

Fig. 6. Histogram showing the distribution of reward prediction error (PE) values. The two outliers (PE values less than -\$20) were excluded from analyses presented in the text.

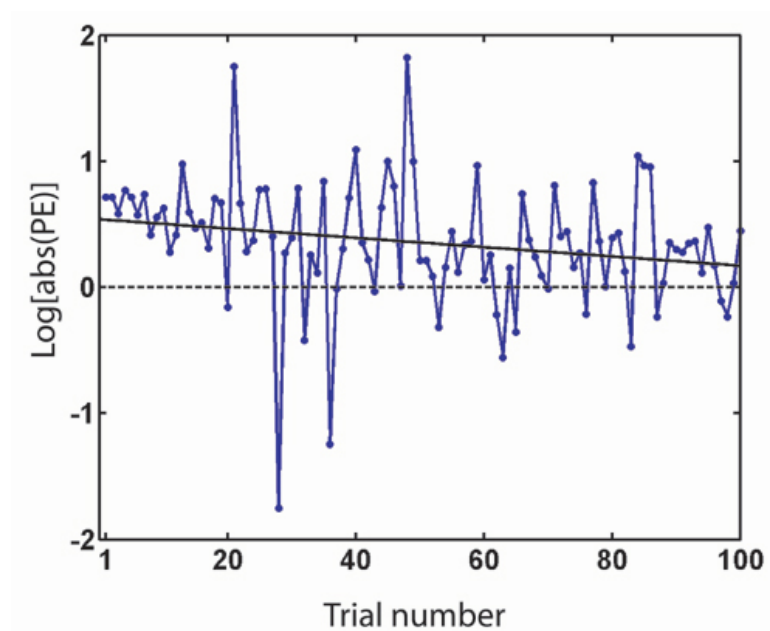
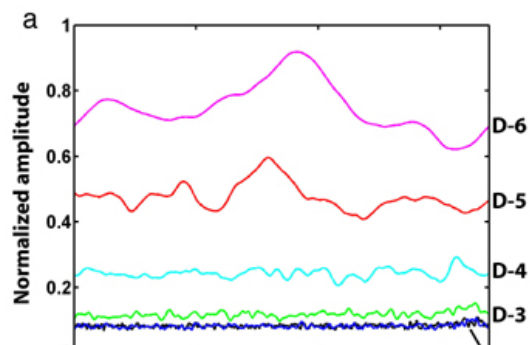


Fig. 7. Time course of reward PE. When the logarithm of the absolute value of PE is plotted against trial number, a significant regression was found ($r = -0.217$, $P = 0.03$, $n = 100$), indicating learning by the patient over time.



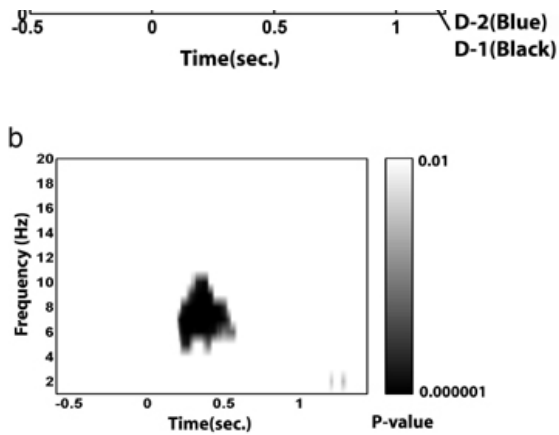


Fig. 8. Amplitudes and phase-locking values recorded at channel 1. (*a*) Mean instantaneous amplitude of field potentials on channel 1 ($n = 91$). D-1 to D-6 correspond to decomposition levels of the discrete wavelet transform. 0 on the x axis represents the time of punishment delivery (or "Wait" cue delivery in nonpunished trials). Roughly, D-1 to D-3 correspond to gamma, D-4 corresponds to beta, D-5 corresponds to alpha, and D-6 corresponds to theta frequency bands. Note the clear increase of amplitude in alpha and theta bands at 200-600 msec. (*B*) Phase-locking values (PLV) of field potentials on channel 1 ($n = 91$). Significant phase concentration occurred in the alpha and theta frequency range from 200-500 msec. For this calculation, zero phase-shift windowed FIR (finite impulse response) band-pass filters of order 100 with center frequency 1 to 20 Hz (1-Hz steps) were applied to the ERP data.

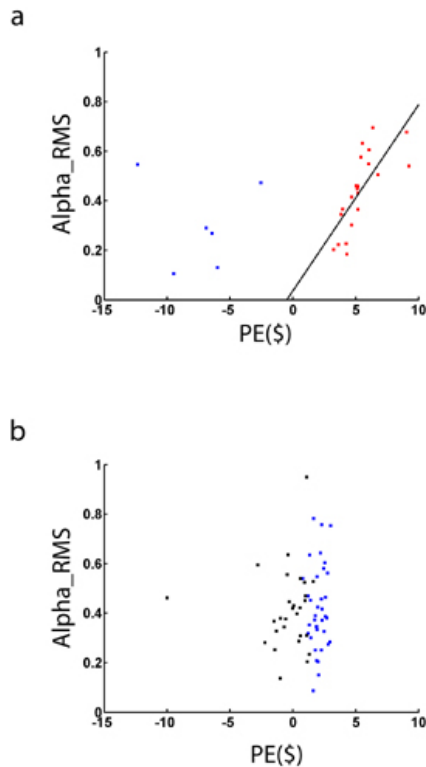


Fig. 9. Relationship between prediction error (PE) and alpha-band ERP amplitude. (*a*) For trials corresponding to choices from risky decks, there was a correlation between PE and ERP amplitude only for that subset in which no punishment was actually obtained (red), but not that subset in which punishment was given (blue). (*b*) PE versus alpha-band ERP amplitude for choices from safe decks, for that subset in which no punishment was given (blue) and that subset in which punishment was given (black); neither showed a significant correlation.



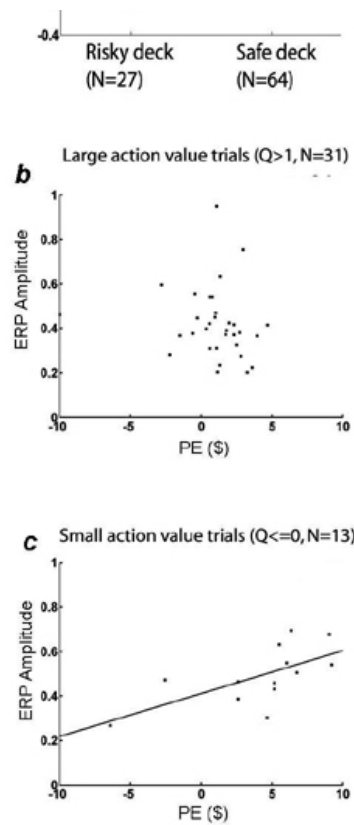


Fig. 10. Analyses of action values (Q). (a) Action values were significantly larger for choices from the risky decks than for choices from the safe decks [$t(89) = 6.2$, $P < 0.001$, $n = 91$]. Error bars are ± 1 SEM. Relationship between PE and alpha-band ERP amplitude (alpha rms value) for those trials that showed large Q values (b) did not correlate, whereas PE and ERP amplitude were correlated for those trials that showed negative Q values (c).

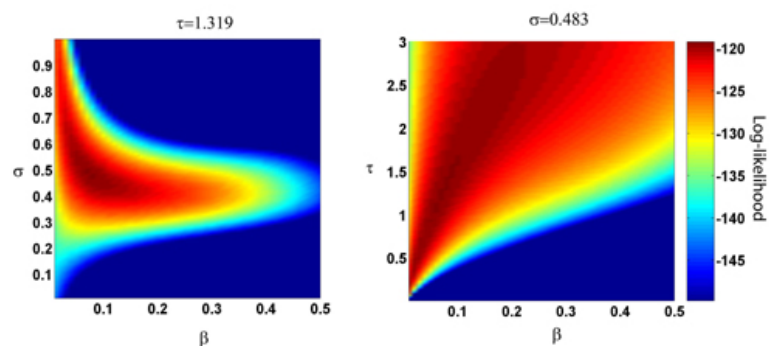


Fig. 11. Log-likelihood maps for various parameter combinations, verifying that a global minimum was achieved. We held either τ or σ fixed and plot the values of the other two parameters. τ , temperature parameter; β , step-size parameter; and σ , sensitivity to punishment. Color encodes log-likelihood of the model.