

Supplementary Methods 1. Fitting procedure

We followed the Bayesian Hierarchical Drift Diffusion Models (Wiecki *et al.*, 2013) recommendations to fit our models. For each model, the starting values were set at the maximum *a posteriori* value to accelerate convergence. Bayesian inference was then performed by drawing 50,000 posterior samples by Markov Chain-Monte Carlo methods. The first 25,000 samples were discarded to limit the influence of starting values on posterior distributions. We retained every 10th sample, to reduce autocorrelation within chains. We performed 20 runs of the same model, which were then combined to generate the final model. Parameter convergence was checked before analysis, by visual examination of the trace, autocorrelation and marginal posterior distribution, and with the Gelman-Rubin R-hat statistic (Gelman and Rubin, 1992) comparing the within-chain and between-chain variances of the 20 different runs of the same model.

Gelman A, Rubin DB. Inference from Iterative Simulation Using Multiple Sequences. *Stat. Sci.* 1992; 7: 457–511.

Wiecki T V, Sofer I, Michael FJ. HDDM: Hierarchical Bayesian estimation of the Drift-Diffusion Model in Python. *Front. Neuroinform.* 2013; 7: 1–10.