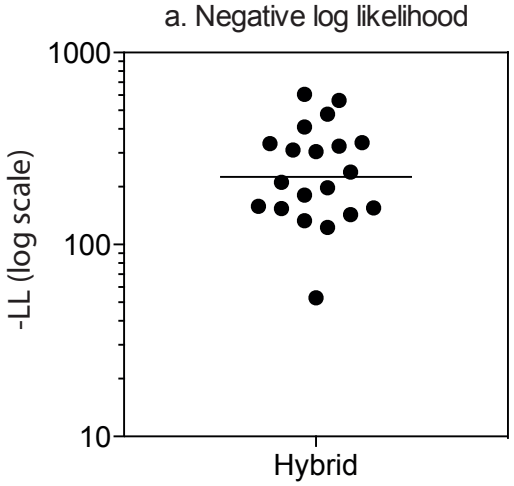
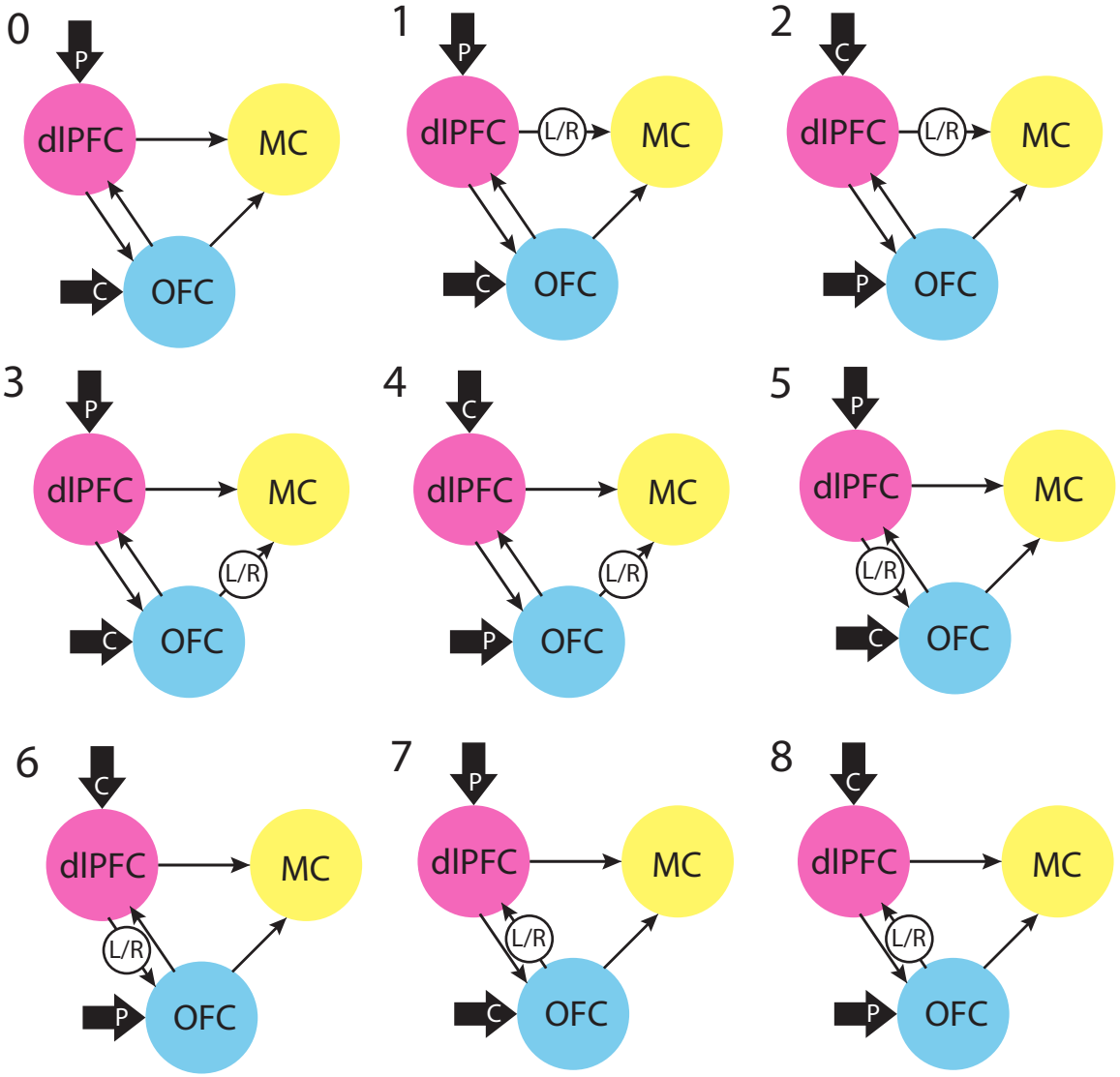
Supplementary Figure 1. Individual Fits



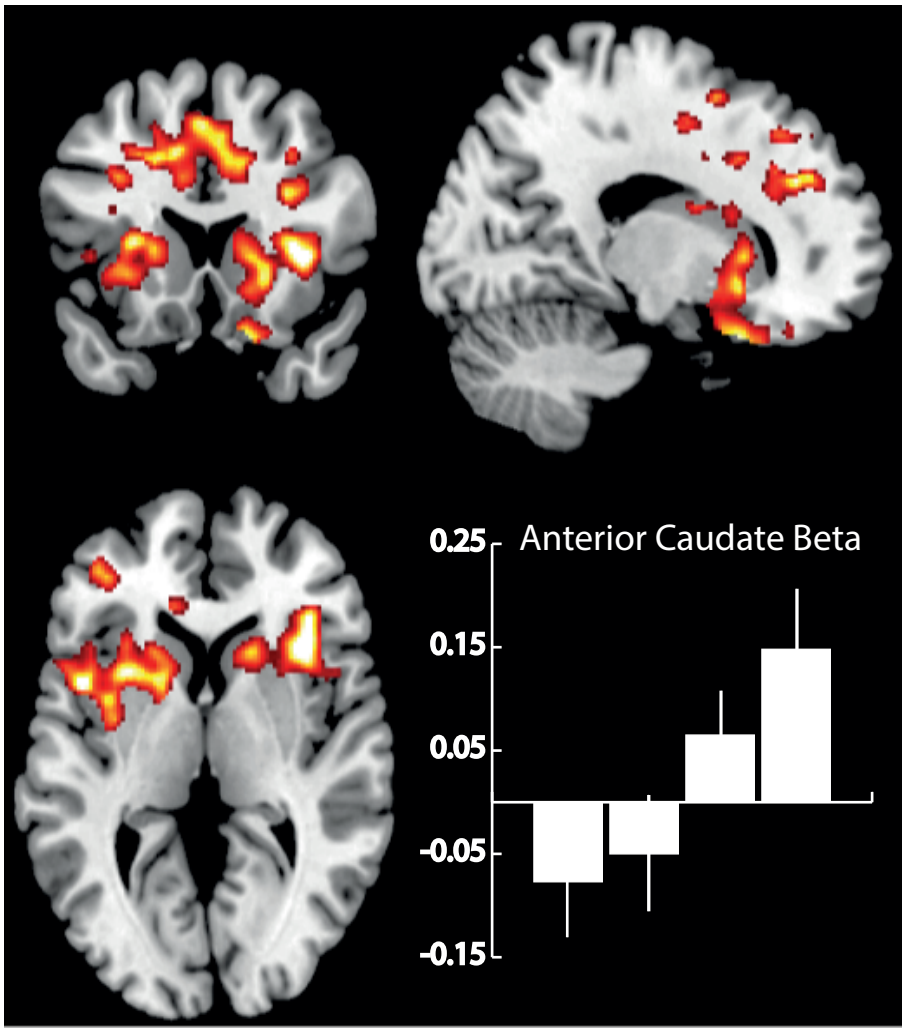
Negative log likelihood (-LL) of the best fit by maximum likelihood for each individual from the hybrid model

Supplementary Figure 2. Models



Dynamic causal models showing regions-of-interest (dlPFC: dorsolateral prefrontal cortex; OFC: medial orbitofrontal cortex; MC: motor cortex), driving inputs (P: relative advantage values; C: chosen action values), and modulatory coupling in an action-specific manner (L/R).

Supplementary Figure 3. Caudate correlations



Regions tracking the local correlation between response rate and reward. Bar chart (above) shows beta values extracted from an anatomical ROI in the anterior caudate linearly increased with the local correlation.

Supplementary Table 1. Estimated parameters and pseudo r^2 for each subject.

Subject	R^2	α	κ	τ_p	τ_q
01	0.612	0.000	2.312	1.769	9.717
02	0.824	0.022	3.535	1.041	6.669
03	0.785	0.025	3.320	0.788	6.392
04	0.807	0.099	3.303	0.649	6.786
05	0.317	0.000	1.453	0.693	3.432
06	0.606	0.001	2.398	0.962	53.059
07	0.820	0.000	3.659	1.741	100.000
08	0.719	0.000	3.012	0.707	2.082
09	0.188	0.001	1.059	0.241	100.000
10	0.698	0.122	2.584	2.130	7.822
11	0.835	0.104	3.554	0.979	5.799
12	0.901	0.106	1.998	0.000	47.204
13	0.663	0.000	2.704	1.241	100.000
14	0.318	0.106	1.477	0.097	1.844
15	0.468	0.072	1.921	0.515	3.389
16	0.341	0.011	-2.088	2.758	67.652
17	0.008	0.029	-0.117	0.375	4.217
18	0.506	0.000	2.028	0.826	14.521
19	0.738	0.052	2.902	1.065	9.352
20	0.796	0.067	3.438	1.972	9.157
Median	0.681	0.050	2.754	0.894	8.489

Pseudo R^2 and individual best fitting model parameters (α, κ and τ) for the hybrid model.

Supplementary Table 2. Pearson correlations (r) between *P* and *D* and π signals

Subject	*P* * *D*	*P* * π
001	0.864	0.520
002	0.340	0.093
003	0.237	-0.050
004	0.077	-0.023
005	0.810	0.287
006	0.661	0.270
007	0.626	0.251
008	0.625	0.126
009	0.749	0.057
010	0.555	0.363
011	0.240	0.087
012	-0.495	-0.621
013	0.762	0.266
014	0.590	0.061
015	0.618	0.187
016	0.831	0.284
017	0.854	0.187
018	0.770	0.238
019	0.076	-0.045
020	-0.016	-0.039

Supplementary Table 3. Recovered *Q*-learning parameters

simulation	α	τ_q	κ
1	0.014	14.256	1.823
2	0.016	13.073	1.814
3	0.016	12.943	1.766
4	0.015	12.979	1.789
5	0.012	15.151	1.821
6	0.012	15.151	1.821
7	0.017	12.949	1.843
8	0.016	12.952	1.793
9	0.015	13.625	1.784
10	0.013	14.357	1.804
11	0.018	11.794	1.833
12	0.015	14.083	1.782
13	0.017	12.455	1.826
14	0.013	14.711	1.797
15	0.011	16.056	1.842
Mean	0.015	13.769	1.809
SD	0.002	1.177	0.023

Mean recovered parameters from the Q-learning model derived from 15 simulations using the same parameters to generate *D* signals for the fMRI analysis (α = 0.015, τ_q = 13.407, and κ = 1.811). The results show the parameters can be recovered accurately from choice data. Note the *P* signal was not dependent on the parameters and so no simulation was performed.