credible interval of the coefficient estimates for hierarchical model fit to all participants across experiment 1, experiment 2, and combined between both experiments. Each column represents a variable in the learning task. For each experiment and the combined data set, first row represents data after combining both delay conditions, and the second row represents the difference between the two delay conditions.

Experiment	Coefficient	Intercept	Play/Pass	Reward Prob.	Reward prob. (other category)	Trial value	Probability *value	Surprise	Uncertainty
Experiment 1	Mean beta	4.8 [4.69, 4.90]	0.078 [0.05, 0.10]	0.047 [0.01, 0.08]	*	-0.03 [-0.052, 0.0001]	0.042 [0.01, 0.07]	0.012 [-0.01,0.03]	-0.019 [-0.04, 0.01]
	Delay difference	-0.14 [-0.23,-0.05]	-0.008 [-0.03,0.02]	-0.043 [-0.07,-0.01]	*	0.003 [-0.03,0.03]	0.019 [-0.02,0.05]	0.01 [-0.01,0.03]	-0.004 [-0.03,0.02]
Experiment 2	Mean beta	4.81 [4.71,4.92]	0.0685 [0.046,0.092]	0.0226 [-0.011,0.056]	-0.0303 [-0.061,-0.001]	0.012 [-0.016,0.043]	0.0073 [-0.019,0.034]	-0.017 [-0.04,0.01]	0.0096 [-0.012,0.032]
	Delay difference	-0.13 [-0.22,-0.03]	0.019 [-0.005,0.018]	-0.013 [-0.046,0.018]	0.01 [-0.020,0.040]	0.03 [-0.016,0.043]	0.01 [-0.015,0.035]	-0.005 [-0.030,0.020]	-0.002 [-0.02,0.020]
Combined	Mean beta	4.81 [4.7,4.9]	0.07 [0.05,0.09]	0.05 [0.03,0.07]	*	0.003 [-0.02,0.02]	0.02 [0.004,0.04]	-0.001 [-0.02,0.01]	-0.001 [-0.02,0.01]
	Delay difference	-0.13 [-0.19,-0.07]	0.00 [-0.01,0.02]	-0.03 [-0.05, -0.002]	*	0.014 [-0.01, 0.03]	0.01 [-0.01, 0.03]	0.00 [-0.01,0.02]	-0.00 [-0.02,0.01]

Supplementary Table 2. Proportion of total posterior density falling on values greater than zero. For example, a value of 1 indicates that all samples from posterior were greater than zero and 0 indicates that no sample from the posterior took a value greater than zero.

Experiment	Coefficient	Intercept	Play/Pass	Reward Prob.	Reward prob. (other category)	Trial value	Probability *value	Surprise	Uncertainty
Experiment 1	Mean beta	1	1	0.995	*	0.0253	0.995	0.89	0.07
	Delay difference	0.002	0.27	0.01	*	0.57	0.87	0.82	0.37
Experiment 2	Mean beta	1	1	0.916	0.022	0.799	0.705	0.086	0.8
	Delay difference	0.004	0.947	0.21	0.74	0.8	0.76	0.36	0.44
Combined	Mean beta	1	1	1	*	0.387	0.994	0.43	0.3
	Delay difference	0	0.68	0.91	*	0.91	0.91	0.68	0.39