Supporting information 1019

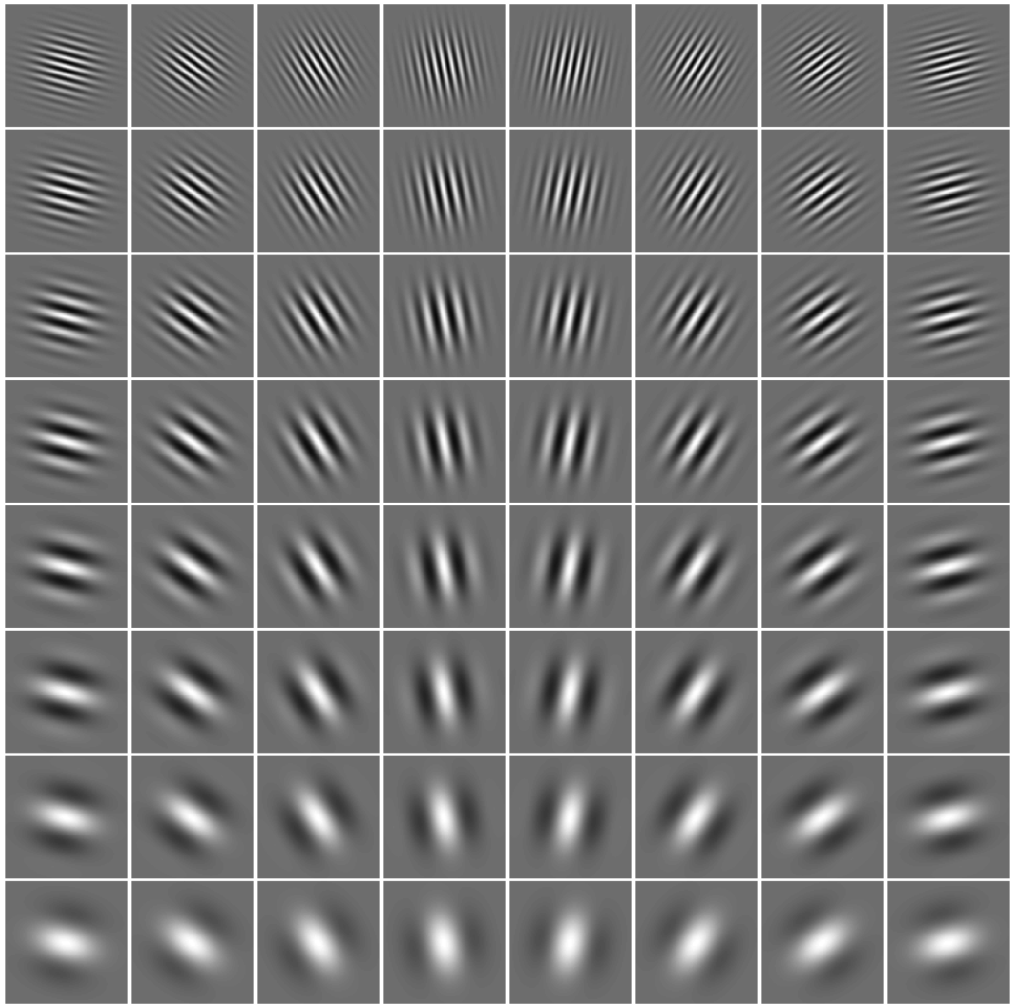

Fig S1. Gabor stimuli. Tilt varies from left to right from 105° to 255° in equally spaced intervals, while stripe frequency increases moving upwards from 1.5 to 15 in log intervals.

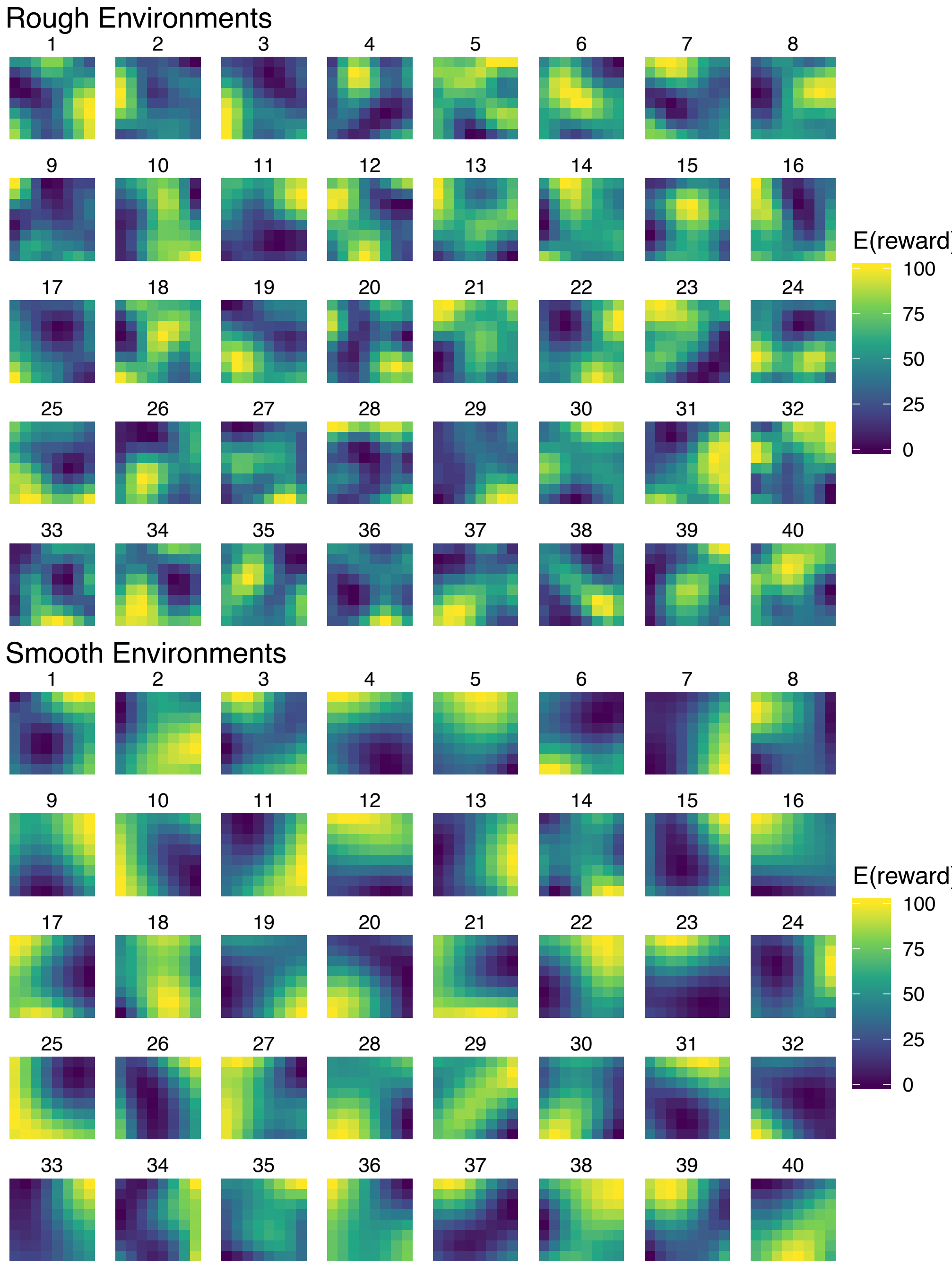

Fig S2. Correlated reward environments. Heatmaps of the reward environments used in both spatial and conceptual domains. The color of each tile represents the expected reward of the bandit, where the x-axis and y-axis were mapped to the spatial location or the tilt and stripe frequency (respectively). All environments have the same minimum and maximum reward values, and the two classes of environments share the same expectation of reward across options.

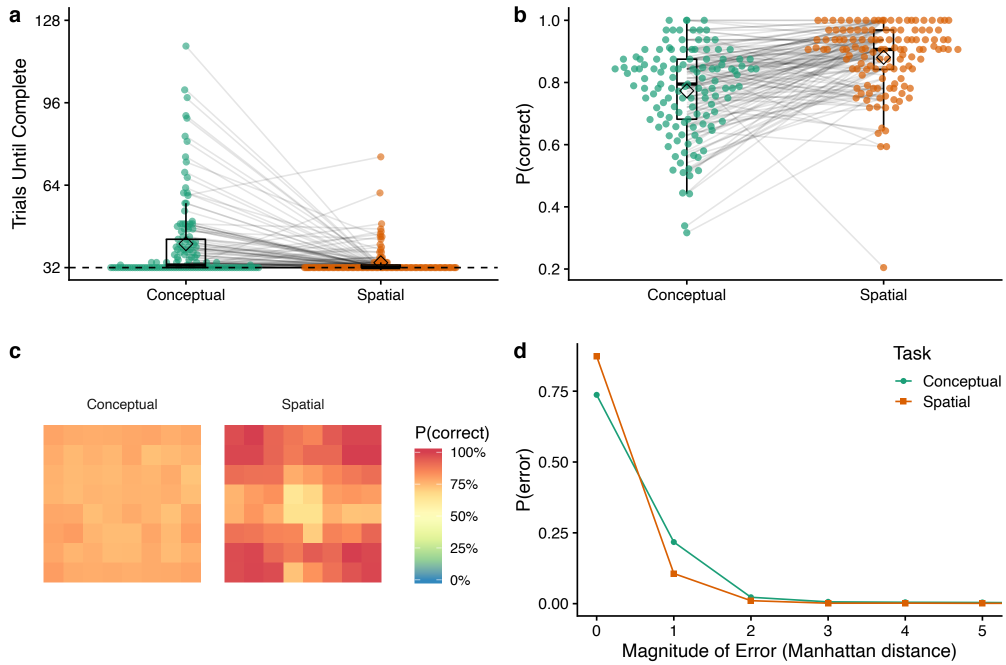

Fig S3. Training Phase. **a**) Trials needed to reach the learning criterion (90% accuracy over 10 trials) in the training phase, where the dotted line indicates the 32 trial minimum. Each dot is a single participant with lines connecting the same participant. Tukey boxplots show median (line) and 1.5x IQR, with diamonds indicating group means. **b**) Average correct choices during the training phase. In the last 10 trials before completing the training phase, participants had a mean accuracy of 95.0% on the spatial task and 92.7% on the conceptual task (difference of 2.3%). In contrast, in the first 10 trials of training, participants had a mean accuracy of 84.1% in the spatial task and 68.8% in the conceptual (difference of 15.4%). **c**) Heatmaps of the accuracy of different target stimuli, where the x and y-axes of the conceptual heatmap indicate tilt and stripe frequency, respectively. **d**) The probability of error as a function of the magnitude of error (Manhattan distance from the correct response). Thus, most errors were close to the target, with higher magnitude errors being monotonically less likely to occur.

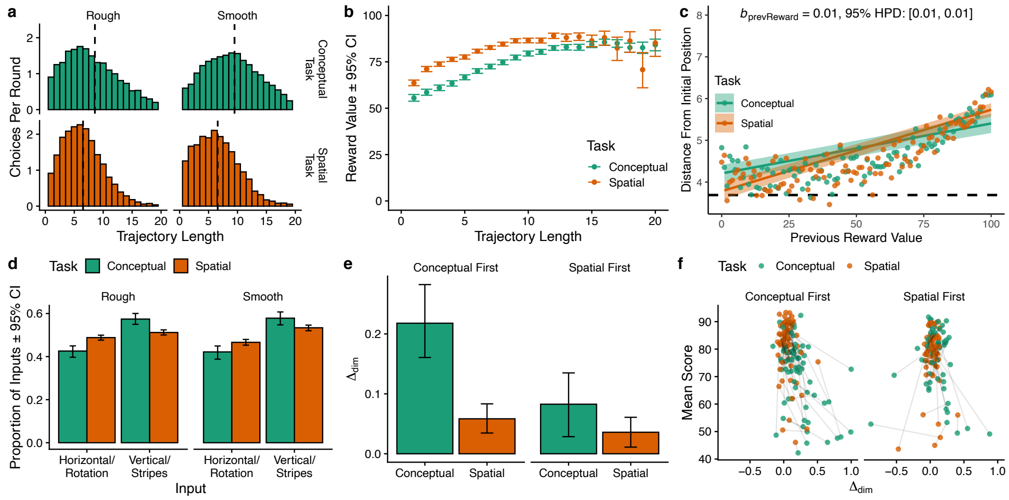

Fig S4. Search Trajectories. **a**) Distribution of trajectory length, separated by task and environment. The dashed vertical line indicates the median for each category. Participants had longer trajectories in the contextual task ($t(128) = -10.7$, $p < .001$, $d = 1.0$, $BF > 100$), but there were no differences across environments ($t(127) = 1.3$, $p = .213$, $d = 0.2$, $BF = .38$). **b**) Average reward value as a function of trajectory length. Longer trajectories were correlated with higher rewards ($r = .23$, $p < .001$, $BF > 100$). Each dot is a mean with error bars showing the 95% CI. **c**) Distance from the random initial starting point in each trial as a function of the previous reward value. Each dot is the aggregate mean, while the lines show the fixed effects of a Bayesian mixed-effects model (see Table S1), with the ribbons indicating the 95% CI. The relationship is not quite linear, but is also found using a rank correlation ($r_\tau = .18$, $p < .001$, $BF > 100$). The dashed line indicates random chance. **d**) Search trajectories decomposed into the vertical/stripe frequency dimension vs. horizontal/tilt dimension. Bars indicate group means and error bars show the 95% CI. We find more attention given to the vertical/stripe frequency dimension in both tasks, with a larger effect for the conceptual task ($F(1, 127) = 26.85$, $p < .001$, $\eta^2 = .08$, $BF > 100$), but no difference across environments ($F(1, 127) = 1.03$, $p = .311$, $\eta^2 = .005$, $BF = 0.25$). **e**) We compute attentional bias as $\Delta_{\text{dim}} = P(\text{vertical/stripe frequency}) - P(\text{horizontal/tilt})$, where positive values indicate a stronger bias towards the vertical/stripe frequency dimension. Attentional bias was influenced by the interaction of task order and task ($F(1, 127) = 8.1$, $p = .005$, $\eta^2 = .02$, $BF > 100$): participants were more biased towards the vertical/stripe frequency dimension in the conceptual task when the conceptual task was performed first ($t(66) = -6.0$, $p < .001$, $d = 0.7$, $BF > 100$), but these differences disappeared when the spatial task was performed first ($t(61) = -1.6$, $p = .118$, $d = 0.2$, $BF = .45$). **f**) Differences in attention and score. Each participant is represented as a pair of dots, where the connecting line shows the change in score and Δ_{dim} across tasks. We found a negative correlation between score and attention for the conceptual task only in the conceptual first order ($r_\tau = -.31$, $p < .001$, $BF > 100$), but not in the spatial first order ($r_\tau = -.07$, $p = .392$, $BF = .24$). There were no relationships between score and attention in the spatial task in either order (spatial first: $r_\tau = .03$, $p = .738$, $BF = .17$; conceptual first: $r_\tau = -.03$, $p = .750$, $BF = .17$).

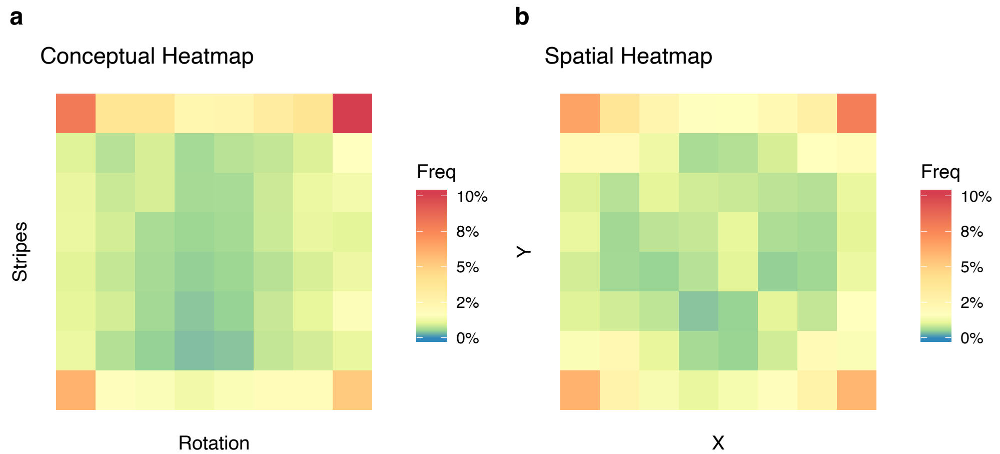

Fig S5. Heatmaps of choice frequency. Heatmaps of chosen options in **a**) the Gabor feature of the conceptual task and **b**) the spatial location of the spatial task, aggregated over all participants. The color shows the frequency of each option centered on yellow representing random chance (1/64), with orange and red indicating higher than chance, while green and blue were lower than chance.

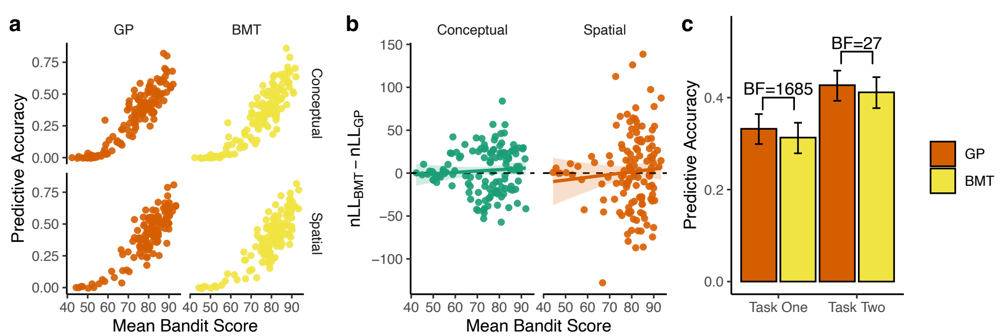

Fig S6. Additional Modeling Results. **a**) The relationship between mean performance and predictive accuracy, where in all cases, the best performing participants were also the best described. **b**) The best performing participants were also the most diagnostic between models, but not substantially skewed towards either model. Linear regression lines strongly overlap with the dotted line at $y = 0$, where participants above the line were better described by the GP model. **c** Model comparison split by which task was performed first vs. second. In both cases, participants were better described on their second task, although the superiority of the GP over the BMT remains, comparing only task one (paired t-test: $t(128) = 4.6$, $p < .001$, $d = 0.10$, $BF = 1685$) or only task two ($t(128) = 3.5$, $p < .001$, $d = 0.08$, $BF = 27$).

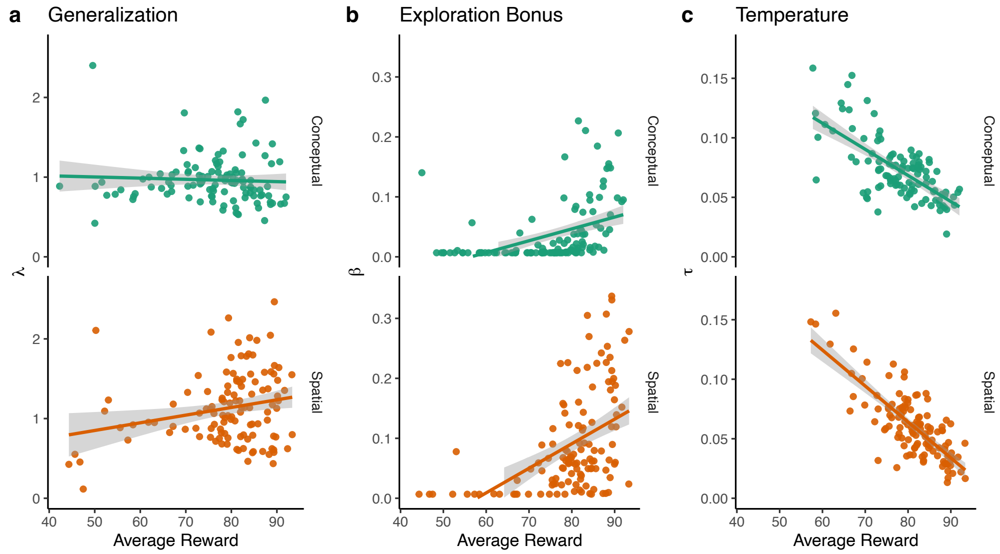

Fig S7. GP parameters and performance. **a**) We do not find a consistent relationship between λ estimates and performance, which were anectdotally correlated the spatial task ($r_\tau = .13$, $p = .030$, $BF = 1.2$) or negatively correlated in conceptual task ($r_\tau = -.22$, $p < .001$, $BF > 100$). **b**) Higher beta estimates were strongly predictive of better performance in both conceptual ($r_\tau = .32$, $p < .001$, $BF > 100$) and spatial tasks ($r_\tau = .31$, $p < .001$, $BF > 100$). **c**) On the other hand, high temperature values predicted lower performance in both conceptual($r_\tau = -.59$, $p < .001$, $BF > 100$) and spatial tasks ($r_\tau = -.58$, $p < .001$, $BF > 100$).

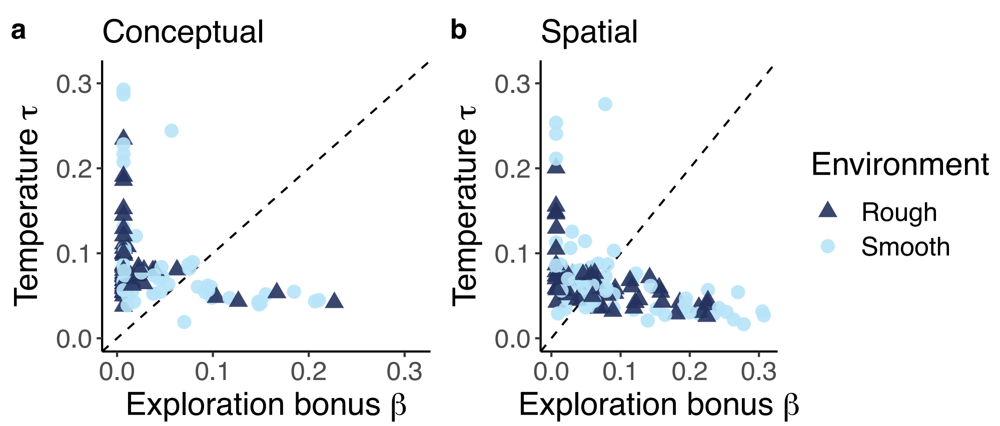

Fig S8. GP exploration bonus and temperature. We check here whether there exists any inverse relationship between directed and undirected exploration, implemented using the UCB exploration bonus β (x-axis) and the softmax temperature τ (y-axis), respectively. Results are split into conceptual (**a**) and spatial tasks (**b**), where each dot is a single participant and the dotted line indicates $y = x$. The upper axis limits are set to the largest 1.5×IQR, for both β and τ, across both conceptual and spatial tasks.

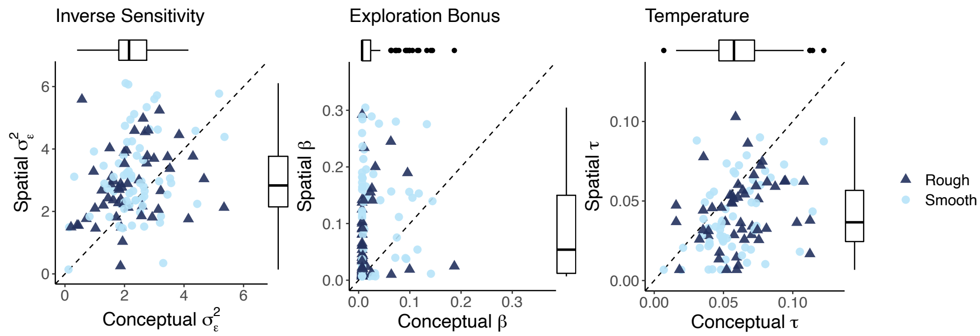

Fig S9. BMT parameters. Each dot is a single participant and the dotted line indicates $y = x$. **a**) We found lower error variance (σ^2_ϵ) estimates in the conceptual task (Wilcoxon signed-rank test: $Z = -4.8$, $p < .001$, $r = -.42$, $BF > 100$), suggesting participants were more sensitive to the reward values (i.e., more substantial updates to their means estimates). Error variance was also somewhat correlated across tasks ($r_\tau = .18$, $p = .003$, $BF = 10$). **b**) As with the GP model reported in the main text, we also found strong differences in exploration behavior in the BMT. We found lower estimates of the exploration bonus in the conceptual task ($Z = -5.9$, $p < .001$, $r = -.52$, $BF > 100$). The exploration bonus was also somewhat correlated between tasks ($r_\tau = .16$, $p = .006$, $BF = 4.8$). **c**) Also in line with the GP results, we again find an increase in random exploration in the conceptual task ($Z = -6.9$, $p < .001$, $r = -.61$, $BF > 100$). Once more, temperature estimates were strongly correlated ($r_\tau = .34$, $p < .001$, $BF > 100$).

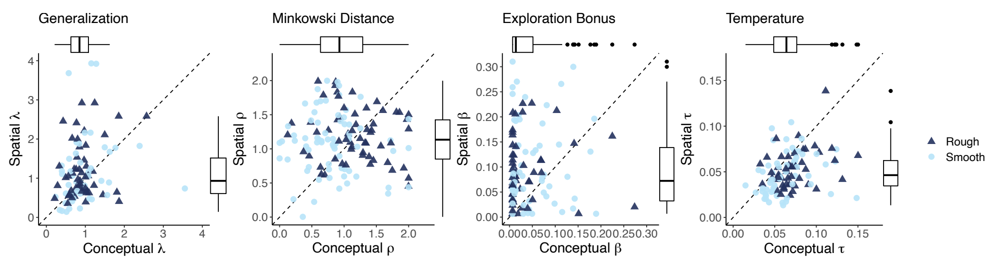

Fig S10. Shepard kernel parameters. We also considered an alternative form of the GP model. Instead of modeling generalization as a function of squared-Euclidean distance with the RBF kernel, we use the Shepard kernel described in [65], where we instead use Minkowski distance with the free parameter $\rho \in [0, 2]$. This model is identical to the GP model reported in the main text when $\rho = 2$. But when $\rho < 2$, the input dimensions transition from integral to separable representations [112]. The lack of clear differences in model parameters motivated us to only include the standard RBF kernel in the main text. **a**) We find no evidence for differences in generalization between tasks ($Z = -1.8$, $p = .039$, $r = -.15$, $BF = .32$). There is also marginal evidence of correlated estimates ($r_\tau = .13$, $p = .026$, $BF = 1.3$). **b**) There is anecdotal evidence of lower ρ estimates in the conceptual task ($Z = -2.5$, $p = .006$, $r = -.22$, $BF = 2.0$). The implication of a lower ρ in the conceptual domain is that the Gabor features were treated more independently, whereas the spatial dimensions were more integrated. However, the statistics suggest this is not a very robust effect. These estimates are also not correlated ($r_\tau = -.02$, $p = .684$, $BF = .12$). **c**) Consistent with all the other models, we find systematically lower exploration bonuses in the conceptual task ($Z = -5.5$, $p < .001$, $r = -.49$, $BF > 100$). There was weak evidence of a correlation across tasks ($r_\tau = .14$, $p = .021$, $BF = 1.6$). **d**) We find clear evidence of higher temperatures in the conceptual task ($Z = -6.3$, $p < .001$, $r = -.56$, $BF > 100$), with strong correlations across tasks ($r_\tau = .41$, $p < .001$, $BF > 100$)

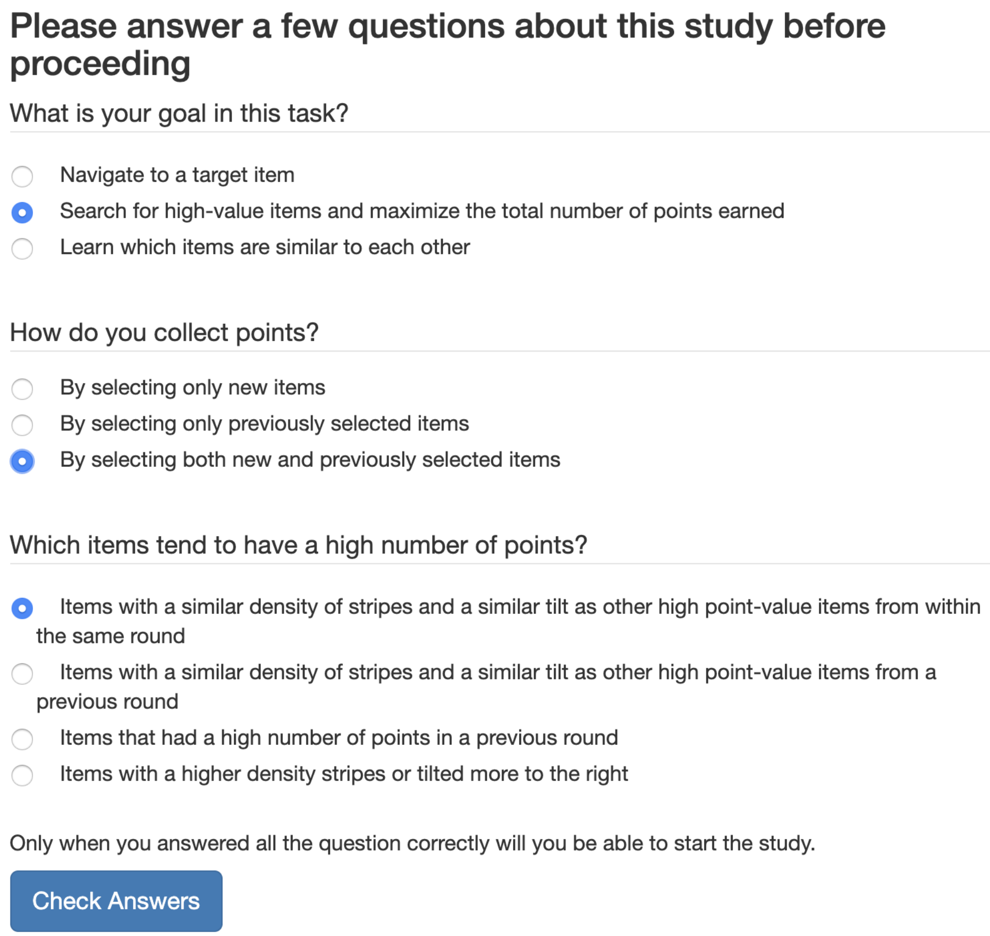

Fig S11. Comprehension questions for the conceptual task. The correct answers are highlighted.

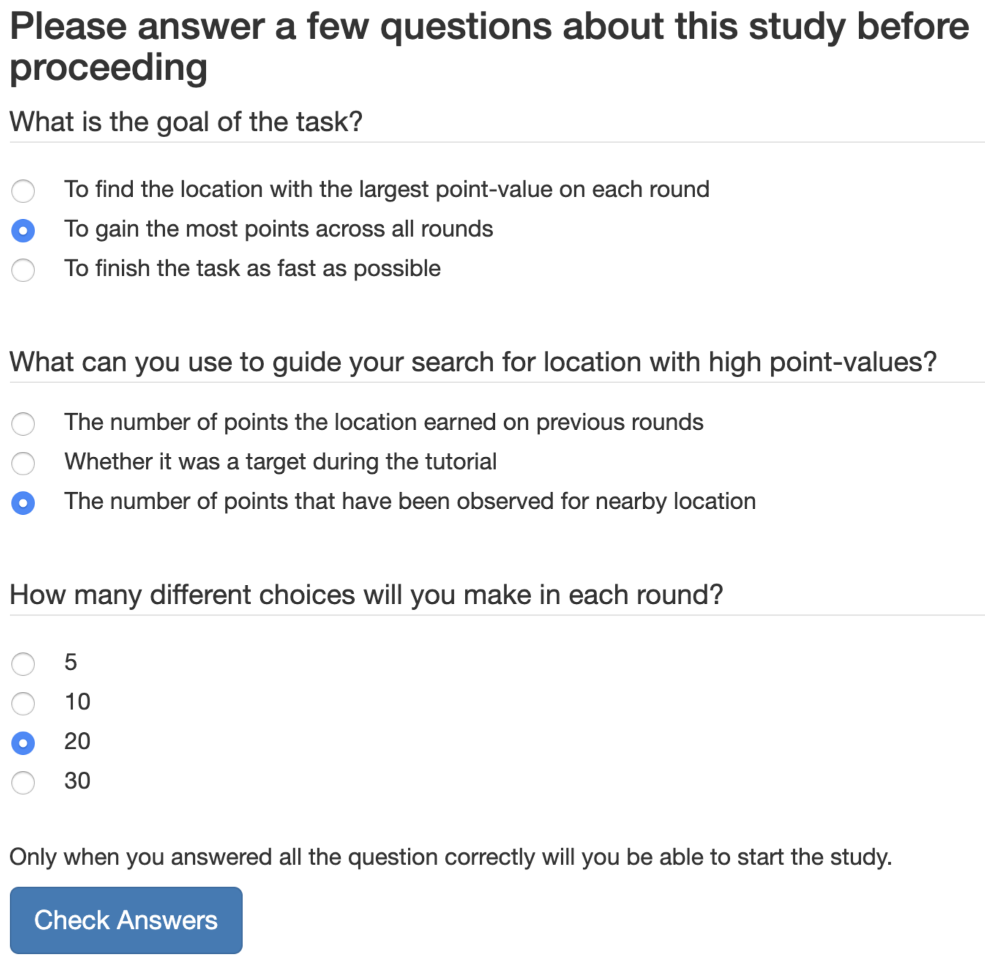

Fig S12. Comprehension questions for the spatial task. The correct answers are highlighted.

Table S1. Mixed Effects Regression Results: Previous Reward

	Distance Between Choices		Distance from Initial Position	
Predictors	*Est.*	*95% HPD*	*Est.*	*95% HPD*
Intercept	7.04	6.77 – 7.31	4.21	4.00 – 4.41
PreviousReward	-0.06	-0.06 – -0.06	0.01	0.01 – 0.01
Spatialtask	1.03	0.68 – 1.38	-0.2043	-0.67 – -0.18
PreviousReward:Spatialtask	-0.01	-0.02 – -0.01	0.01	0.004 – 0.01
Random Effects				
σ^2	1.10		1.08	
τ_{00}	7.22		8.34	
N	129		129	
Observations	44118		44118	
Bayesian R^2	.539		.118	

Note: We report the posterior median (Est.) and 95% highest posterior density (HPD) interval. σ^2 indicates the individual-level variance and τ_{00} indicates the variation between individual intercepts and the average intercept. See Methods for full specification of model structure and priors.

Table S2. Mixed Effects Regression Results: Bonus round judgments

	Model Prediction		Model Uncertainty	
Predictors	*Est.*	*95% HPD*	*Est.*	*95% HPD*
Intercept	12.42	10.45 – 14.34	0.92	0.86 – 0.98
ParticipantJudgment	0.82	0.75 – 0.89	-0.02	-0.02 – -0.01
Spatialtask	1.34	-0.33 – 3.04	0.002	-0.08 – 0.08
ParticipantJudgment:Spatialtask	0.03	-0.04 – 0.11	-0.01	-0.02 – 0.001
Random Effects				
σ^2	183.49		0.06	
τ_{00}	346.35		0.03	
N	129		129	
Observations	2580		2580	
Bayesian R^2	.437		.674	

Note: We report the posterior median (Est.) and 95% highest posterior density (HPD) interval. In the first model (Model Prediction), participant judgments in the range [1,100] are used to predict the GP posterior mean, whereas the second model (Model Uncertainty) uses confidence judgments in the range [1,11] to predict the GP posterior variance. All GP posteriors are computed based on individual participant λ-values, estimated from the corresponding bandit task. σ^2 indicates the individual-level variance and τ_{00} indicates the variation between individual intercepts and the average intercept. See Methods for full specification of model structure and priors.