Rule	'Relevant' VD	'Irrelevant' VD	'Relevant'-'Irrelevant' VD
ADD	-5.40; *0.0001*	-0.79*; n.s.*	-2.41*; 0.05*
MEAN	-7.63; *0.0001*	-2.31; *0.05*	-2.42; *0.05*
MULTIPLY	-2.21; *0.05*	1.15; *n.s.*	-3.48; *0.005*

Supplementary Table 1

Effect of alternative rules on reaction time regression.

One possible explanation for our results is that subjects did not in fact integrate across the probabilities in a Bayes optimal fashion, due to the way we instructed our subjects to perform the task. We therefore considered whether alternative rules for cue combination (mean, multiplication or sum) would explain away our differential effects on 'relevant' and 'irrelevant' attributes (supplementary information). Here we show these alternative rule combinations and their influence on subjects' log(reaction times). T-statistic (18 d.o.f.) is first number given; associated p-value is second number given; the model used was the same as in the main text. We considered several alternative rules for combining pS (stimulus-based probability of reward) and pA (action-based probability of reward). These were: 'ADD' ($pO_i = pS_i + pA_i$); 'MULTIPLY' ($pO_i = pS_i * pA_i$); or 'MEAN' ($pO_i = (pS_i + pA_i)/2$). The integrated value difference, using these rules (as opposed to those actually used to determine reward probabilities), was found to influence subjects' RTs (subjects were faster with higher integrated value difference, rightmost column in table). However, irrespective of which rule was used, the value difference on one attribute – the 'relevant' value difference – consistently had a greater influence on reaction time than the 'irrelevant' value difference.

Region	*Cluster size*	*Corrected p*	*Z-Max*
L IPS	1406	0.00231	3.65

Supplementary Table 2

Activations surviving whole-brain correction for contrast shown in main figure 5 (cluster forming threshold: Z>2.3).

Region	*Cluster size*	*Corrected p*	*Z-Max*	*Z-Max [X,Y,Z] (mm)*
Dorsal MFC	1032	0.00543	-3.85	*[0,38,42]*
R Angular gyrus	965	0.00813	-3.28	*[50,-44,44]*
L Anterior LOC/Angular gyrus	846	0.017	-3.52	*[-38,-60,44]*

Supplementary Table 3

Activations surviving whole-brain correction for contrast shown in main figure 7 (cluster forming threshold: Z<-2.3).

Supplementary Note on Computational Modeling

Model comparison with "random attribute" model. We also compared the hierarchical competition model with a "random attribute" model, in which the relevant attribute was selected stochastically on each trial. This model is identical in form to the hierarchical competition model, except for the specification of $w_i(t)$. Instead of $w_i(t)$ being a dynamic variable changing through time (by comparing within-attribute differences, see eq. 3 in computational model), it is randomly set to either 1 for the stimulus attribute and 0 for the action attribute, or vice-versa, on each trial (with equal probability). Thus, the model captures a selective attention to a randomly selected attribute on each trial. This model has one fewer parameter than the hierarchical competition model, as by definition there is no between-attribute competition parameter β. The reduction in model complexity is accounted for by the Bayesian Information Criterion (BIC) used to compare models.

To fit both this model and the hierarchical competition model to data, model parameters were estimated via a grid search across parameter space. At each point in the grid, for each of the 200 possible trials, a probability distribution of the likelihood of each response was constructed, by counting the responses of the model using an empirical simulation of 100 repetitions of each trial (In the random attribute model, as the stochasticity of attribute choice differed in each of these repetitions, this probability distribution empirically marginalises over attentional stochasticity). Parameters were selected for each subject by selecting the point in the grid that maximised the likelihood of the observed choices.

At certain points in parameter space, some actions might never be selected by the model, forcing the likelihood of that particular choice to 0. Consequently, we also included an additional 'lapse-rate' parameter in both models thereby accounting for lapses in subject performance (where they occasionally select very unlikely choices) by adding a very small probability ε to the likelihood of each option:

$$p'^{(C=o)} = \frac{p(C=o)}{1+\varepsilon} + \frac{\varepsilon}{3}$$

Note that the lapse rate is directly akin to the epsilon parameter in ε-greedy or ε-softmax action selection algorithms. In practice, the influence of this parameter was small: it was estimated to be 0 in 8/19 subjects, below 0.05 in 9/19 subjects, and below 0.1 in the remaining 2 subjects.

Predictions of 'attribute comparison' fMRI data from computational model (main Fig. 4). We simulated predictions of fMRI timeseries from the 'attribute comparison' node of the computational model (eq. 3 in modelling section, methods). Predicted activity in this node on each trial was generated under the following assumptions: (i) activity would be greatest when competition between attributes was highest, i.e. $w_i(t)$ is close to 0.5, and not close to 1; (ii) activity would accumulate across time until the decision was made by the hierarchical model. These assumptions were implemented with the following function, applied on each trial:

$$r = \sum_{t=1}^{tdec} e^{-|w_i(t)-0.5|}$$

where r is the predicted activity, and $tdec$ is the decision time of the model. The exponential function chosen is simply a useful transformation to reflect high activity when $w_i(t)$ is close to 0.5. It is shown graphically in supplementary Fig. 14. As the function is symmetric around 0.5, either attribute can be selected for i at each timepoint.

Each trial's activity was convolved with a canonical hemodynamic response function (generated in SPM12), and autocorrelated observation noise was then added. A general linear model was fit to these data, containing a constant term, the probability of the chosen and best unchosen options on relevant and irrelevant attributes, and the trial reaction time. This was repeated 10 times (to produce 10 simulated 'subjects'), and the mean +/- s.e. of parameter estimates of the regression are plotted in main Fig. 4.