Supplemental Information for

Executive function assigns value to novel goal-congruent outcomes

Samuel D. McDougle, Ian C. Ballard, Beth Baribault, Sonia J. Bishop, & Anne G.E. Collins

Figure S1

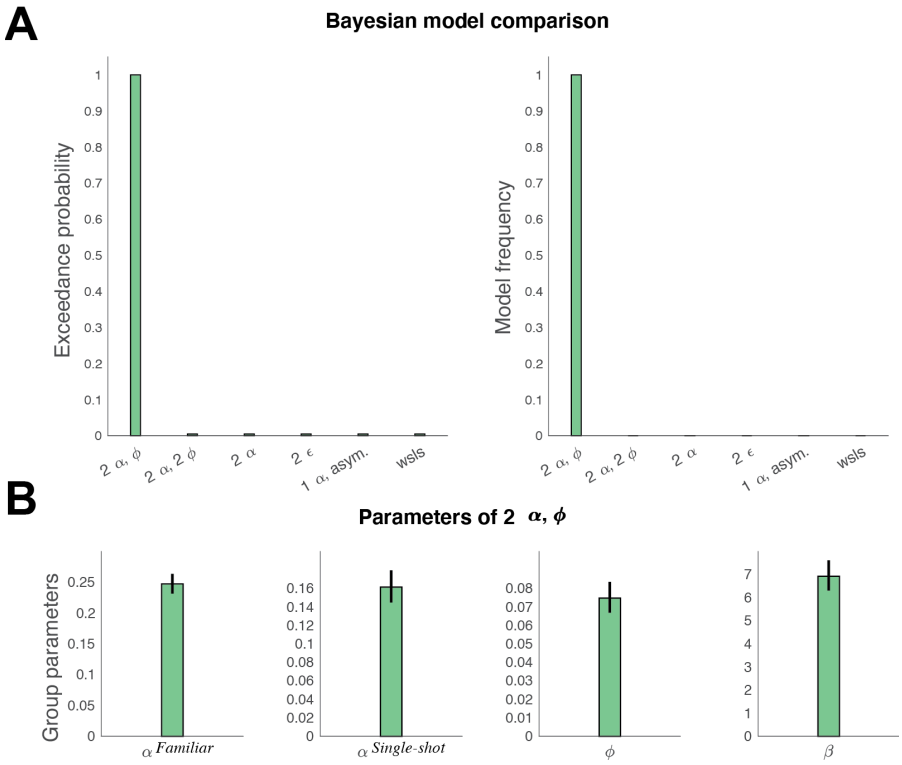

Figure S1. *Model comparisons, and parameters for the winning model.* Models were fit and compared using the Hierarchical Bayesian Estimation (Piray et al., 2019), implemented in the CBM toolbox (https://payampiray.github.io/cbm). In this method (HBI), models are compared and free parameters are estimated simultaneously. Recent work using this method has shown that fits estimated by HBI are more precise and recoverable than competing methods, and model comparison is robust and less biased towards simplicity (Piray et al., 2019). We note briefly that similar behavioral and fMRI results were obtained when we used traditional maximum likelihood estimation methods; however, clearer model comparison results and more interpretable parameter estimates were obtained in our HBI analysis. For our HBI procedure we implemented a prior variance parameter of v = 6.00. To compare learning rate parameters across feedback conditions (Figure 1F, Main Text) we could not perform standard frequentist tests due to the hierarchical procedure; thus, in a follow-up HBI analysis we implemented the Single-shot condition learning rate as a free parameter that was added or subtracted to the Familiar condition learning rate and performed an “HBI t-test” on the the resulting parameter fit relative to zero. (Further details

concerning the HBI analysis procedure, and its open-source toolbox, are given in Piray et al., 2019.) **(A)** The winning model (2 α, Φ) had two separate learning rate (α) parameters (one for each feedback condition) a single decay parameter (Φ), and a single inverse temperature (β) parameter (see *Methods*). The remaining models (in order of appearance on graph) had: two separate learning rates and decay parameters (2 α, 2 Φ), two separate learning rates but no decay parameters (2 α), a single learning rate but separate decision noise parameters for each feedback condition (2 ε), separate/asymmetric learning rates for positive and negative outcomes shared across feedback conditions (1 α, asym.), and finally a win-stay-lose-shift model with α fixed at 1 but two separate inverse temperature parameters for the conditions. For all models except the win-stay-lose-shift model, a single inverse temperature parameter was fit. **(B)** Group-level parameters results of the hierarchical fitting procedure for the winning model (2 α, Φ), with hierarchically-specified error bars (Piray et al., 2019).

Figure S2

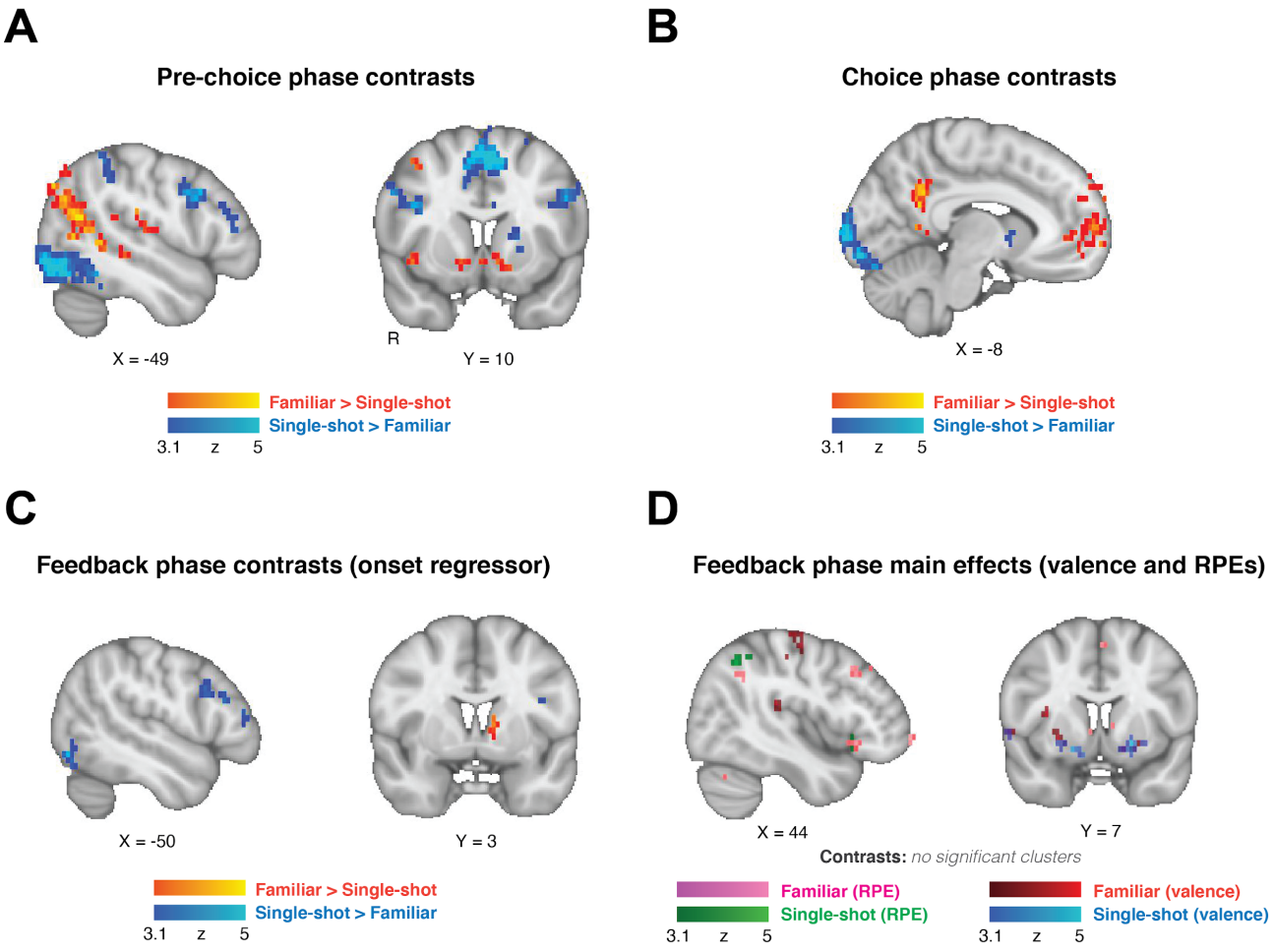

Figure S2. *Whole-brain contrasts.* Contrasts between feedback conditions (Single-shot/Familiar) for the **(A)** pre-choice phase, **(B)** choice phase, and **(C)** Feedback onset. **(D)** Main effects of valence and reward prediction error (RPE) modulators of feedback onset, separated by condition. We note that no significant clusters emerged for either the valence contrast or the RPE contrast. For all whole-brain analyses we imposed a family-wise error cluster-corrected threshold of $p < 0.05$ (FSL FLAME 1) with a cluster-forming threshold of $p < 0.001$. We note that the increased dorso-medial striatal activity seen in the feedback onset contrast (panel C, coronal) for the Familiar versus Single-shot condition is consistent with expected confounding of average reward rate with the unsigned feedback onset regressor: Indeed, the effect size of this contrast within the dorso-medial striatum was correlated, between-subjects, with behavioral differences in performance (i.e., reward rate) between the two feedback conditions (Familiar > Single-shot; Spearman's $\rho = 0.39$, $p = 0.040$).

Figure S3

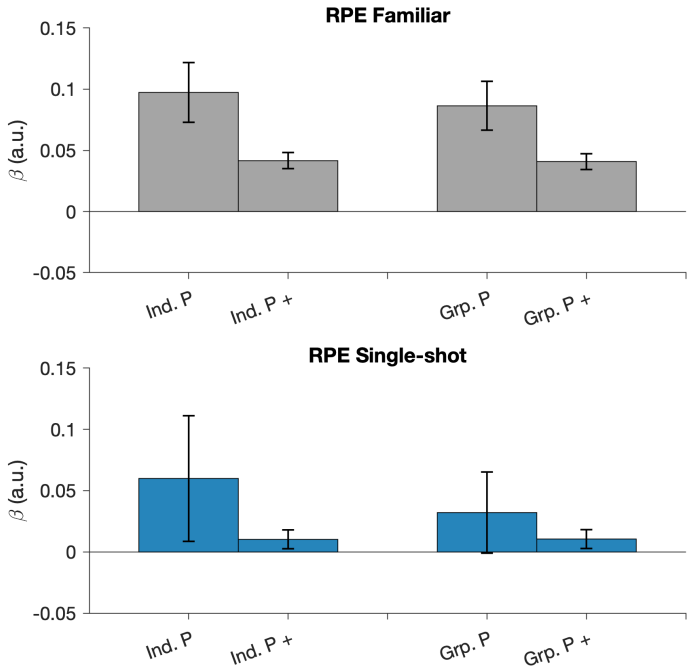

Figure S3. *Dorsal-medial striatum reward prediction error (RPE) effect sizes using alternative GLMs.* We further explored the effect of RPEs in the striatum by generating striatal ROIs in the Familiar condition (and testing effects in the held-out Single-shot condition) with RPE regressors generated either using individually fit parameters for each subject ('Ind. P', same as Figure 3 in main text), individually fit parameters for each subject with separate regressors for positive versus negative RPEs ('Ind. P +'; positive RPE effects are depicted), group-level parameters applied uniformly to each subject's behavior to generate RPE regressors ('Grp. P'), and group-level parameters applied uniformly with separate regressors for positive versus negative RPEs ('Grp P +'; positive RPE effects are depicted). All striatal RPE effects in the Familiar condition were significant (all *p*s < 0.01), but none were in the Single-shot condition (all *p*s > 0.18).

Figure S4

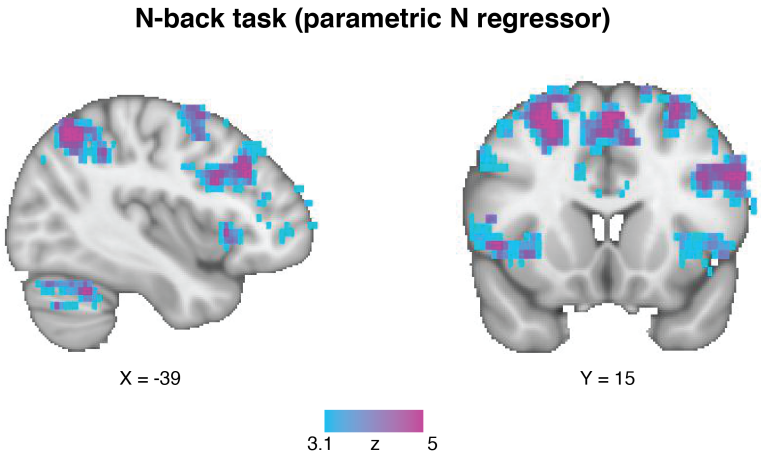

Figure S4. *N-back task whole-brain main effect of load.* Group-level block-designed main effect of *N* (*Ns* used: 1, 2, 3), using a family-wise error cluster-corrected threshold of $p < 0.05$ (FSL FLAME 1) with a cluster-forming threshold of $p < 0.001$.

Figure S5

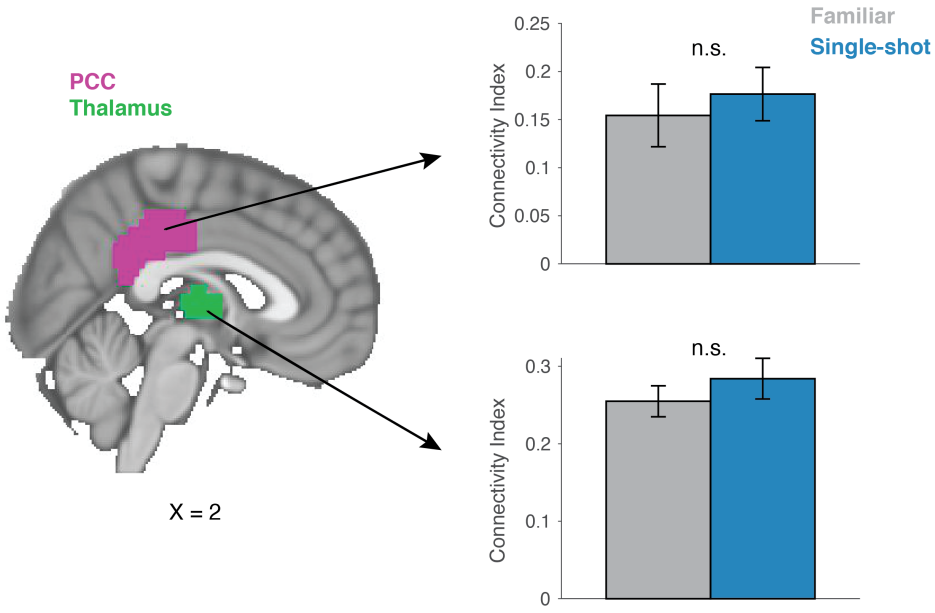

Figure S5. *Control connectivity analyses in two anatomical ROIs.* We conducted a beta series correlation analysis (see *Methods*) between PFC activity at pre-choice and feedback-locked activity (on successful trials) in two additional regions, to control for global activity effects not specifically related to activations in our task-based ROIs (Figure 5). We chose two control anatomical ROIs in regions (posterior cingulate cortex, pink; thalamus, green) shown to be active in reinforcement learning tasks but about which we had no *a priori* hypotheses. Bilateral ROIs were defined anatomically on the MNI brain using the Harvard-Oxford atlas (threshold: 0.50).

Figure S6

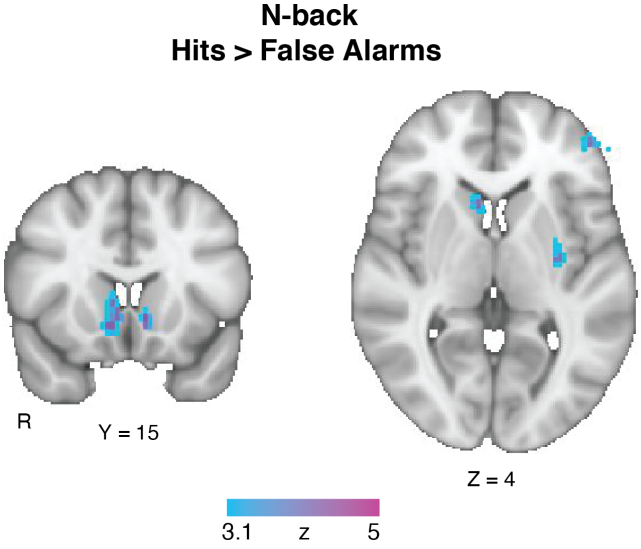

Figure S6. *Event-related striatal activation during n-back task performance for correct trials > false alarm trials.* This analysis acts as a replication of Satterthwaite et al., 2012, showing striatal activation for correct trials despite there being no explicit performance feedback. Depicted activations reflect the group-level event-related contrast of hit trials > false alarms, collapsed over all levels of *N*. We note that both types of trials require a button press, thus controlling for movement confounds. The GLM used in this analysis was identical to that described in the *Methods* section of the main text (GLM 5), with additional spike regressors for each type of event (i.e., hits, false alarms, correct rejections, and misses; convolved with the canonical double-gamma HRF). As in all of our other whole-brain analyses, we used a family-wise error cluster-corrected threshold of $p < 0.05$ (FSL FLAME 1) with a cluster-forming threshold of $p < 0.001$.

Table S1

Model	Parameter					
	α	*α+*	*α-*	*Φ*	*ε*	*β*
2 α, Φ	Fam, SS	-	-	Single	-	Single
2 α, 2 Φ	Fam, SS	-	-	Fam, SS	-	Single
2 α	Fam, SS	-	-	-	-	Single
2 ε	Single	-	-	-	Fam, SS	Single
1 α, asym.	-	Single	Single	-	-	Single
wsls	Fixed at 1	-	-	-	-	Fam, SS

Table S1. *Model specifications.* This table above describes the free parameters included in each of the models compared in our hierarchical fitting procedure. 'Single' = single parameter used for both the Familiar and Single-shot feedback conditions; Fam, SS = unique parameters used for each feedback condition; α = learning rate; α+ = learning rate for successful outcomes; α- = learning rate for unsuccessful outcomes; Φ decay parameter; ε = decision noise; β inverse temperature. See *Methods* for further details.