Figures and figure supplements

Lateral orbitofrontal cortex promotes trial-by-trial learning of risky, but not spatial, biases

Christine M Constantinople et al
Figure 1. Behavioral task: Rats performing the task exhibit stable performance over months, but also trial-by-trial learning dynamics. (A) Example trial: rat initiates a trial by nose-poking and fixating in center. On each side, light flashes and click rates convey reward probability and water volume, respectively. One side (here, the right port) offers guaranteed reward (“safe”); safe and risky sides vary randomly over trials. (B) Relationship between flashes and probability, and click rates and reward volumes (6, 12, 24, or 48 µL) in one version of the task. Risky side could have rewarded probability between 0–1 (increments of 0.1). (C) Offered reward volumes and probabilities. (D) Behavioral performance in units of ‘efficiency’ for five representative rats in the final training stage (Materials and methods). We compared the average expected value (reward x probability) per trial the rat received compared to an agent choosing randomly, or one that always chose the option with the greater expected value (‘ideal performance’). The dashed line is criterion performance for each rat (see ‘Materials and methods’). (E) Percent of trials one rat chose the safe option for each of the four safe volumes. Axes show probability and volume of risky alternatives. (F) Difference in probability of choosing the safe option following guaranteed rewards and risky rewards (relative to the mean probability of choosing safe) for all rats (black is mean). Rats were more likely to gamble following risky rewards (p=8.35e-16, paired t-test). (G) The magnitude of the risky win-stay bias exhibits graded dependence on the reward probability of the gamble (mean across rats). p=0.0035 of slope parameter of least-squares regression line (dashed line). The riskier the gamble that won, the more likely that rats will choose to gamble again. See also Figure 1—figure supplement 1. (H) Change in the probability of repeating left or right choices following rewarded or unrewarded trials. Asterisks indicate that rats’ ‘win-stay’ biases were significantly different from zero (p=2.06e-13, paired t-test), as were their ‘lose-switch’ biases (p=2.65e-15).

DOI: https://doi.org/10.7554/eLife.49744.002
Figure 1—figure supplement 1. Supplemental behavioral analyses. (A) Distribution of inter-trial intervals (ITIs) for three representative rats. Trials were self-paced; rats were free to initiate trials within 100–200 ms of the preceding trial. If rats terminated the trial early by breaking center fixation, they were penalized with a time-out penalty (those trials are not shown). (B) Average number of trials per session for each rat, excluding trials that were terminated prematurely. Mean of this distribution (368 trials/session) is shown by the red arrow. (C) Mean behavioral performance across rats, including >2.5 million trials. Percent of trials all rats chose the safe option for each of the four safe side volumes. Axes show the probability and volume of risky alternatives. Mean performance across 36 rats (normalized to max before averaging). (D) Estimates of conditional probabilities in finite sequential data can have small biases (Miller and Sanjurjo, 2015). If this bias were driving sequential effects in our data, such as increased willingness to take risks following risky wins, we reasoned that computing this bias from random flips (of the same length as our data) of a weighted coin would also reveal an effect. Therefore, we generated random choices for each rat with a generative probability corresponding to the mean probability of choosing the safe option for that rat. We then calculated the change in the probability of choosing the safe option based on reward history for the simulated choices; the same number of trials that were used in Figure 1F were applied to this analysis. There was no observable risky win-stay bias in the simulated dataset, indicating that the effect we observed did not reflect biased estimates of conditional probabilities. (E) Difference in probability of choosing the safe option following guaranteed rewards and risky rewards of different probabilities (relative to the mean probability of choosing safe) for simulated data, as in B. Randomly simulated choices with the same sample sizes as the data (Figure 1G) did not exhibit a bias for risky choices with a graded dependence on reward probability. p=0.80 of slope parameter of least-squares regression line (dashed line). Therefore, the risky win-stay bias we observe, with graded dependence on reward probability, does not reflect biased estimation of conditional probabilities. (F) Difference in probability of choosing the safe option following guaranteed rewards, or risky unrewarded choices. There was no systematic, significant change in probability of choosing safe following unrewarded trials (paired t-test comparing change in probability of choosing safe). DOI: https://doi.org/10.7554/eLife.49744.003
Figure 2. IOFC encodes reward history during the cue period. (A) IOFC neuron with activity aligned to trial initiation. This neuron's firing rate reflected whether the previous trial was rewarded. (B) Mean encoding of reward history (discriminability or d') across IOFC neurons that exhibited significantly different spike counts based on reward history. Mean ± s.e.m. See also Figure 2—figure supplement 1. (C) Fraction of neurons with significantly different spike counts based on reward history, with more spikes following unrewarded (no rew > rew) or rewarded (rew > no rew) trials. (D) Schematic of analysis (TCA/PARAFAC) used to discover low dimensional descriptions of trial-by-trial population dynamics. See also Figure 2—figure supplement 2. (E) Result of TCA/PARAFAC from one recording session. Y-axis is in arbitrary units (A.U.; see Materials and methods). (F) Mean (± s.d.) shuffle-corrected reward (blue) and no-reward (black) triggered averages of trial factors across all sessions (see Materials and methods). (G) Correlation between trial factors and reward history for each session. Gray bars indicate significance.

DOI: https://doi.org/10.7554/eLife.49744.004
Figure 2—figure supplement 1. Method for identifying putatively identical waveforms over days. (A) Distribution of $d_1$ and $d_2$ values comparing waveforms across rats produces a null distribution (gray, see Materials and methods). Distribution of values comparing waveforms within rats from subsequent recording sessions (red). Dashed lines are empirically chosen thresholds. (B) Subplot from panel A. (C) Neuron that was identified as putatively identical across four recording sessions. Raster plots only show 150 trials (out of 400–600 each day) for display purposes (upper panels). PSTHs (derived from all trials) are shown below (lower panels). (D) Figure 2B was reproduced combining putatively identical units recorded over multiple days. Mean discriminability index ($d'$) depending on whether the previous trial was rewarded or not, computed in 50 ms bins. Error bars are ± s.e.m.

DOI: https://doi.org/10.7554/eLife.49744.005
Figure 2—figure supplement 2. TCA/PARAFAC tensor decomposition applied to neural data. (A–C) Method used to determine model rank. We performed 20 random initializations and compared the similarity of the factors recovered from each iteration to those recovered from the previous one. We show the distribution of similarity indices for example recording sessions that were determined to be rank 1, 2, and 3 (A, B, C, respectively). Black lines are mean ± s.e.m. The majority of the data was either rank 1 or 2 (50/105 sessions were rank 1, 50/105 were rank 2, 5/105 were rank 3), so for simplicity, we fit a rank one model to each session. (D) Four neurons from the recording session in panel 3D; firing rates are plotted when the trial factor was high (>85th percentile) or low (<15th percentile). (E) Mean (± s.d.) shuffle-corrected reward (blue) and no-reward (black) triggered averages of trial factors across all sessions (see Materials and methods), excluding cells that had significantly different spike counts following rewarded or unrewarded trials. (F) Mean (± s.d.) shuffle-corrected reward (blue) and no-reward (black) triggered averages of trial factors, excluding random subsets of cells, of the same number that were excluded in panel E. (G) Distribution of simultaneously recorded units across all recording sessions. (H) Relationship between the Pearson’s correlation between trial factors and reward history, and number of units recorded in each session. (I) Relationship between the absolute magnitude of the Pearson’s correlation between trial factors and reward history, and number of units recorded in each session.

DOI: https://doi.org/10.7554/eLife.49744.006
Figure 3. Optogenetic perturbation of IOFC during the cue period does not affect spatial or risky trial history biases. (A) Schematic of bilateral optogenetic perturbations. For CaMKIIa-eNpHR3.0 rats (n = 8), we used continuous illumination of a green laser for photoinhibition. For Pvalb-iCre-ChR2 rats (n = 5), a blue laser was pulsed at 20 Hz. See also Figure 3—figure supplement 1. While the schematic shows a 3 s trial, trial durations were variable (2.6–3.35 s); photoinhibition persisted for the duration of the cue period. (B) Histological section from Pvalb-iCre-ChR2 rats also stained for DAPI and parvalbumin (PV) immunoreactivity. (C) Virus injection in a wild type rat expressing CaMKIIa-eNpHR3.0. Location of fibers were estimated by damage at brain surface and fiber tracks. (D) Magnitude of spatial win-stay and lose-switch biases (difference in probability of repeating a left or right choice) on control and laser trials. Error bars are normal approximation of 95% confidence intervals (Materials and methods). (E) Magnitude of risky win-stay bias (difference in probability of choosing the safe option following safe or risky rewards) on control and laser trials.

DOI: https://doi.org/10.7554/eLife.49744.007
Figure 3—figure supplement 1. Characterization of photoinhibition in Pvalb-iCre-ChR2 rats. (A) Representative unit recorded from Pvalb-iCre rats expressing ChR2. Three epochs of photoinhibition (blue lines) reliably suppressed spiking activity. Blue laser was pulsed at 20 Hz, 10 ms pulse width, for 8 s. (B) Normalized activity of the cell shown in A; mean ± s.e.m. over 30 photoinhibition epochs. (C) Mean suppression over 65 recorded units from two rats. (D) Activity change of each unit plotted as a function of its distance from the optical fiber. Units were recorded in four tracks, 250, 500, 750, or 1000 μm from the fiber tip. Robust photoinhibition was observed in all tracks. (E) Example injection site shown in Figure 4C, inset shows putative fiber track. (F) Percent of parvalbumin-immunoreactive cells that co-expressed eYFP in Pvalb-iCre rats expressing eYFP-ChR2 (left), and fraction of eYFP-expressing cells co-labeled for parvalbumin immunoreactivity (right).

DOI: https://doi.org/10.7554/eLife.49744.008
Figure 4. At time of choice report, IOFC neurons represent risk, reward, and left/right choice. (A) Example IOFC neuron with activity aligned to when the rat left the center poke to report his choice. This neuron’s firing rate reflected whether the rat chose the risky (magenta) or safe (black) option on the current trial, analyzing rewarded trials only. (B) Mean $d'$ across IOFC neurons with significantly different spike counts on trials with risky or safe choices. See also Figure 4—figure supplement 1. (C,D) Mean z-scored firing rate of neurons in panel B aligned to entering the center poke (C), or leaving it to report choice (D). (E) Fraction of neurons in panels B-D that preferred trials when rats made risky or safe choices. Higher firing rates on trials in which rats chose the safe reward could reflect encoding of decision confidence or reward expectation (Lak et al., 2014). (F) Mean $d'$ reflecting whether rats chose the left/right ports, or whether rats received reward, averaged across neurons with significantly different spike counts on those trials. See also Figure 4—figure supplement 1. (G) Venn diagram of overlap between neurons whose activity differentiated between left/right choices and rewarded/unrewarded trials. (H) Fraction of neurons in panels F,G preferring left/right choices or rewarded/unrewarded trials.

DOI: https://doi.org/10.7554/eLife.49744.009
Figure 4—figure supplement 1. Results do not depend on whether units are treated independently over days. (A, B) Figure 4B (A) and 4F (B) were reproduced combining putatively identical units recorded over multiple days. Mean discriminability index ($d'$) depending on whether the rat chose the safe or risky option on rewarded trials only (A), chose left or right (B, purple), or was rewarded (B, yellow), computed in 50 ms bins. Error bars are ± s.e. m. (C) Of the units with significantly different spike counts on trials in which rats chose risky or safe, the fraction selective (or not) for choosing the left or right port.

DOI: https://doi.org/10.7554/eLife.49744.010
Figure 5. Photoinhibition of IOFC at the time of choice report selectively eliminates the risky win-stay bias. (A) For choice reporting period perturbations, the laser was triggered when rats left the center poke, and persisted for 4 s into the inter-trial-interval. See also Figure 5—figure supplement 1. (B) Spatial win-stay/lose-switch biases following photoinhibition during the choice reporting period; sham rats also exhibited a significant reduction in lose-switch biases, and trended towards a reduction in win-stay biases. Control data are replotted from Figure 3D. (C) Magnitude of the risky win-stay bias following choice reporting period inactivations. Control data are replotted from Figure 3E. Error bars are 95% confidence intervals. DOI: https://doi.org/10.7554/eLife.49744.011
Figure 5—figure supplement 1. Photoinhibition during the choice reporting period does not affect baseline performance, but selectively reduces the risky win–stay bias. (A) Psychometric performance for each CaMKIIa-eNpHR3.0 rat on control trials (black) and trials following photoinhibition during the choice period. These plots include all trials, regardless of trial history, so the elimination of the risky win-stay bias is not evident. (B) Psychometric performance for each Pvalb-iCre-ChR2 rat on control trials (black) and trials following photoinhibition during the choice period. (C) Difference in logistic regression coefficients (control - photoinhibition) parameterizing different choice biases. Data are mean ± standard deviation across rats. Asterisks indicate significant Bonferroni-corrected p-value from one-way ANOVA (p=0.0063).

DOI: https://doi.org/10.7554/eLife.49744.012