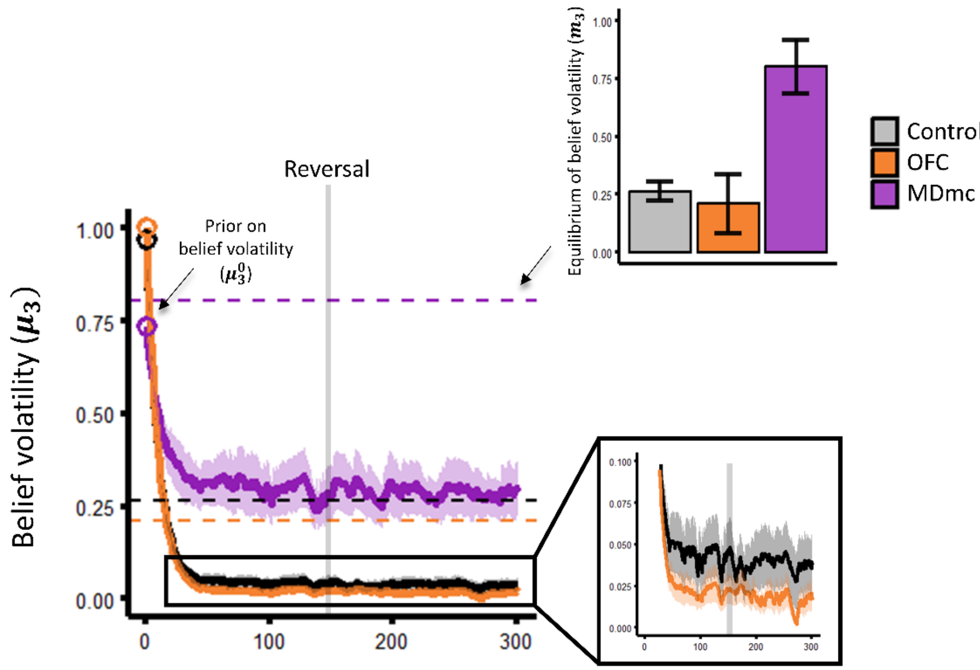
Cell Reports, Volume *43*

Supplemental information

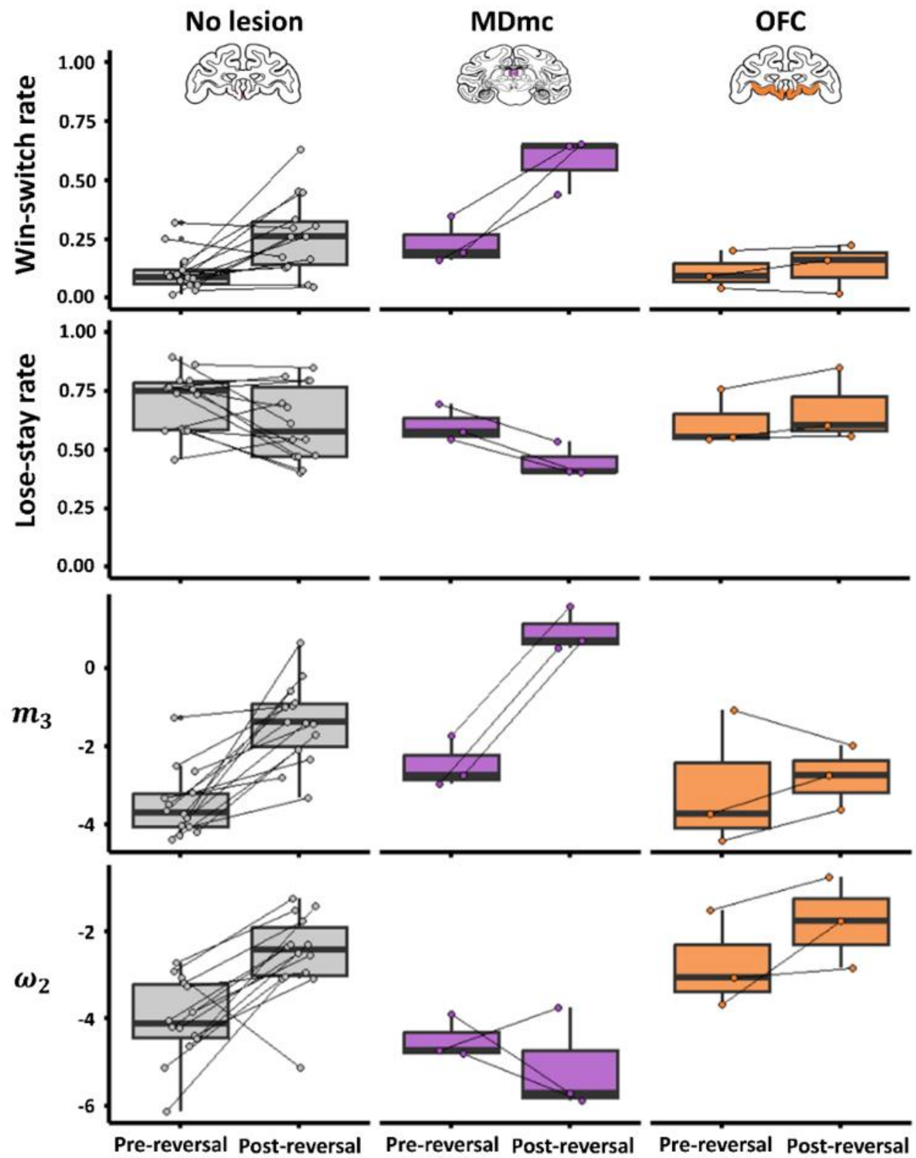
Lesions to the mediodorsal thalamus, but not orbitofrontal cortex, enhance volatility beliefs linked to paranoia

Praveen Suthaharan, Summer L. Thompson, Rosa A. Rossi-Goldthorpe, Peter H. Rudebeck, Mark E. Walton, Subhojit Chakraborty, Maryann P. Noonan, Vincent D. Costa, Elisabeth A. Murray, Christoph D. Mathys, Stephanie M. Groman, Anna S. Mitchell, Jane R. Taylor, Philip R. Corlett, and Steve W.C. Chang

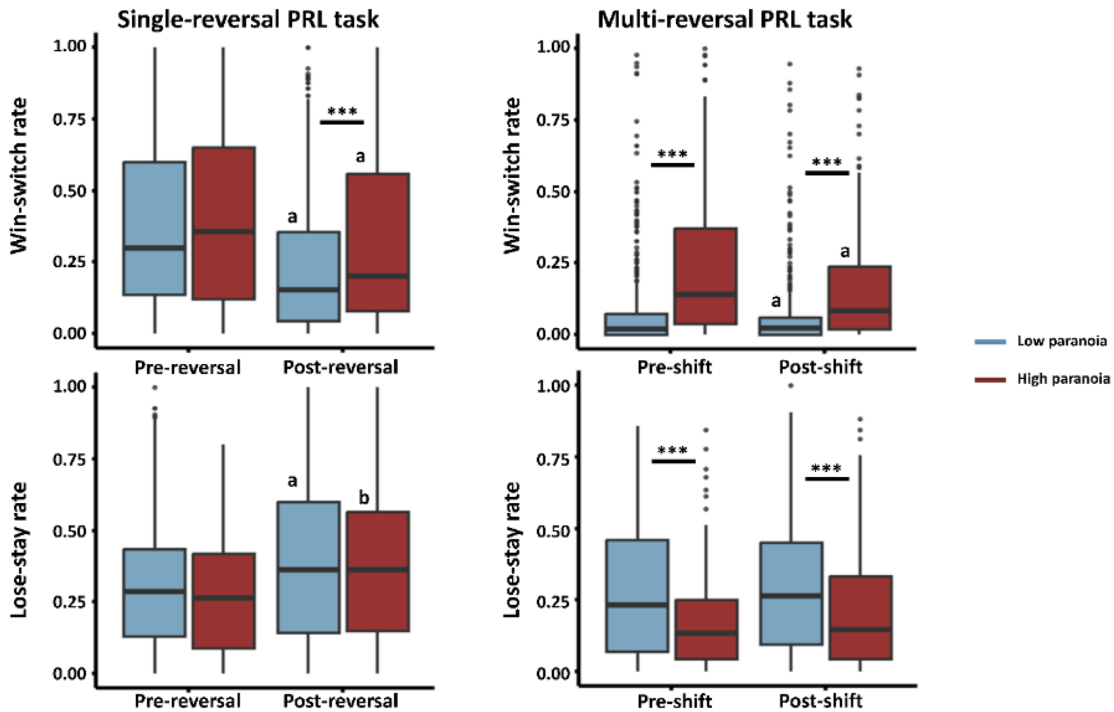
Supplementary Figures



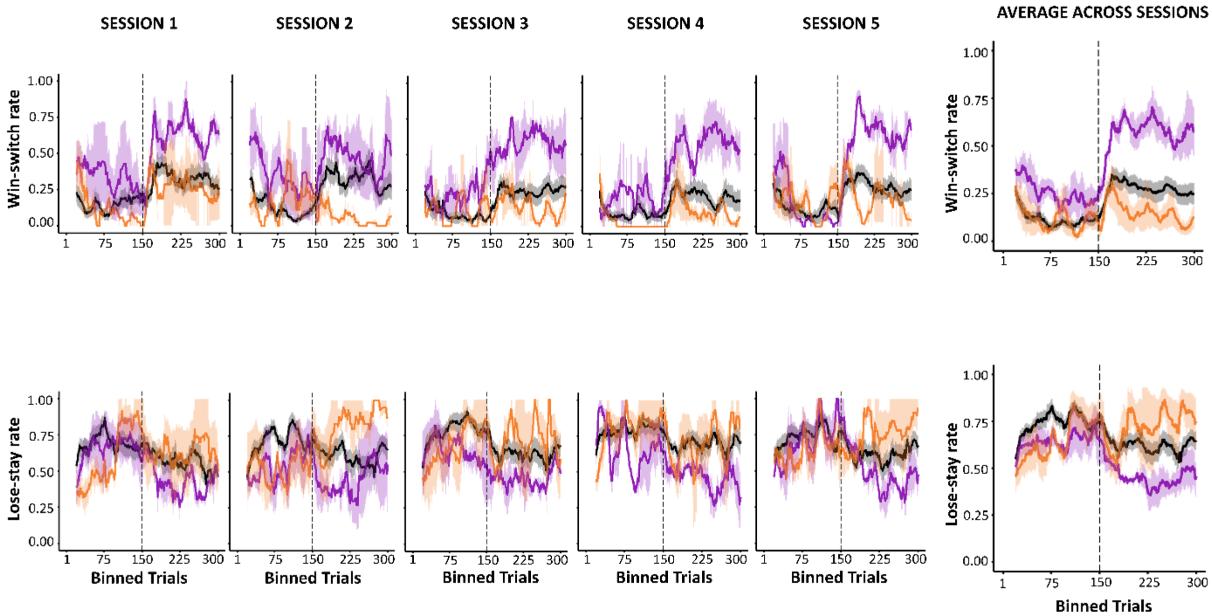
Supplementary Figure S1. An intuition of the relationship between components of belief volatility. We illustrate lesion-specific effects on monkey's perception of changes in the environment (μ_3). We observe monkeys with lesions to the MDmc have lower (p = 0.08) initial anticipation for changes in the environment (μ_3^0) relative to the control and OFC-lesioned monkeys. However, while the control and OFC-lesioned monkeys quickly become more certain of the environment, MDmc-lesioned monkeys persist and stabilize with greater uncertainty, driven largely by their equilibrium state of their belief on environmental volatility (m_3).



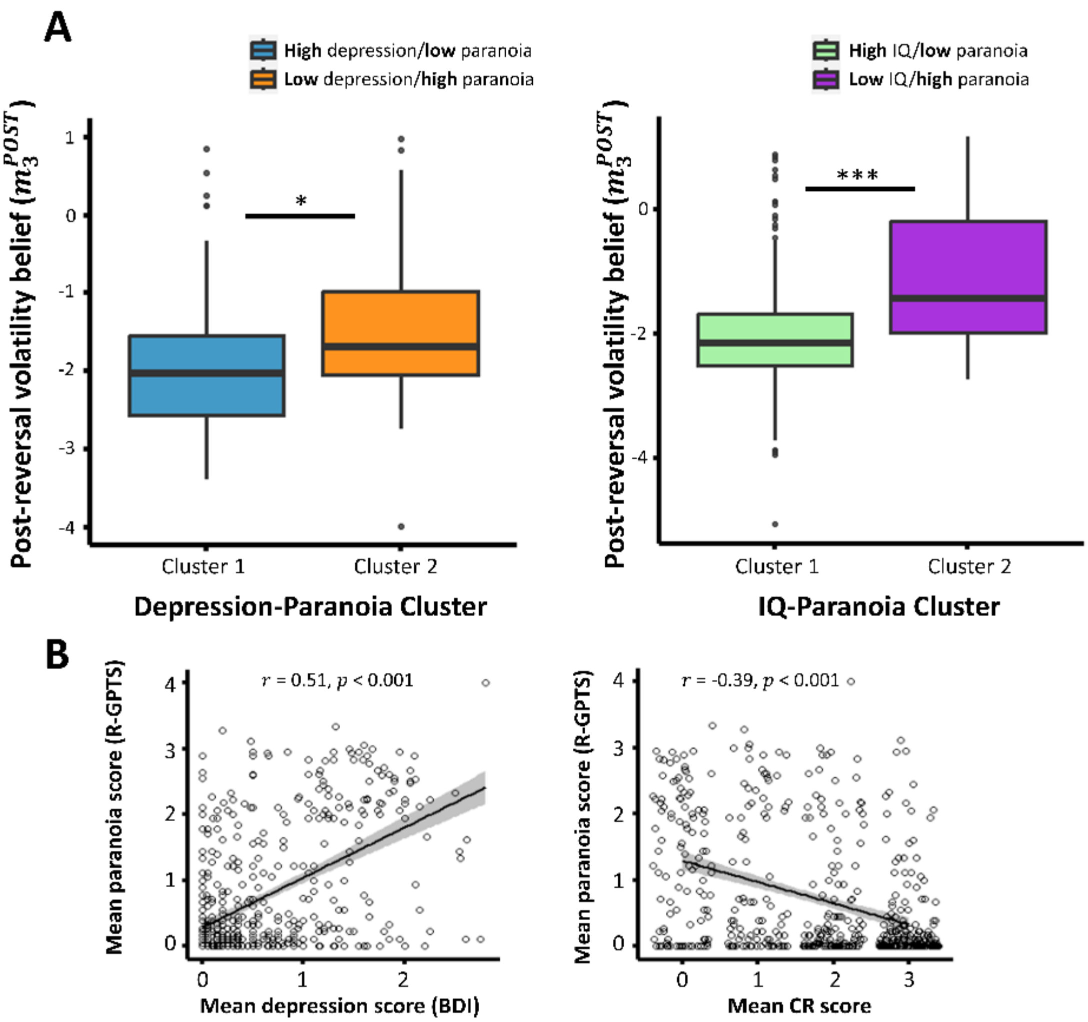
Supplementary Figure S2. Single-subject comparison of PRL task behavior and belief-updating across lesion groups. We show, with connecting lines, the pre- and post-reversal change in behavior and beliefs for every individual monkey.



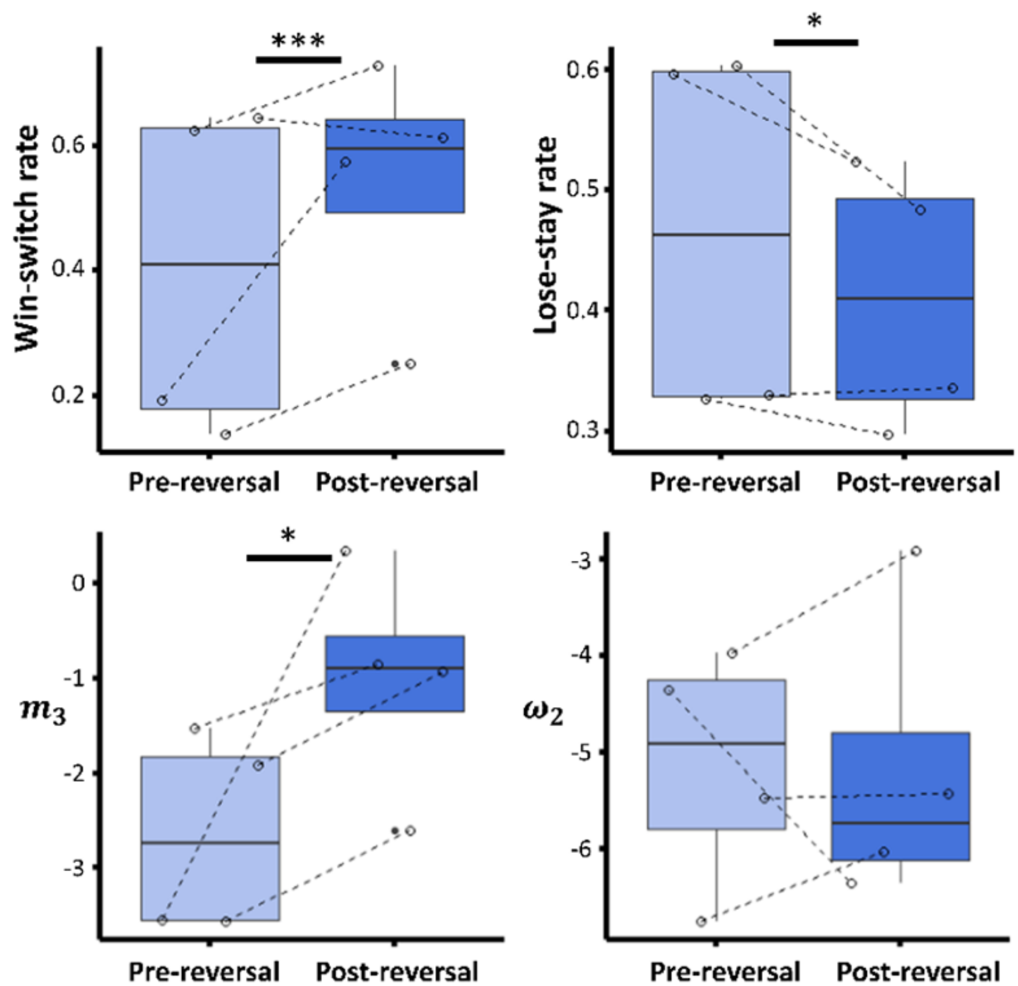
Supplementary Figure S3. Paranoia group differences in PRL task behavior between single- and multi-reversal versions. The multi-reversal version shows more sensitivity to differences in switch and stay behavior between paranoia groups. Just as we observed greater win-switching and fewer lose-stay behavior in MDmc-lesioned monkeys, individuals with paranoia show similar behavior more prominently in the multi-reversal version than the single-reversal version.



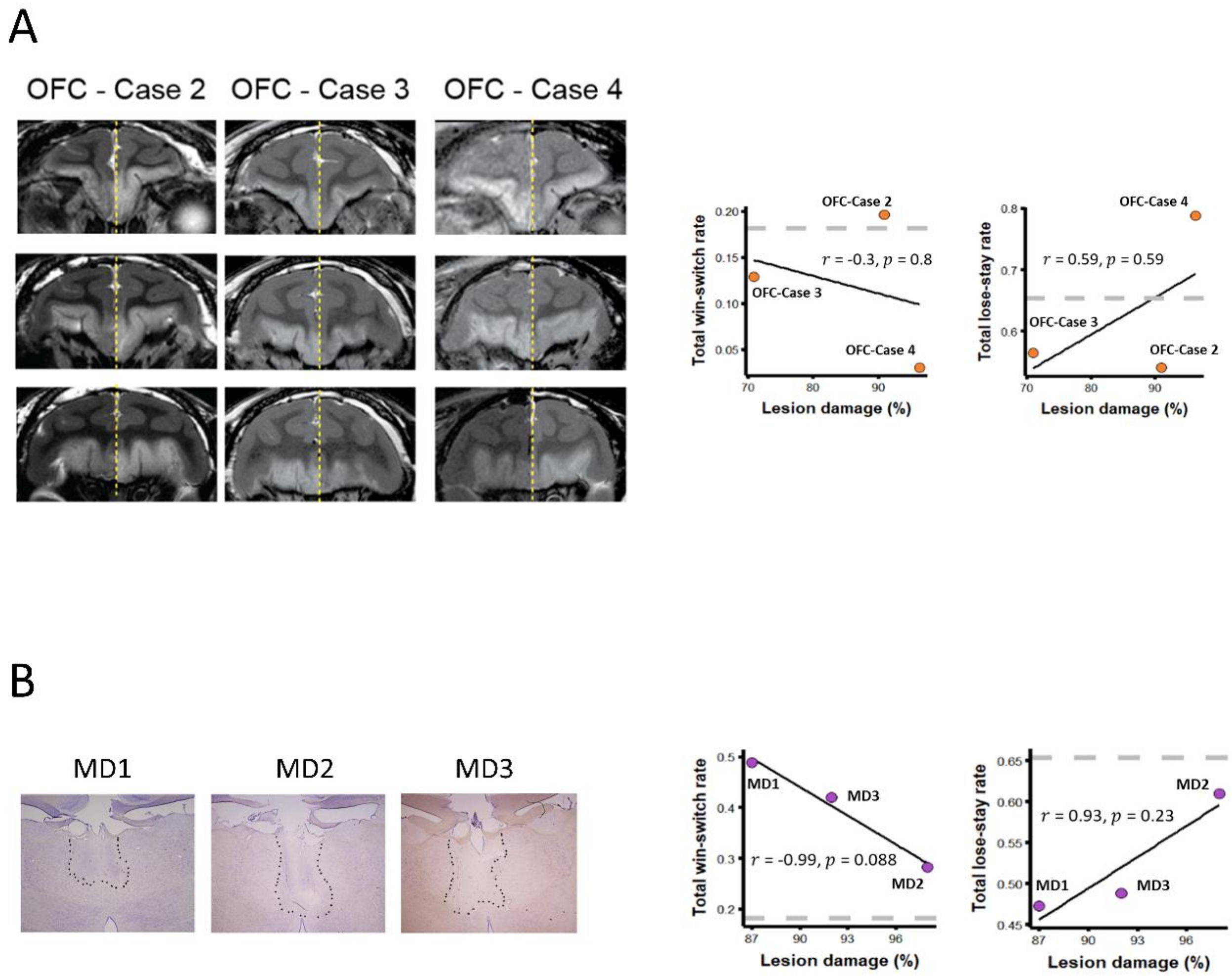
Supplementary Figure S4. Probabilistic reversal learning task performance across multiple sessions in monkeys. The upper and lower panels show the win-switch and lose-stay behavioral measures, respectively, across five individual sessions (Session 1-5), with lag-trial bins on the x-axis and behavioral metric on the y-axis. The rightmost graphs present the average behavior across all sessions. This repeated measures design mirrors the multi-reversal human PRL task structure, thereby capturing the increased volatility characteristic of the task. The congruent patterns of behavior between the monkey multiple-session data and the human multi-reversal PRL task suggest a comparable elicitation of behavioral responses to volatility. These similarities support the conclusion that tasks with higher volatility may more effectively differentiate behavioral tendencies associated with levels of paranoia.



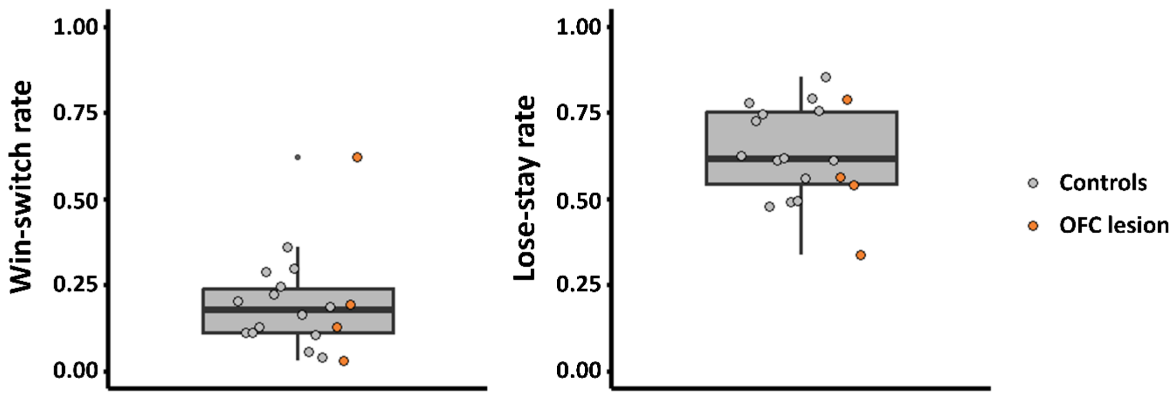
Supplementary Figure S5. Association of volatility belief with depression and cognitive reflection. Panel A (left) illustrates the difference of belief volatility in two clusters: (i) individuals with clinical depression and nonclinical paranoia and (ii) individuals with nonclinical depression and clinical paranoia, revealing that higher paranoia correlates with greater belief volatility, independent of depressive clinical symptoms (χ^2 = 5.47, p = 0.019). Panel A (right) illustrates similar differences but in two different clusters: (i) individuals with higher cognitive reflection (or higher approximate IQ - scores on a series of three cognitive reflection questions; refer to Table 7 in prior work[3]) and nonclinical paranoia and (ii) individuals with lower IQ with clinical paranoia, revealing again that higher paranoia correlates with greater belief volatility, independent of IQ (χ^2 = 83.82, p < 0.001). Panel B illustrates correlation between paranoia and depression (left) and paranoia and IQ (right).



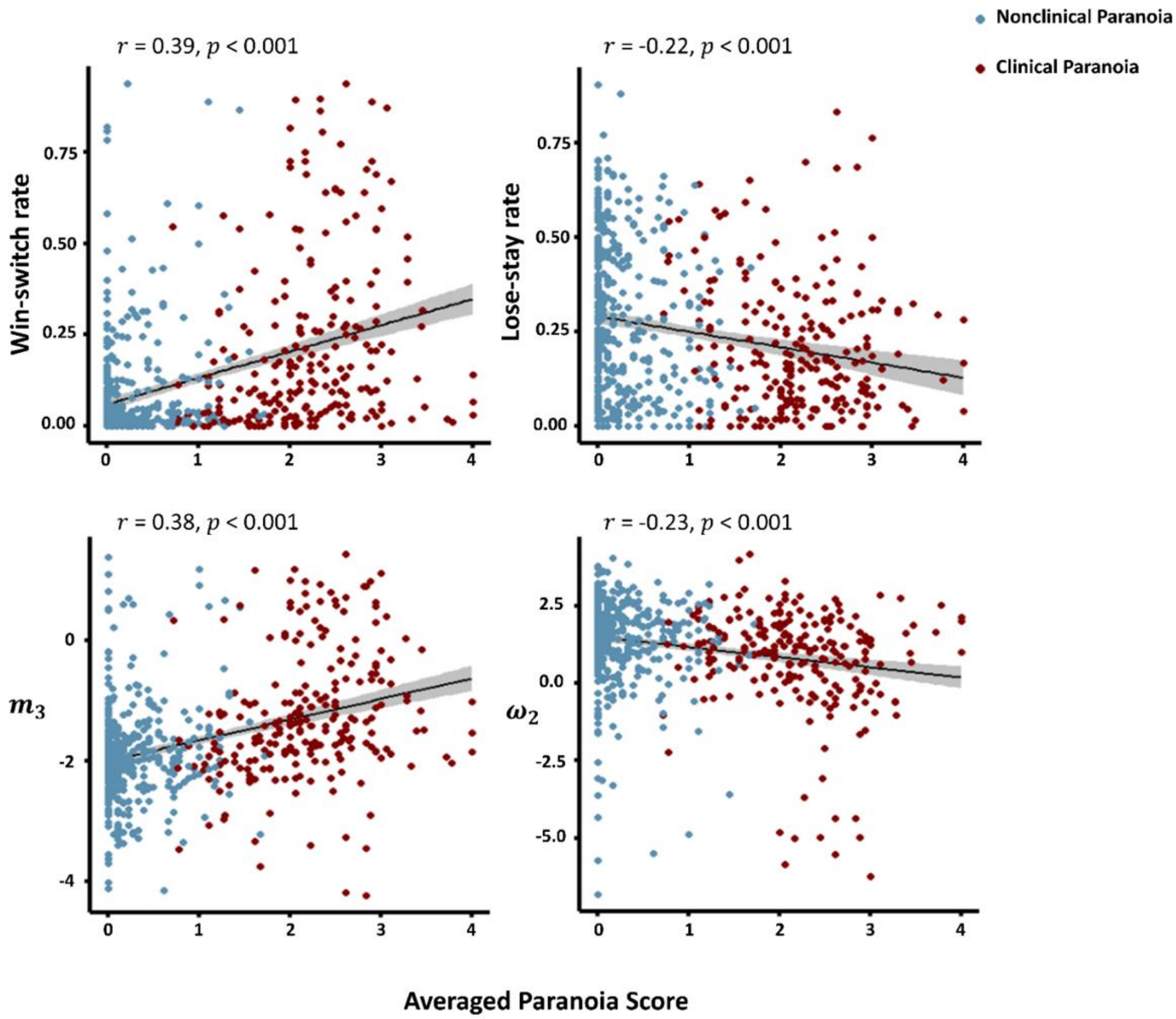
Supplementary Figure S6. Ventrolateral Prefrontal Cortex (VLPFC) lesion effect on PRL task behavior and belief-updating. We observe VLPFC lesion effects in behavior and belief-updating similar to MDmc lesion effects; that is, elevated win-switching (χ^2 = 79.99, p < 0.001) and belief volatility (χ^2 = 4.17, p = 0.041) after the reversal. A likelihood ratio test comparing models with and without the reversal term was used to assess the effect of reversal phase on behavior and belief-updating in VLPFC-lesioned monkeys. Refer to prior study for details on surgical and lesion procedures of the VLPFC[22].



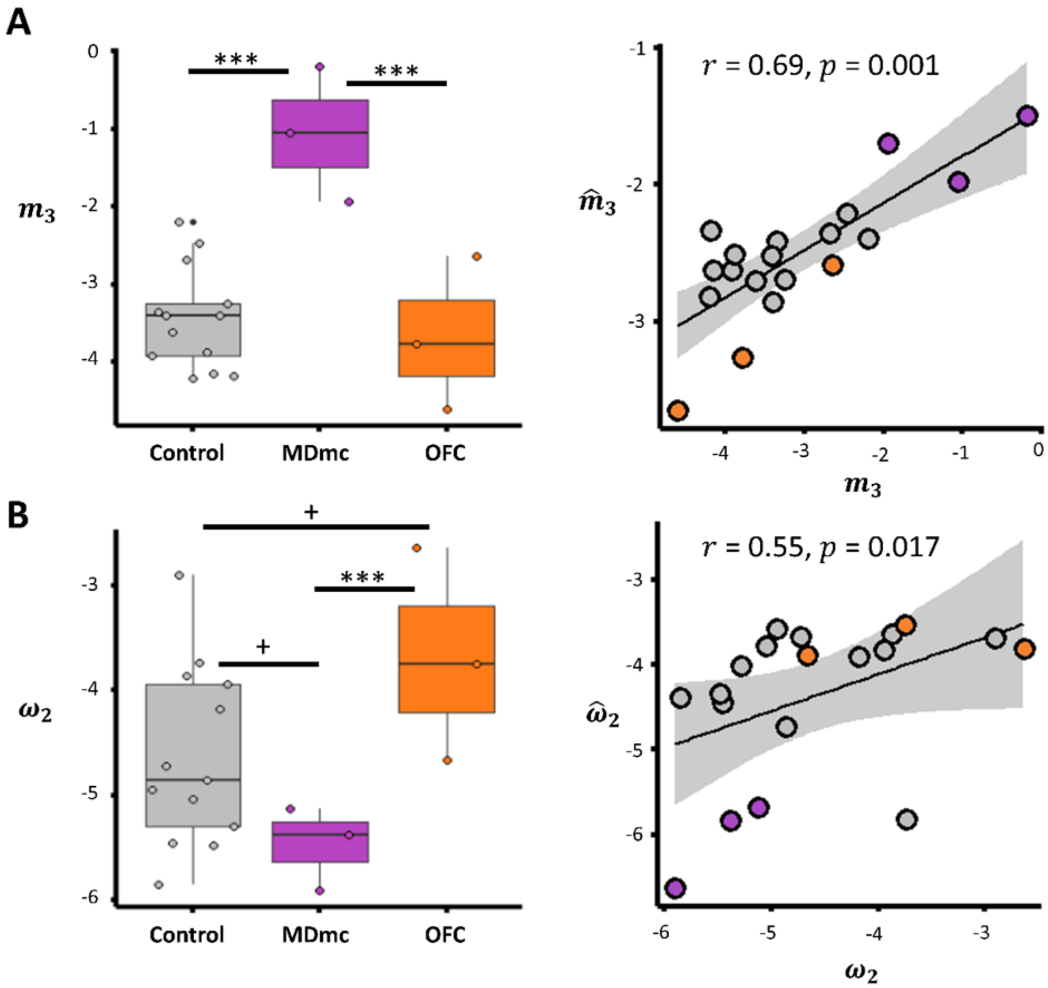
Supplementary Figure S7. Relationship between lesion damage and PRL task behavior. (A) Panel A (left) displays MRI scans illustrating the extent of lesions in the orbitofrontal cortex (OFC). Panel A (right) are two scatter plots of the correlation between the percent lesion damage in the OFC and task behavior. (B) Panel B (left) illustrates the extent of the MDmc lesion, histologically. Panel B (right) are two scatter plots of the correlation between the percent lesion damage in the MDmc and task behavior. Dotted lines represent average behavior in control monkeys. Images of MDmc and OFC lesions were adopted from prior studies[13,22].



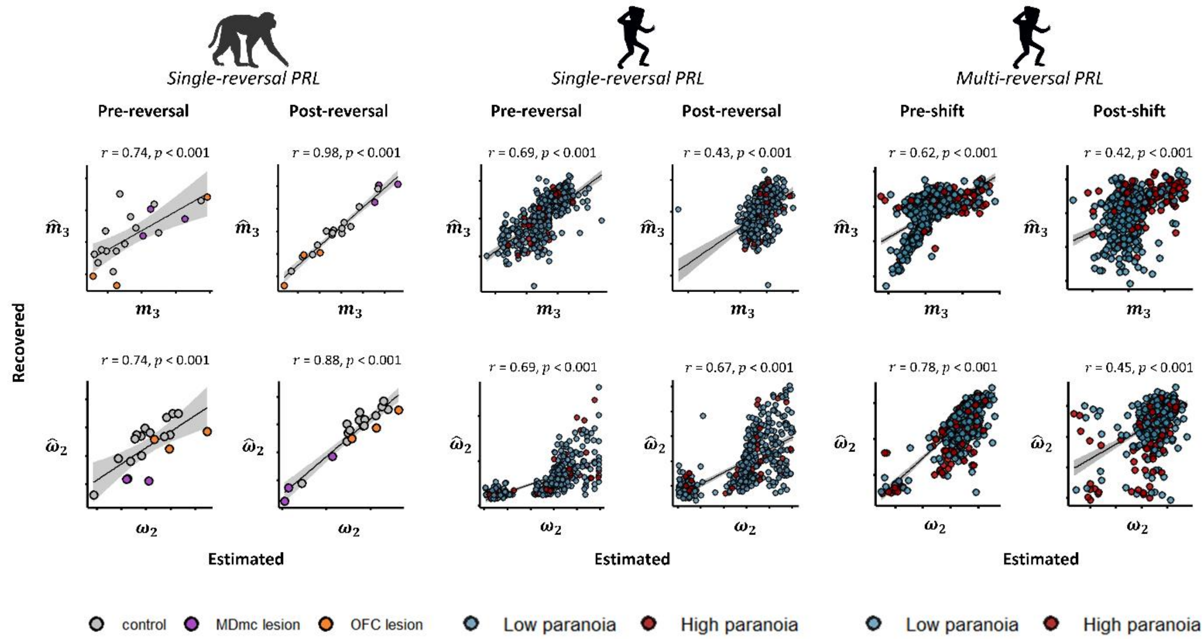
Supplementary Figure S8. Exclusion criteria of one monkey outlier based on noisy win-switching behavior. We observe a particular monkey in the OFC lesion group behaved differently to its relative members. A Interquartile Range (IQR) test suggests that monkey is an outlier when including all the unoperated controls and OFC-lesioned monkeys.



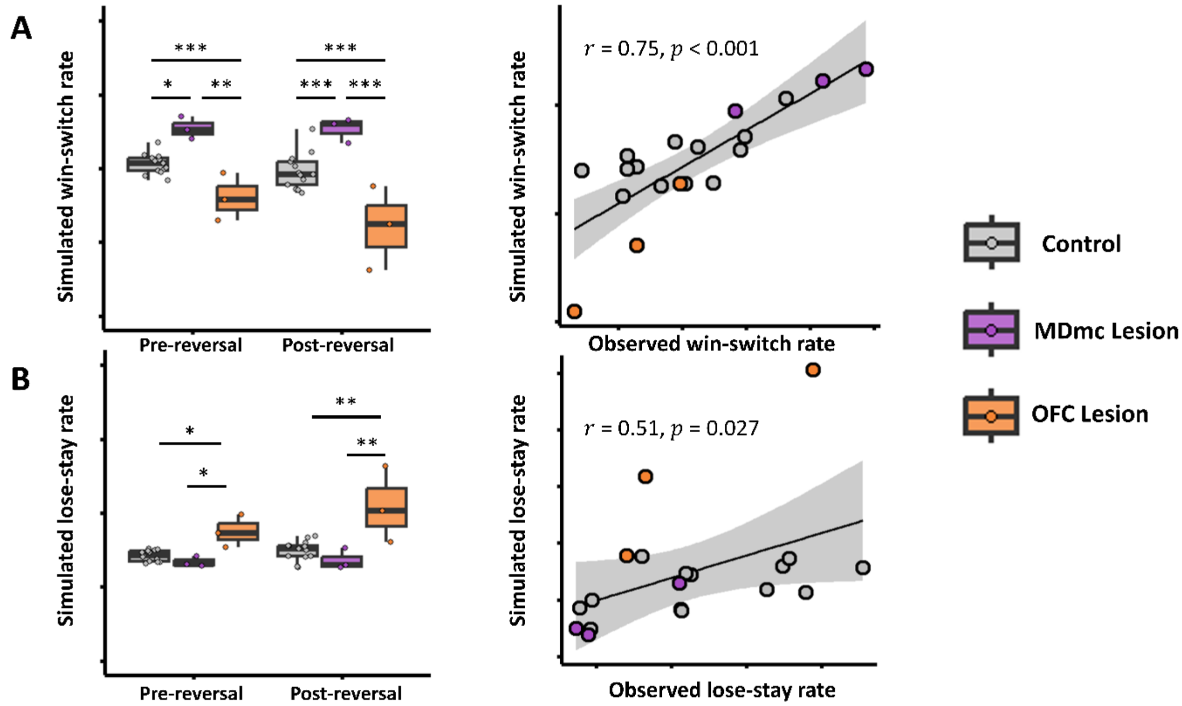
Supplementary Figure S9. Relationship between PRL task behavior and belief-updating to paranoia. We show that individuals with paranoia who performed the PRL task typically exhibit elevated win-switching and belief volatility with diminished lose-staying and value learning.



Supplementary Figure S10. Hierarchical Gaussian Filter model fit to the entire dataset. We show that when we fit the model to the entire dataset (i.e., using all trial information, instead of splitting), we still maintain the primary findings. (A) Belief volatility differences across lesion groups; MDmc-lesioned monkeys exhibit elevated belief volatility compared to both control monkeys (χ^2 = 32.64, p < 0.001) and OFC-lesioned monkeys (χ^2 = 17.71, p < 0.001). (B) Value learning rate differences across lesion groups; OFC-lesioned monkeys exhibit increased value learning compared to MDmc-lesioned monkeys (χ^2 = 12.06, p < 0.001), but we observe a trend between controls and OFC-lesioned monkeys (χ^2 = 3.28, p = 0.07) and controls and MDmc-lesioned monkeys (χ^2 = 3.01, p = 0.08). This suggests that our modeling approach presented in the paper, using priors from the first half to the second half of the experiment, does not impact the results. Parameter recovery supports these findings.



Supplementary Figure S11. Reliable model parameter recovery between monkeys and humans for each experimental task design. These parameter recovery plots echo the group-specific (lesion and paranoia) effects on belief parameters. The correlation measures underscore the reliability of the model and supports the parameter interpretations presented in our study.



Supplementary Figure S12. Model validation using simulated choice data. We show similar behavioral effects across groups using simulated choice data derived from the parameter recovery approach. (A) Simulated win-switch behavior recapitulate lesion group effects before reversal (control vs MDmc: χ^2 = 6.04, p = 0.014; control vs OFC: χ^2 = 7.97, p = 0.005; MDmc vs OFC: χ^2 = 16.57, p < 0.001) and after reversal (control vs MDmc: χ^2 = 16.07, p < 0.001; control vs OFC: χ^2 = 33.60, p < 0.001; MDmc vs OFC: χ^2 = 57.04, p < 0.001). (B) Simulated lose-stay behavior recapitulate lesion group effects before reversal (control vs OFC: χ^2 = 6.08, p = 0.014; MDmc vs OFC: χ^2 = 6.01, p = 0.014) and after reversal (control vs OFC: χ^2 = 10.13, p = 0.001; MDmc vs OFC: χ^2 = 10.17, p = 0.001). Correlations reveal model reliably simulates win-switch and lose-stay behavior observed in the actual data.

Supplementary Table

Table 1. Prior mean and variance of the perceptual and response models of the Hierarchical Gaussian Filter

Perceptual: *The model of how we see the world*		
Parameter	**Prior mean**	**Prior variance**
κ	0.5	1
ω	-5	16
θ	0.5	1
μ_2^0	0	0
σ_2^0	1	0
μ_3^0	1	1
σ_3^0	0.1	1
m_3	1	1
ϕ	0.1	0
Response: *The model of what choice we make*		
Parameter	**Prior mean**	**Prior variance**
β	48	1