Reduced top-down connectivity as an underlying mechanism for psychotic experiences in healthy people

Authors: Ilvana Dzafic[1,3]*, Roshini Randeniya[1], Marta I. Garrido[1,2,3,4]

Affiliations

[1]Queensland Brain Institute, University of Queensland, Brisbane, Australia

[2]School of Mathematics and Physics, University of Queensland, Brisbane, Australia

[3]Australian Research Council Centre of Excellence for Integrative Brain Function, Australia

[4]Centre for Advanced Imaging, University of Queensland, Brisbane, Australia

***Correspondence should be addressed to:**

Dr Ilvana Dzafic; i.dzafic@uq.edu.au

Abstract

Perception results from our brain's ability to make predictive models of sensory information. Recently, it has been proposed that psychotic traits may be linked to impaired predictive processes. Here, we examine the brain dynamics underlying prediction formation in a population of healthy individuals with a range of psychotic experiences. We designed a novel paradigm, which incorporated both stable and volatile sound sequences by manipulating their probability. We measured prediction error with electroencephalography and gauged prediction formation explicitly by behaviourally recording sensory 'regularity' learning errors. Critically, we show that top-down frontotemporal connectivity is a neural mechanism by which impaired regularity learning influences psychotic experiences. These findings further our understanding of the neurobiological underpinnings of prediction formation and provide evidence for a continuum of psychosis in the healthy, non-clinical population.

Keywords: prediction, volatility, psychotic, mediation, effective connectivity

One Sentence Summary

Healthy individuals with psychotic experiences have impaired sensory learning, mediated by reduced top-down frontotemporal connectivity.

In a stable environment, sensory perception is facilitated by prior beliefs about what is likely to happen next (*1, 2*). By estimating the probability of forthcoming events, we can form a predictive model about the world and its regularities (*3*). When circumstances are 'volatile', such that previously learnt regularities change, having a flexible predictive model is more advantageous (*4-6*). Previous literature has shown that healthy individuals are able to optimally estimate environmental volatility (*6*), adopting a greater learning rate in the face of ever changing, volatile circumstances (*6, 7*). This motivates exploratory behaviour and continuous updating, as well as suppression of top-down prior beliefs (*8*). However, this state of constant learning is inefficient as a long-term strategy in stable environments (*9, 10*). Stable environments allow for the development of a robust predictive model, which simultaneous enables an efficient encoding of sensory stimuli while minimising the spending of cognitive resources (*5*). Poor estimation of environmental volatility has been found to have negative consequences on social cognition and decision-making in individuals with autism, anxiety and schizophrenia (*11-13*). These patient groups have aberrant representations of volatility; either over-estimating it, leading to imprecise, weak predictive models, or under-estimating it, leading to rigid and maladaptive predictive models (*14*).

Emerging theoretical accounts of psychosis postulate that psychotic experiences arise due to an impairment in the brain's predictive ability to infer internal and external sensations (*14-16*). Individuals experiencing psychosis (such as schizophrenia) may misattribute saliency to irrelevant sensory information leading to the formation of unfounded odd beliefs. There are many converging lines of evidence that support the theory that psychosis arises due to an impaired predictive model (*17, 18*). The most robust and replicable empirical evidence for this arises from attenuated neurophysiological responses to surprising sounds embedded in a sequence of predictable sounds, in so-called oddball paradigms (*19-21*). This is thought to reflect a sensory *prediction error* deficit that results from a failure to form accurate predictions

about forthcoming predictable stimuli (*15, 22, 23*). Reduced prediction error (PE) response in schizophrenia has been linked to alterations in brain connectivity between the frontal and the temporal cortex as well as within the inferior frontal gyrus (IFG) and auditory cortex (*24, 25*). Alterations in the neurophysiology of PEs have been shown to increase as psychotic traits increase, suggesting that the degree of PE aberrancy aligns on a continuum of psychosis (*26, 27*) – the idea that nonclinical individuals in the general healthy population may display a range of psychotic traits. However, the underlying brain networks, presumably also altered along the continuum remain unknown. The psychosis continuum comprises the full spectrum of psychotic experiences, from healthy individuals who experience a range of psychotic-like experiences, to prodromal individuals with subclinical symptoms, and to those with florid psychosis at the very end of the spectrum (*28, 29*). One of the benefits of investigating psychosis in relation to neural dynamics (*30*) and PE response (*31, 32*) on the healthy end of the spectrum is the possibility to eschew the confounds of medication and illness severity.

The aims of this study were three-fold. The first aim was to examine the relationship between regularity learning and sensory PE measured with electroencephalography (EEG). Secondly, we wanted to elucidate the neural dynamic underpinnings of prediction formation during regularity learning. For this purpose, we developed a novel auditory oddball task with either fixed sound probabilities (stable conditions) or varying sound probabilities (volatile conditions; see Figure S1). The third aim was to investigate whether aberrations in prediction formation are aligned on a continuum of psychosis. Specifically, we examined the relationship between regularity learning, PE responses and the brain networks engaged in prediction violations. We hypothesised that disruptions in intrinsic and top-down brain dynamics (*24*) mediate the influence of impaired prediction formation on increased psychotic experiences.

Results and Discussion

Our first aim was to compare the strength of the predictive models established in stable and volatile in a regularity learning task. For this purpose we examined the event-related potential (ERP) recorded at frontocentral channel (Fz), and in line with the vast oddball literature we found that responses to deviant sounds were larger than responses elicited by standard sounds, regardless of volatility ($F(1,30) = 45.33$, $p < 0.001$, $\eta^2 = 0.60$). Moreover, we found a significant interaction between PE response and volatility ($F(1,30) = 11.06$, $p = 0.002$, $\eta^2 = 0.27$). Critically, a follow-up analysis revealed that PEs were larger under the stable compared to the volatile conditions, $t(30)=-3.33$, $p = 0.002$, $d = -0.60$ (see Figure 1a and 1b). Increased PEs in stable conditions, compared to volatile, have been identified in previous studies (*33, 34*). PEs signal a violation between what the brain predicts will happen and what is actually experienced. As such, PEs are fundamental teaching signals that drive updating of the brain's predictive model of the sensed world (*35, 36*). Thus, greater PE responses to regularity violations indicate a stronger (i.e., more precise) prediction model in *stable* than volatile conditions (*36, 37*).

We next examined differences in mean percentage errors in probability estimation (regularity learning error) and mean confidence ratings during stable and volatile conditions. The data show that participants had fewer errors in regularity learning during stable conditions ($M = 9.30\%$, $SE = 0.87$), compared to volatile conditions ($M = 12.38\%$, $SE = 1.07$), $t(30) = -2.40$, $p = 0.023$, $d = -0.43$. In addition, participants had greater confidence in their probability estimates during stable conditions ($M = 2.57$, $SE = 0.07$), compared to volatile conditions ($M = 2.19$, $SE = 0.05$), $t(30) = 4.78$, $p < 0.001$, $d = 0.86$ (see Figure 1c). This shows that regularity learning and confidence are enhanced in stable, more predictable environments, than in more volatile, less predictable, environments. In addition, we asked whether regularity learning errors were related to the degree of PE response. Pearson's correlations and Bayesian analysis

revealed a very strong, significant correlation between regularity learning errors and PEs (at the ERP level) in stable conditions (p = 0.003 ($p_{adjusted}$ < 0.01), BF_{+0} = 33.99; see Figure 1d, Table S2), hence demonstrating that greater PEs in *stable conditions* are associated with better sensory regularity learning. Regularity learning is the process by which the brain learns the statistical structure in the environment and forms predictive models of what is likely to happen next (*38-41*). Previous studies have demonstrated that individuals are able to implicitly learn the statistical structure of sensory events in the environment (*3, 23*). Crucially, by simultaneously recording PE responses and behaviourally measuring regularity learning, we show for the first time that greater sensory PEs are associated with improved explicit ability to gauge the sensory regularities within one's environment.

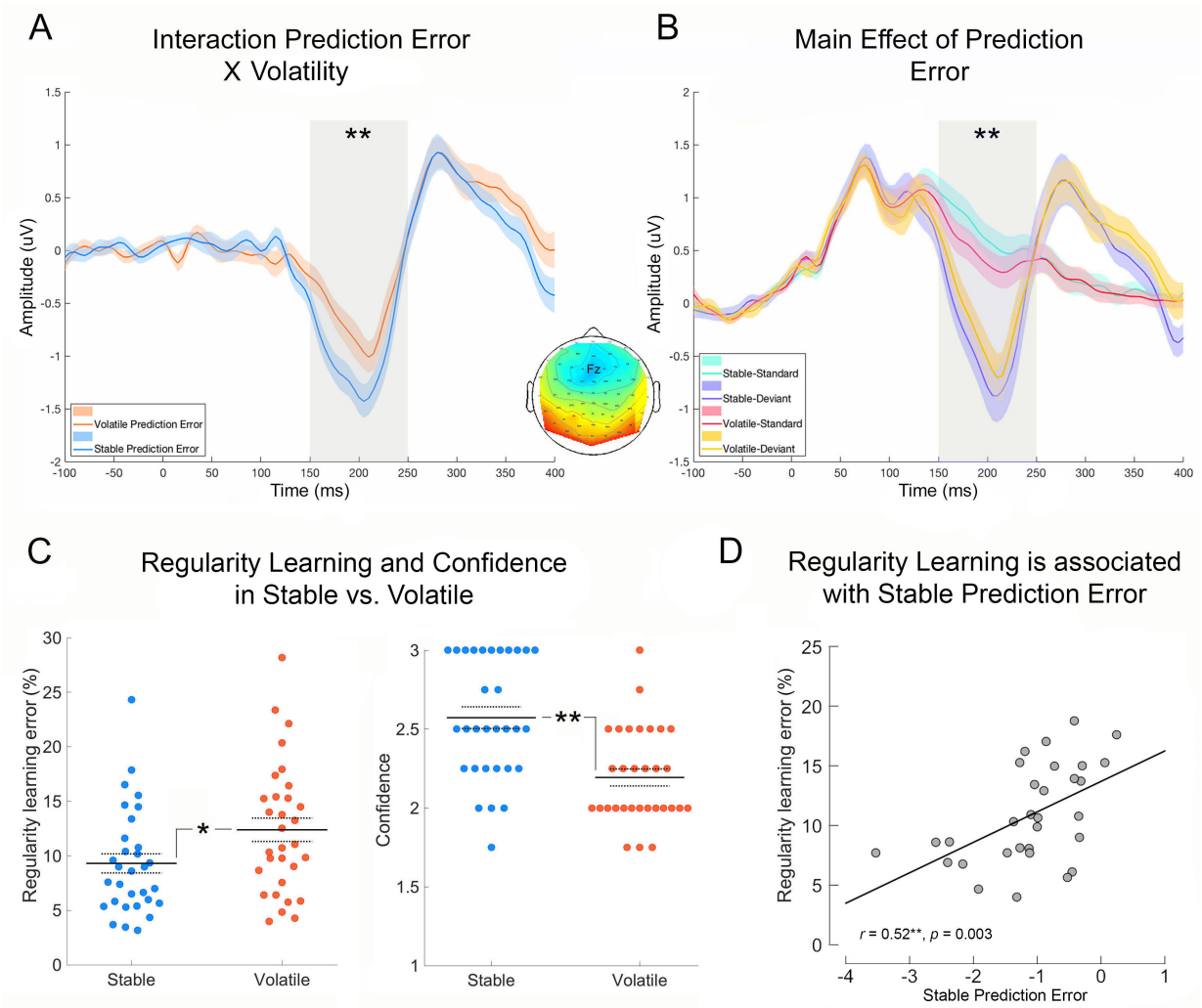


Fig. 1. Greater prediction errors and regularity learning indicate a more precise prediction model in *stable* than *volatile* conditions. A) Significant interaction between volatility (stable – blue – versus volatile – orange) and brain PE responses, B) Significant main effect of PE, showing brain responses evoked by standards (green, red) and deviants (purple and yellow) in the context of stable and volatile conditions. Solid lines represent mean and lighter shading represent standard error of the mean, C) Fewer regularity learning errors and greater confidence in stable (blue) than volatile (orange) conditions, D) Significant correlation between regularity learning errors and prediction errors in stable conditions. * $p < 0.05$; ** $p < 0.001$.

To further investigate the sensory PEs evoked by regularity violations with fewer spatial and temporal constraints, we ran a general linear model for the whole spatiotemporal volume of brain activity. Firstly, we replicated previous auditory oddball findings by showing a significant main effect of PE response (standard sounds vs deviant sounds) peaking at 205 ms (peak-level $F = 170.92$, $p_{FWE} < 0.001$), 290 ms (peak-level $F = 240.99$, $p_{FWE} < 0.001$; frontocentral and occipitoparietal channels), and 25 ms (right frontal channels; peak-level $F = 25.26$, $p_{FWE} = 0.004$). Moreover, we found a significant interaction between PE response and volatility, at 165 ms over occipitocentral channels (peak-level $z = 4.27$, $p_{FWE} = 0.015$, see Figure 2a). Next, we asked whether regularity learning errors were related to neuronal activity. To address this question, we conducted a spatiotemporal multiple regression analysis at the interaction between PEs and volatility (Stable PEs > Volatile PEs) with regularity learning error as the predictor variable. Our data show that a decrease in regularity learning errors significantly predicted an increase in brain activity at 165 ms (peak-level $z = 3.64$, cluster-level $p_{FWE} = 0.034$, see Figure 2b). In order to determine where in the brain this effect came from we used source reconstruction techniques (*42*), which uncovered an increased activity in the right superior frontal gyrus (peak-level $z = 2.19$, $p_{uncorrected} = 0.014$) and the right fusiform gyrus (peak-level $z = 1.89$, $p_{uncorrected} = 0.029$, see Figure S2b). This finding demonstrates that the difference in PEs in stable and volatile environments (ability to attune to volatility) increases as regularity learning improves, associated with activity in right frontotemporal regions. This finding is concordant with the idea that healthy individuals are optimally attuned to different

environments, such that in volatile environments there is a greater reliance on local probability (smaller PE) (*43*), whereas stable environments enable stronger neuronal representations, or more precise predictive models (larger PE), of global regularities (*44*).

Source-level analysis revealed that stable PEs engaged frontoparietal regions, such as the middle frontal gyrus (peak-level $z = 4.23$, $p_{FWE} = 0.02$), the primary motor area (peak-level $z = 4.45$, $p_{FWE} = 0.009$), and the inferior parietal lobule (peak-level $z = 4.56$, $p_{FWE} = 0.006$). In comparison, volatile PEs engaged occipitoparietal regions, such as the precuneus (peak-level $z = 5.05$, $p_{FWE} = 0.001$) and the middle occipital gyrus (peak-level $z = 4.39$, $p_{FWE} = 0.011$, see Figure 2c). These results are in keeping with prior studies suggesting that higher hierarchical frontal regions (engaged for stable PEs) are associated with formation and representation of prior beliefs (*45-47*) and activity in inferior parietal regions is associated with the evaluation of prior beliefs (*48, 49*). In comparison, lower hierarchical occipital regions (engaged for volatile PEs) are associated with sensory processing (*46*), which drive prior belief updating.

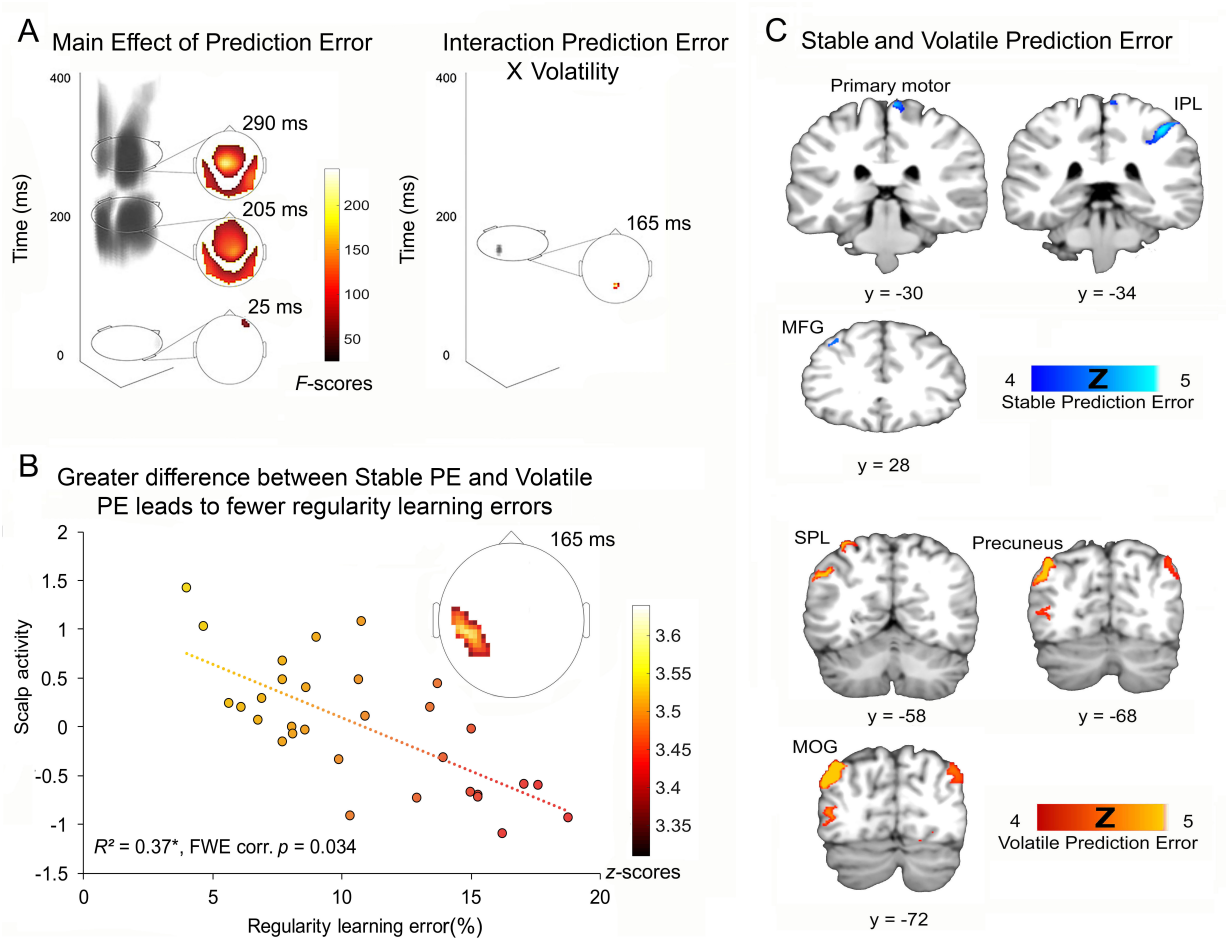


Fig. 2. Brain responses underlying volatility attuning and regularity learning. A) Spatiotemporal univariate statistical analysis revealed a significant main effect of PE (left column) over frontal and occipitoparietal channels; and PE x volatility interaction (right column) over occipitocentral channels. B) Spatiotemporal multiple regression analysis revealed a negative relationship between regularity learning errors and spatiotemporal activity during the interaction (Stable PE > Volatile PE) at 165ms. C) Source reconstruction analysis revealed significant clusters for stable PEs in frontoparietal regions versus for volatile PEs in occipitoparietal regions. IPL = inferior parietal lobule; MFG = middle frontal gyrus; SPL = superior parietal lobule; MOG = middle occipital gyrus. All maps are displayed at $p < 0.05$, FWE whole-volume corrected.

The network architecture underlying PE response has been extensively studied previously (*50, 51*), with robust findings demonstrating that a three-level hierarchical brain model underlies the generation of PEs evoked in auditory oddball paradigms. In the current study, we focused on the pattern of connections that best *differentiates* PE responses under stable and volatile environments. The high temporal resolution of EEG data enables improved estimation of the underlying neurobiological interactions, providing insights into the brain's

effective connectivity (*52, 53*). Here, we were interested in, 1) the effect of contextual volatility on neuronal responses, and 2) the effective connectivity related to psychotic traits.

Bayesian model comparison was performed on thirty-six different dynamic causal models (see Figure S3), which were based on the functional brain architecture shown to underlie PE responses (*50, 51*). Results from Bayesian model selection using random effects family-level analysis indicated that the best model included connections amongst six *a priori* defined regions, with inputs to left and right primary auditory cortex (A1); intrinsic connections within the A1; bilateral connections between: A1 and superior temporal gyri (STG), STG and inferior frontal gyri (IFG); as well as lateral connections between left A1 and right A1, and left STG and right STG (model 7; see Figure 3a). In the optimal model, larger PEs in stable compared to volatile blocks were caused by enhanced modulations in backward, forward and intrinsic connections (see Figure 3a). Forward connections are thought to convey PEs, whereas backward connections covey predictions (i.e., beliefs about sensory input), and intrinsic connections emulate local adaptation of neuronal responses and are thought to reflect the precision (strength) of neuronal representations (*53, 54*). This finding is in keeping with the predictive coding account of the mechanisms underlying perception of an auditory oddball sequence (*55*), and suggest that more precise models about sensory input are enabled by greater brain connectivity in stable than volatile PEs.

In order to test the evidence for a continuum of psychosis, we examined the altered neural dynamics, behaviour and neurophysiology related to psychotic traits in the general healthy population. First, we examined brain connectivity estimates by applying Bayesian model averaging across all models (weighted by their probability) and participants. Critically, we found a strong, significant correlation between psychotic experiences and top-down connectivity from the right IFG to STG (frontotemporal) ($p = 0.005$ ($p_{adjusted} < 0.05$), $BF_{-0} = 18.28$; see Table S3). This shows that a greater degree of psychotic traits in healthy people was

associated with weaker top-down connectivity from inferior frontal to superior temporal regions. Precisely the same connection has previously been found aberrant in patients with schizophrenia (*24*), as well as high-risk individuals with a genetic predisposition for schizophrenia (*56*), and is aligned with the dysconnectivity hypothesis for schizophrenia, observed particularly between frontotemporal regions (*57, 58*). Next, we asked if aberrations in behaviour (greater regularity learning errors) and neurophysiology (attenuated PE) are also aligned on the psychosis continuum. Pearson's and Bayesian correlations were conducted on psychotic experiences, regularity learning errors, as well as PEs (at the ERP level) in stable and volatile conditions. We found a moderate, significant correlation between psychotic experience and errors in regularity learning ($p = 0.028$, $BF_{+0} = 4.37$; see Table S2), meaning that healthy individuals with greater psychotic experiences were worse at learning about sensory regularities.

Our final analysis explored if the top-down frontotemporal connection, which was weaker in individuals with more psychotic experiences, was the underlying mechanism by which regularity learning errors influenced psychotic traits. For this purpose, we employed a mediation analysis, which establishes the mechanism that enables a predictor to influence an outcome. Multiple regressions were conducted to asses each component of the mediation analysis (see Figure 3c). The results demonstrated that regularity learning error was a significant predictor of top-down frontotemporal connectivity ($b = -0.009$, $p = 0.02$), and that top-down frontotemporal connectivity was a significant predictor of psychotic experience ($b = -4.98$, $p = 0.04$), supporting the mediation hypothesis. Regularity learning was no longer a significant predictor of psychotic experiences after controlling for the mediator, top-down frontotemporal connectivity ($b = 0.08$, $p = 0.17$), consistent with a full mediation (see Figure 3b). The results indicate a significant indirect effect (*ab* = total effect - direct effect) of regularity learning on psychotic experience through top-down frontotemporal connectivity (*ab*

= 0.05, Bias Corrected and Accelerated Bootstrap (BCA) CI [0.004, 0.14], P_M = 0.38 - percent mediation: percent of the total effect accounted for by the indirect effect). Critically, these findings identify top-down frontotemporal connectivity as a mechanism by which poorer regularity learning influences increased severity of psychotic experiences in healthy people.

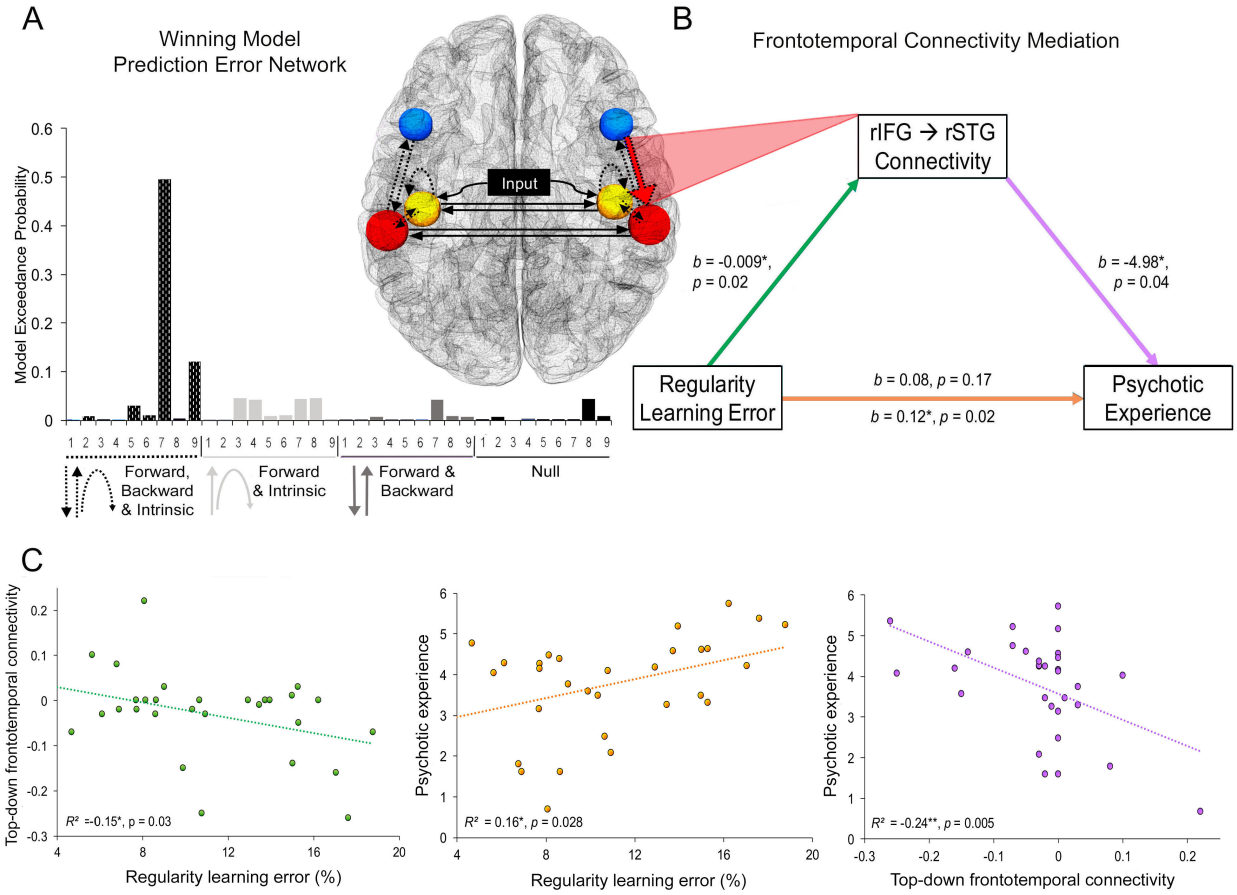


Fig. 3. Top-down frontotemporal connectivity mediates the relationship between regularity learning errors and severity of psychotic experiences in healthy people. A) 36 competing connectivity models tested the modulation of connections for stable > volatile PEs. The winning model architecture had connections between all six regions bilaterally (model 7), and included volatility-dependent (stable > volatile) modulations in forward, backward and intrinsic connections, B) The mediation analysis revealed a significant full mediation, with an indirect effect of regularity learning on psychotic experience through top-down frontotemporal connectivity, C) Regression plots, demonstrating that regularity learning errors predict top-down frontotemporal connectivity (green) and psychotic experiences (orange), and that top-down frontotemporal connectivity predicts psychotic experiences (purple).

In the current study, we explored the brain mechanisms underpinning regularity

learning under uncertainty, and the relationship between disruptions to predictive processes and psychotic experiences in healthy individuals. We found that individuals learn better and their brain PE responses are greater during stable than volatile conditions. At the neural level, there is a greater engagement of higher hierarchical regions, such as the middle frontal gyrus, as well as greater modulation of intrinsic, forward and backward connections. Importantly, our data show that aberrations in the brain's predictive model are aligned on a continuum of psychosis, in the sense that healthy people with more psychotic traits have poorer regularity learning abilities, driven by weaker top-down frontotemporal connectivity. Our findings have implications for understanding the neurobiological underpinnings of impaired prediction formation, with the potential to inform the application of neuromodulation therapies for psychosis targeting the frontotemporal network.

References and Notes

1. M. T. Sherman, A. K. Seth, A. B. Barrett, R. Kanai, Prior expectations facilitate metacognition for perceptual decision. *Conscious Cogn* **35**, 53-65 (2015).
2. M. Moutoussis, P. Fearon, W. El-Deredy, R. J. Dolan, K. J. Friston, Bayesian inferences about the self (and others): a review. *Conscious Cogn* **25**, 67-76 (2014).
3. M. I. Garrido, M. Sahani, R. J. Dolan, Outlier responses reflect sensitivity to statistical structure in the human brain. *PLoS Comput Biol* **9**, e1002999 (2013).
4. C. Mathys, J. Daunizeau, K. J. Friston, K. E. Stephan, A bayesian foundation for individual learning under uncertainty. *Front Hum Neurosci* **5**, 39 (2011).
5. S. Vossel *et al.*, Spatial attention, precision, and Bayesian inference: a study of saccadic response speed. *Cereb Cortex* **24**, 1436-1450 (2014).
6. T. E. Behrens, M. W. Woolrich, M. E. Walton, M. F. Rushworth, Learning the value of information in an uncertain world. *Nat Neurosci* **10**, 1214-1221 (2007).
7. A. O. Diaconescu *et al.*, Inferring on the intentions of others by hierarchical Bayesian learning. *PLoS Comput Biol* **10**, e1003810 (2014).
8. P. Dayan, J. Y. Angela, in *Advances in neural information processing systems*. (2003), pp. 173-180.
9. K. Friston, The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience* **11**, 127-138 (2010).
10. P. Schwartenbeck, T. FitzGerald, R. J. Dolan, K. Friston, Exploration, novelty, surprise, and free energy minimization. *Frontiers in psychology* **4**, (2013).
11. L. Deserno *et al.*, Overestimating environmental volatility increases switching behavior and is linked to activation of dorsolateral prefrontal cortex in schizophrenia. *bioRxiv*, 227967 (2017).
12. M. Browning, T. E. Behrens, G. Jocham, J. X. O'Reilly, S. J. Bishop, Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nat Neurosci* **18**, 590-596 (2015).
13. R. P. Lawson, C. Mathys, G. Rees, Adults with autism overestimate the volatility of the sensory environment. *Nat Neurosci* **20**, 1293-1299 (2017).
14. A. R. Powers, C. Mathys, P. R. Corlett, Pavlovian conditioning-induced hallucinations result from overweighting of perceptual priors. *Science* **357**, 596-600 (2017).
15. R. A. Adams, K. E. Stephan, H. R. Brown, C. D. Frith, K. J. Friston, The computational anatomy of psychosis. *Front Psychiatry* **4**, 47 (2013).

16. R. A. Adams, Q. J. Huys, J. P. Roiser, Computational Psychiatry: towards a mathematically informed understanding of mental illness. *J Neurol Neurosurg Psychiatry* **87**, 53-63 (2016).
17. D. Dima, D. E. Dietrich, W. Dillo, H. M. Emrich, Impaired top-down processes in schizophrenia: a DCM study of ERPs. *Neuroimage* **52**, 824-832 (2010).
18. L. E. Hong *et al.*, Refining the predictive pursuit endophenotype in schizophrenia. *Biol Psychiatry* **63**, 458-464 (2008).
19. M. Bodatsch *et al.*, Prediction of psychosis by mismatch negativity. *Biol Psychiatry* **69**, 959-966 (2011).
20. R. Naatanen, T. Shiga, S. Asano, H. Yabe, Mismatch negativity (MMN) deficiency: a break-through biomarker in predicting psychosis onset. *Int J Psychophysiol* **95**, 338-344 (2015).
21. D. Umbricht, S. Krljes, Mismatch negativity in schizophrenia: a meta-analysis. *Schizophr Res* **76**, 1-23 (2005).
22. M. Aghamolaei, K. Zarnowiec, S. Grimm, C. Escera, Functional dissociation between regularity encoding and deviance detection along the auditory hierarchy. *Eur J Neurosci* **43**, 529-535 (2016).
23. F. Lecaignard, O. Bertrand, G. Gimenez, J. Mattout, A. Caclin, Implicit learning of predictable sound sequences modulates human brain responses at different levels of the auditory hierarchy. *Front Hum Neurosci* **9**, 505 (2015).
24. D. Dima, S. Frangou, L. Burge, S. Braeutigam, A. C. James, Abnormal intrinsic and extrinsic connectivity within the magnetic mismatch negativity brain network in schizophrenia: a preliminary study. *Schizophr Res* **135**, 23-27 (2012).
25. S. Ranlund *et al.*, Impaired prefrontal synaptic gain in people with psychosis and their relatives during the mismatch negativity. *Human brain mapping* **37**, 351-365 (2016).
26. R. Randeniya, L. K. L. Oestreich, M. I. Garrido, Sensory prediction errors in the continuum of psychosis. *Schizophr Res*, (2017).
27. H. Verdoux, J. van Os, Psychotic symptoms in non-clinical populations and the continuum of psychosis. *Schizophr Res* **54**, 59-65 (2002).
28. P. DeRosse, K. H. Karlsgodt, Examining the Psychosis Continuum. *Curr Behav Neurosci Rep* **2**, 80-89 (2015).
29. J. van Os, U. Reininghaus, Psychosis as a transdiagnostic and extended phenotype in the general population. *World Psychiatry* **15**, 118-124 (2016).

30. S. K. Keedy, J. L. Reilly, J. R. Bishop, P. J. Weiden, J. A. Sweeney, Impact of antipsychotic treatment on attention and motor learning systems in first-episode schizophrenia. *Schizophr Bull* **41**, 355-365 (2015).
31. M. Devrim-Ucok, H. Y. Keskin-Ergen, A. Ucok, Mismatch negativity at acute and post-acute phases of first-episode schizophrenia. *Eur Arch Psychiatry Clin Neurosci* **258**, 179-185 (2008).
32. K. C. Chen *et al.*, P300 waveform and dopamine transporter availability: a controlled EEG and SPECT study in medication-naive patients with schizophrenia and a meta-analysis. *Psychol Med* **44**, 2151-2162 (2014).
33. J. Todd *et al.*, Mismatch negativity (MMN) to pitch change is susceptible to order-dependent bias. *Front Neurosci-Switz* **8**, (2014).
34. J. Todd, A. Provost, G. Cooper, Lasting first impressions: a conservative bias in automatic filters of the acoustic environment. *Neuropsychologia* **49**, 3399-3405 (2011).
35. R. P. N. Rao, D. H. Ballard, Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci* **2**, 79-87 (1999).
36. M. I. Garrido, J. M. Kilner, K. E. Stephan, K. J. Friston, The mismatch negativity: a review of underlying mechanisms. *Clin Neurophysiol* **120**, 453-463 (2009).
37. M. I. Garrido, J. M. Kilner, S. J. Kiebel, K. J. Friston, Dynamic causal modeling of the response to frequency deviants. *J Neurophysiol* **101**, 2620-2631 (2009).
38. N. Barascud, M. T. Pearce, T. D. Griffiths, K. J. Friston, M. Chait, Brain responses in humans reveal ideal observer-like sensitivity to complex acoustic patterns. *Proc Natl Acad Sci U S A* **113**, E616-625 (2016).
39. R. Dale, N. D. Duran, J. R. Morehead, Prediction during statistical learning, and implications for the implicit/explicit divide. *Adv Cogn Psychol* **8**, 196-209 (2012).
40. A. Tavano, A. Widmann, A. Bendixen, N. Trujillo-Barreto, E. Schroger, Temporal regularity facilitates higher-order sensory predictions in fast auditory sequences. *Eur J Neurosci* **39**, 308-318 (2014).
41. A. Bendixen, U. Roeber, E. Schroger, Regularity extraction and application in dynamic auditory stimulus sequences. *J Cogn Neurosci* **19**, 1664-1677 (2007).
42. J. Mattout, M. Pélégrini-Issac, L. Garnero, H. Benali, Multivariate source prelocalization (MSP): use of functionally informed basis functions for better conditioning the MEG inverse problem. *NeuroImage* **26**, 356-373 (2005).

43. A. E. Radillo, A. Veliz-Cuba, K. Josic, Z. P. Kilpatrick, Evidence Accumulation and Change Rate Inference in Dynamic Environments. *Neural Comput*, 1-50 (2017).
44. J. Todd *et al.*, What controls gain in gain control? Mismatch negativity (MMN), priors and system biases. *Brain Topogr* **27**, 578-589 (2014).
45. G. Barbalat, N. Bazargani, S. J. Blakemore, The Influence of Prior Expectations on Emotional Face Perception in Adolescence. *Cereb Cortex* **23**, 1542-1551 (2013).
46. C. Summerfield, E. Koechlin, A neural representation of prior information during perceptual inference. *Neuron* **59**, 336-347 (2008).
47. R. A. Adams, Q. J. M. Huys, J. P. Roiser, Computational Psychiatry: towards a mathematically informed understanding of mental illness. *J Neurol Neurosur Ps* **87**, 53-63 (2016).
48. A. R. O'Connor, S. Han, I. G. Dobbins, The Inferior Parietal Lobule and Recognition Memory: Expectancy Violation or Successful Retrieval? *J Neurosci* **30**, 2924-2934 (2010).
49. M. Corbetta, G. Patel, G. L. Shulman, The reorienting system of the human brain: From environment to theory of mind. *Neuron* **58**, 306-324 (2008).
50. B. Opitz, T. Rinne, A. Mecklinger, D. Y. von Cramon, E. Schroger, Differential contribution of frontal and temporal cortices to auditory change detection: fMRI and ERP results. *Neuroimage* **15**, 167-174 (2002).
51. M. I. Garrido, J. M. Kilner, S. J. Kiebel, K. E. Stephan, K. J. Friston, Dynamic causal modelling of evoked potentials: a reproducibility study. *Neuroimage* **36**, 571-580 (2007).
52. M. I. Garrido *et al.*, Repetition suppression and plasticity in the human brain. *Neuroimage* **48**, 269-279 (2009).
53. S. J. Kiebel, M. I. Garrido, K. J. Friston, Dynamic causal modelling of evoked responses: the role of intrinsic connections. *Neuroimage* **36**, 332-345 (2007).
54. K. Friston, Functional integration and inference in the brain. *Prog Neurobiol* **68**, 113-143 (2002).
55. M. I. Garrido *et al.*, The functional anatomy of the MMN: a DCM study of the roving paradigm. *Neuroimage* **42**, 936-944 (2008).
56. K. M. Larsen *et al.*, Altered auditory processing and effective connectivity in 22q11. 2 deletion syndrome. *Schizophrenia research*, (2018).
57. A. Meyer-Lindenberg *et al.*, Evidence for abnormal cortical functional connectivity during working memory in schizophrenia. *Am J Psychiatry* **158**, 1809-1817 (2001).

58. K. E. Stephan, K. J. Friston, C. D. Frith, Dysconnection in schizophrenia: from abnormal synaptic plasticity to failures of self-monitoring. *Schizophr Bull* **35**, 509-527 (2009).

59. R. L. Loewy, C. E. Bearden, J. K. Johnson, A. Raine, T. D. Cannon, The prodromal questionnaire (PQ): preliminary validation of a self-report screening measure for prodromal and psychotic syndromes. *Schizophr Res* **79**, 117-125 (2005).

60. L. A. Weber *et al.*, in *Organization for human brain mapping*. (Geneva, Switzerland, 2016).

61. T. D. Wager, M. C. Keller, S. C. Lacey, J. Jonides, Increased sensitivity in neuroimaging analyses using robust regression. *Neuroimage* **26**, 99-113 (2005).

62. Z. Dienes, Bayesian Versus Orthodox Statistics: Which Side Are You On? *Perspect Psychol Sci* **6**, 274-290 (2011).

63. E. J. Wagenmakers *et al.*, Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychon Bull Rev*, (2017).

64. Z. Šidák, Rectangular confidence regions for the means of multivariate normal distributions. *Journal of the American Statistical Association* **62**, 626-633 (1967).

65. J. Todd *et al.*, Deviant matters: duration, frequency, and intensity deviants reveal different patterns of mismatch negativity reduction in early and late schizophrenia. *Biol Psychiatry* **63**, 58-64 (2008).

66. K. Friston *et al.*, Multiple sparse priors for the M/EEG inverse problem. *Neuroimage* **39**, 1104-1120 (2008).

67. V. Litvak, K. Friston, Electromagnetic source reconstruction for group studies. *Neuroimage* **42**, 1490-1498 (2008).

68. M. Kiang, D. L. Braff, J. Sprock, G. A. Light, The relationship between preattentive sensory processing deficits and age in schizophrenia patients. *Clinical Neurophysiology* **120**, 1949-1957 (2009).

69. J. Todd *et al.*, Mismatch negativity (MMN) to pitch change is susceptible to order-dependent bias. *Front Neurosci-Switz* **8**, (2014).

70. J. Todd, A. Provost, G. Cooper, Lasting first impressions: a conservative bias in automatic filters of the acoustic environment. *Neuropsychologia* **49**, 3399-3405 (2011).

71. S. J. Kiebel, M. I. Garrido, R. Moran, C. C. Chen, K. J. Friston, Dynamic causal modeling for EEG and MEG. *Hum Brain Mapp* **30**, 1866-1876 (2009).

72. B. Opitz, T. Rinne, A. Mecklinger, D. Y. von Cramon, E. Schroger, Differential contribution of frontal and temporal cortices to auditory change detection: fMRI and ERP results. *Neuroimage* **15**, 167-174 (2002).

73. M. I. Garrido, J. M. Kilner, S. J. Kiebel, K. E. Stephan, K. J. Friston, Dynamic causal modelling of evoked potentials: a reproducibility study. *Neuroimage* **36**, 571-580 (2007).

74. M. I. Garrido *et al.*, The functional anatomy of the MMN: a DCM study of the roving paradigm. *Neuroimage* **42**, 936-944 (2008).

75. W. D. Penny *et al.*, Comparing families of dynamic causal models. *PLoS Comput Biol* **6**, e1000709 (2010).

76. K. J. Preacher, A. F. Hayes, SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behav Res Methods Instrum Comput* **36**, 717-731 (2004).

77. J. M. Dickey, B. Lientz, The weighted likelihood ratio, sharp hypotheses about chances, the order of a Markov chain. *The Annals of Mathematical Statistics* **41**, 214-226 (1970).

Acknowledgements

The authors thank the participants for their time.

Funding:

This work was funded by the Australian Research Council Centre of Excellence for Integrative Brain Function (ARC Centre Grant CE140100007) and a University of Queensland Fellowship (2016000071) to MIG, as well as a University of Queensland International Research Scholarship to RR.

Author contributions:

Author ID designed the paradigm, conducted the experiment, analysed the data and wrote the first draft of the manuscript. Author MIG assisted in the design of the paradigm; authors MIG and RR assisted in the analysis of data. All authors contributed to and have approved the final manuscript.

Competing interests:

Authors declare no competing interests.

Data and materials availability:

The article's supporting data and materials have been made available. Please find raw data files and the behavioural/connectivity scores here:

https://espace.library.uq.edu.au/view/UQ:724759.

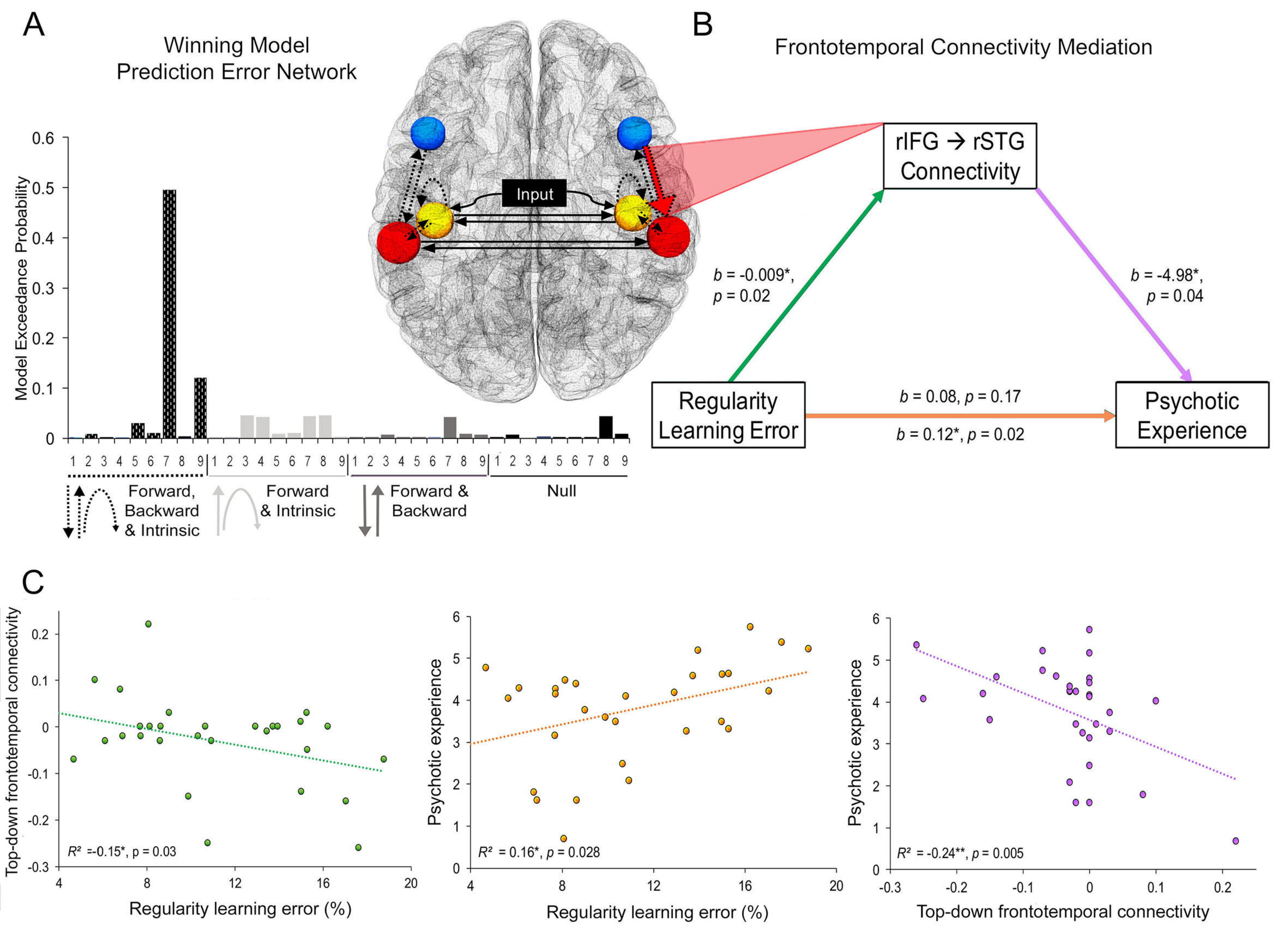
Supplementary Materials

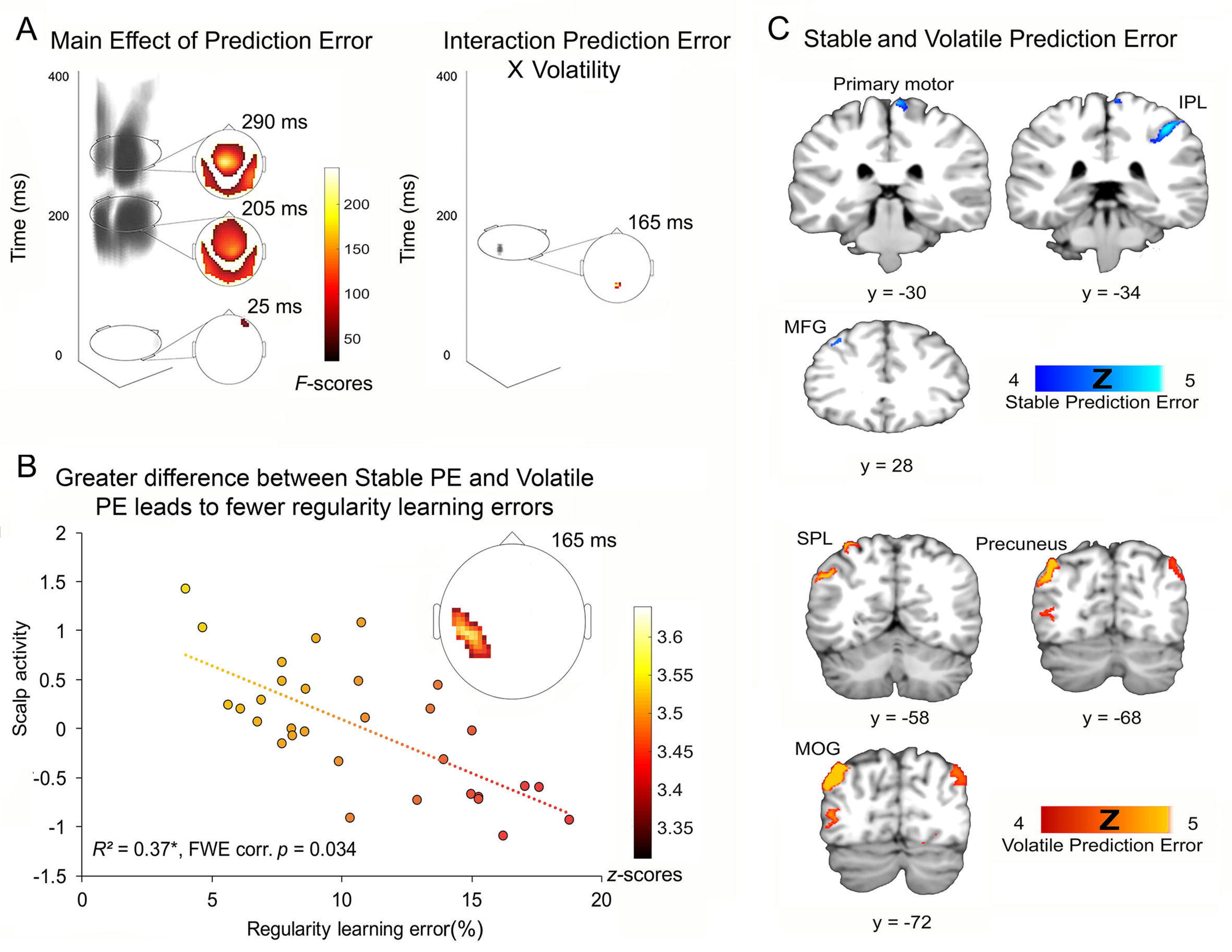
Materials and Methods

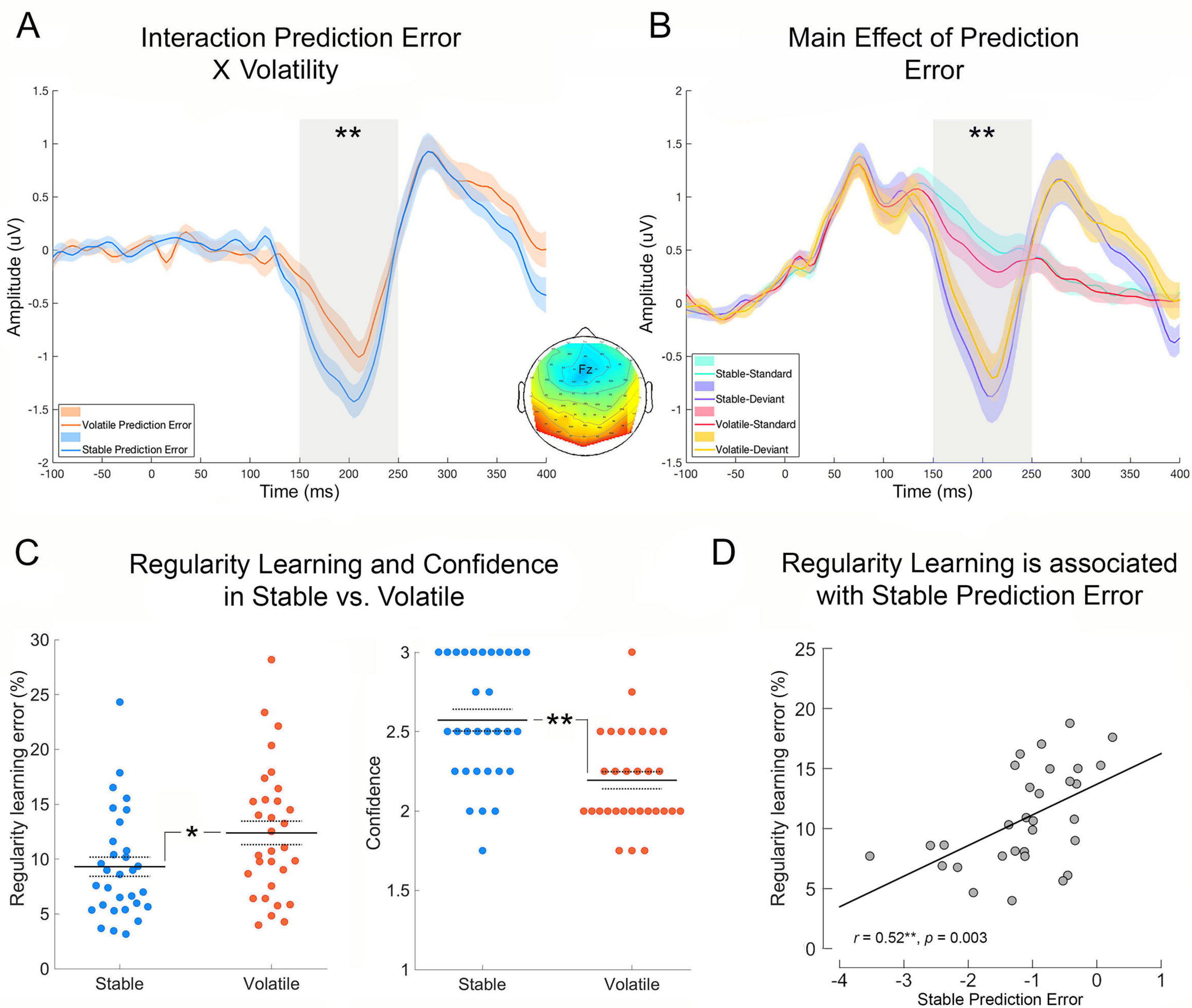
Figures S1-S3

Tables S1-S3

References (59-77)


A
Winning Model
Prediction Error Network
Model Exceedance Probability
0.6
0.5
0.4
0.3
0.2
0.1
0
1 2 3 4 5 6 7 8 9 1 2 3 4 5 6 7 8 9 1 2 3 4 5 6 7 8 9 1 2 3 4 5 6 7 8 9
Forward, Backward & Intrinsic
Forward & Intrinsic
Forward & Backward
Null
Input
B
Frontotemporal Connectivity Mediation
rIFG → rSTG Connectivity
b = -0.009*, p = 0.02
b = -4.98*, p = 0.04
Regularity Learning Error
b = 0.08, p = 0.17
b = 0.12*, p = 0.02
Psychotic Experience
C
Top-down frontotemporal connectivity
R² =-0.15*, p = 0.03
Regularity learning error (%)
Psychotic experience
R² = 0.16*, p = 0.028
Regularity learning error (%)
Psychotic experience
R² = -0.24**, p = 0.005
Top-down frontotemporal connectivity


A
Main Effect of Prediction Error
Time (ms)
400
200
0
290 ms
205 ms
25 ms
200
150
100
50
F-scores
Interaction Prediction Error
X Volatility
Time (ms)
400
200
0
165 ms
C
Stable and Volatile Prediction Error
Primary motor
IPL
y = -30
y = -34
MFG
y = 28
4
Z
5
Stable Prediction Error
SPL
Precuneus
y = -58
y = -68
MOG
y = -72
4
Z
5
Volatile Prediction Error
B
Greater difference between Stable PE and Volatile
PE leads to fewer regularity learning errors
Scalp activity
2
1.5
1
0.5
0
-0.5
-1
-1.5
0
5
10
15
20
Regularity learning error(%)
165 ms
3.6
3.55
3.5
3.45
3.4
3.35
z-scores
R² = 0.37*, FWE corr. p = 0.034


A
Interaction Prediction Error X Volatility
**
Amplitude (uV)
Time (ms)
Volatile Prediction Error
Stable Prediction Error
Fz
B
Main Effect of Prediction Error
**
Amplitude (uV)
Time (ms)
Stable-Standard
Stable-Deviant
Volatile-Standard
Volatile-Deviant
C
Regularity Learning and Confidence in Stable vs. Volatile
Regularity learning error (%)
*
Stable
Volatile
Confidence
**
Stable
Volatile
D
Regularity Learning is associated with Stable Prediction Error
Regularity learning error (%)
r = 0.52**, p = 0.003
Stable Prediction Error