The Bayesian Brain and Predictive Coding as Models for Psychological Distress: A Synthesis of the Evidence

1. Introduction

The "Bayesian brain" and "predictive coding" frameworks propose that the brain is fundamentally a prediction machine, constantly generating and updating models of the world to minimize prediction errors between expected and actual sensory input. These models have gained significant traction in explaining psychological distress and psychiatric disorders, including psychosis, depression, and anxiety. The core idea is that maladaptive inferences—arising from aberrant weighting of prior beliefs versus sensory evidence—can lead to symptoms such as hallucinations, delusions, and affective disturbances. Recent research supports the utility of these frameworks for unifying cognitive, neurobiological, and computational perspectives on mental illness, while also highlighting challenges and the need for further empirical validation (Sterzer et al., 2018; Heinz et al., 2018; Adams et al., 2013; Adams et al., 2014; Gilbert et al., 2022; Corlett et al., 2009; Barrett & Simmons, 2015; Smith et al., 2020; Harding et al., 2024; Qela et al., 2025; Huys et al., 2020; Van De Cruys & Van Dessel, 2021; Petzschner et al., 2017; Fong et al., 2020; Yamashita, 2021; Sterzer et al., 2016).

2. Methods

A comprehensive search was conducted across over 170 million research papers in Consensus, including Semantic Scholar, PubMed, and related sources. The search strategy included 20 targeted queries across 8 thematic groups, focusing on foundational theory, clinical applications, mechanistic decomposition, and critiques of the Bayesian brain and predictive coding in psychological distress. In total, 1,023 papers were identified, 705 were screened, 435 were deemed eligible, and the 50 most relevant papers were included in this review.

Search Strategy

Identification N = 1023 → **Screening** N = 705 → **Eligibility** N = 435 → **Included** N = 50

Identification
- Identified 1010 papers that matched 20 Consensus searches — + 1010 identified
- Identified 13 papers from citation graph exploration — + 13 identified

Screening
- Removed papers with missing abstracts
- Removed duplicates — − 318 removed

Eligibility
- Removed papers with low semantic relevance to each search — − 270 removed

Included
- Selected the top 50 highest quality papers after final ranking — − 385 removed

FIGURE 1 Flow of papers through the search and selection process.

Eight unique search groups were used, spanning foundational theory, clinical application, mechanistic decomposition, and interdisciplinary expansion.

3. Results

3.1. Core Principles and Theoretical Foundations

- **Bayesian Brain Hypothesis:** The brain is conceptualized as an inference machine, using probabilistic (Bayesian) models to combine prior beliefs with sensory evidence, updating beliefs to minimize prediction error (Adams et al., 2013; Adams et al., 2014; Corlett et al., 2009; Barrett & Simmons, 2015; Smith et al., 2020; Petzschner et al., 2017; Yamashita, 2021; Sterzer et al., 2016).
- **Predictive Coding:** This framework posits a hierarchical system where higher-level predictions are compared to lower-level sensory input, and mismatches (prediction errors) drive learning and perception (Sterzer et al., 2018; Adams et al., 2013; Adams et al., 2014; Gilbert et al., 2022; Corlett et al., 2009; Smith et al., 2020; Harding et al., 2024; Qela et al., 2025; Fong et al., 2020; Sterzer et al., 2016).
- **Precision Weighting:** The relative confidence (precision) assigned to priors versus sensory data is crucial; aberrant precision weighting is implicated in various psychiatric symptoms (Sterzer et al., 2018; Adams et al., 2013; Adams et al., 2014; Gilbert et al., 2022; Corlett et al., 2009; Harding et al., 2024; Qela et al., 2025; Sterzer et al., 2016).

3.2. Application to Psychosis and Schizophrenia

- **Psychosis:** Aberrant predictive coding is proposed to underlie psychotic symptoms, with either weakened or overly strong priors leading to hallucinations and delusions. Disrupted precision encoding can result in maladaptive learning and false inferences (Sterzer et al., 2018; Heinz et al., 2018; Adams et al., 2013; Adams et al., 2014; Corlett et al., 2009; Harding et al., 2024; Qela et al., 2025; Fong et al., 2020; Sterzer et al., 2016).
- **Empirical Evidence:** Studies using computational modeling, neuroimaging, and pharmacological manipulations (e.g., NMDA antagonists) support the role of hierarchical prediction errors and precision weighting in psychosis (Sterzer et al., 2018; Heinz et al., 2018; Adams et al., 2013; Corlett et al., 2009; Harding et al., 2024; Qela et al., 2025; Weber et al., 2019; Fong et al., 2020; Sterzer et al., 2016).

3.3. Application to Depression, Anxiety, and Other Disorders

- **Depression:** Disrupted reward prediction error signaling and impaired generative predictions are linked to anhedonia and negative cognitive bias. Predictive coding models explain how maladaptive priors about the self and the world contribute to depressive symptoms (Gilbert et al., 2022; Barrett & Simmons, 2015; Smith et al., 2020; Van De Cruys & Van Dessel, 2021; Manjaly & Iglesias, 2019; Fong et al., 2020; Yamashita, 2021).
- **Anxiety and OCD:** Excessive uncertainty and maladaptive belief updating, as formalized in Bayesian models, are implicated in anxiety and obsessive-compulsive symptoms (Fradkin et al., 2020; Van De Cruys & Van Dessel, 2021; Van Den Bergh et al., 2020; Yamashita, 2021).
- **Transdiagnostic Utility:** Predictive coding deficits are observed across a range of neuropsychiatric disorders, including autism and mood disorders, with disorder-specific patterns of impairment (Qela et al., 2025; Van De Cruys & Van Dessel, 2021; Bolis et al., 2017; Fong et al., 2020).

3.4. Mechanistic and Clinical Implications

- **Neurobiological Correlates:** Predictive coding is linked to specific neural circuits (e.g., hippocampal-prefrontal-striatal networks) and neurotransmitter systems (e.g., dopamine, NMDA receptors) (Heinz et al., 2018; Adams et al., 2013; Adams et al., 2014; Gilbert et al., 2022; Corlett et al., 2009; Harding et al., 2024; Qela et al., 2025; Weber et al., 2019; Fong et al., 2020; Sterzer et al., 2016).
- **Computational Psychiatry:** Bayesian and predictive coding models are increasingly used to formalize and simulate psychiatric symptoms, offering mechanistic insights and potential for personalized interventions (Adams et al., 2015; Smith et al., 2020; Huys et al., 2020; Petzschner et al., 2017; Manjaly & Iglesias, 2019; Huys et al., 2016; Yamashita, 2021).
- **Limitations and Critiques:** Challenges include the heterogeneity of findings, difficulty in empirically constraining models, and the need for more direct clinical applications (Sterzer et al., 2018; Harding et al., 2024; Qela et al., 2025; Haarsma et al., 2020; Fong et al., 2020; Yamashita, 2021).

Key Papers

Paper	Focus	Methodology	Key Results
(Sterzer et al., 2018)	Predictive coding in psychosis	Review, computational	Aberrant precision weighting explains hallucinations, delusions
(Adams et al., 2013)	Computational anatomy of psychosis	Simulation, theory	False inferences from reduced prior precision; links to NMDA/dopamine
(Gilbert et al., 2022)	Predictive coding in depression	Review, neuroimaging	Disrupted prediction error signaling underlies depressive symptoms
(Qela et al., 2025)	Predictive coding in neuropsychiatric disorders	Systematic review	Impairments in predictive coding across schizophrenia, autism, mood disorders
(Yamashita, 2021)	Predictive processing in psychiatry	Review	Theory-driven models explain symptoms via prediction error minimization

FIGURE 2 Comparison of key studies on Bayesian brain and predictive coding models in psychological distress.

Top Contributors

Type	Name	Papers
Author	Karl J. Friston	(Adams et al., 2013; Adams et al., 2014; Adams et al., 2015; Smith et al., 2020; Petzschner et al., 2017; Weber et al., 2019; Manjaly & Iglesias, 2019; FitzGerald et al., 2015; Smith et al., 2019)
Author	P. Fletcher	(Sterzer et al., 2018; Heinz et al., 2018; Griffin & Fletcher, 2017; Corlett et al., 2009; Harding et al., 2024)
Author	Rick A Adams	(Sterzer et al., 2018; Adams et al., 2013; Adams et al., 2014; Adams et al., 2015; Fradkin et al., 2020)
Journal	*Frontiers in Psychiatry*	(Adams et al., 2013; Katthagen et al., 2022; Manjaly & Iglesias, 2019)
Journal	*Biological Psychiatry*	(Sterzer et al., 2018; Petzschner et al., 2017)
Journal	*Schizophrenia bulletin*	(Heinz et al., 2018)

FIGURE 3 Authors & journals that appeared most frequently in the included papers.

4. Discussion

The Bayesian brain and predictive coding frameworks offer a powerful, unifying account of psychological distress, integrating cognitive, neurobiological, and computational levels of explanation. They provide mechanistic models for how aberrant inference and maladaptive learning can give rise to symptoms across a range of psychiatric disorders, from psychosis to depression and anxiety (Sterzer et al., 2018; Adams et al., 2013; Gilbert et al., 2022; Corlett et al., 2009; Barrett & Simmons, 2015; Smith et al., 2020; Harding et al., 2024; Qela et al., 2025; Van De Cruys & Van Dessel, 2021; Petzschner et al., 2017; Fong et al., 2020; Yamashita, 2021; Sterzer et al., 2016). Empirical studies using neuroimaging, computational modeling, and pharmacological interventions support key predictions of these models, such as the role of precision weighting and hierarchical prediction errors (Sterzer et al., 2018; Heinz et al., 2018; Adams et al., 2013; Gilbert et al., 2022; Corlett et al., 2009; Harding et al., 2024; Qela et al., 2025; Weber et al., 2019; Fong et al., 2020; Sterzer et al., 2016).

However, the field faces challenges, including heterogeneity in findings (e.g., whether priors are weakened or strengthened in psychosis), the need for more precise empirical constraints, and the translation of computational models into clinical practice (Sterzer et al., 2018; Harding et al., 2024; Qela et al., 2025; Haarsma et al., 2020; Fong et al., 2020; Yamashita, 2021). There is also ongoing debate about the specificity of predictive coding deficits to particular disorders and the best ways to measure and manipulate these processes in humans.

Claims and Evidence Table

Claim	Evidence Strength	Reasoning	Papers
Aberrant predictive coding underlies psychosis	Strong	Strong theoretical and empirical support, especially for hallucinations and delusions	(Sterzer et al., 2018; Heinz et al., 2018; Adams et al., 2013; Adams et al., 2014; Corlett et al., 2009; Harding et al., 2024; Qela et al., 2025; Fong et al., 2020; Sterzer et al., 2016)
Predictive coding explains depressive and anxiety symptoms	Strong	Disrupted prediction error signaling and maladaptive priors linked to symptoms	(Gilbert et al., 2022; Barrett & Simmons, 2015; Smith et al., 2020; Fradkin et al., 2020; Van De Cruys & Van Dessel, 2021; Manjaly & Iglesias, 2019; Fong et al., 2020; Yamashita, 2021)
Precision weighting is a key mechanism in psychiatric symptoms	Strong	Computational and neurobiological evidence across disorders	(Sterzer et al., 2018; Adams et al., 2013; Adams et al., 2014; Gilbert et al., 2022; Corlett et al., 2009; Harding et al., 2024; Qela et al., 2025; Sterzer et al., 2016)
Predictive coding deficits are transdiagnostic	Moderate	Impairments observed in schizophrenia, autism, mood disorders	(Qela et al., 2025; Van De Cruys & Van Dessel, 2021; Bolis et al., 2017; Fong et al., 2020)
Computational models offer mechanistic insights and clinical potential	Moderate	Used for simulation, diagnosis, and intervention development	(Adams et al., 2015; Smith et al., 2020; Huys et al., 2020; Petzschner et al., 2017; Manjaly & Iglesias, 2019; Huys et al., 2016; Yamashita, 2021)
Empirical constraints and clinical translation remain challenging	Moderate	Heterogeneity, measurement, and application issues persist	(Sterzer et al., 2018; Harding et al., 2024; Qela et al., 2025; Haarsma et al., 2020; Fong et al., 2020; Yamashita, 2021)

FIGURE Key claims and support evidence identified in these papers.

5. Conclusion

The Bayesian brain and predictive coding models provide a robust, mechanistic framework for understanding psychological distress, with strong support across psychosis, depression, and other disorders. They unify cognitive, neurobiological, and computational perspectives, though further empirical and clinical work is needed to refine and apply these models.

Research Gaps

Disorder/Context	Psychosis	Depression	Anxiety/OCD	Autism	Clinical Application
Predictive coding deficits	8	6	5	4	3
Precision weighting mechanisms	7	5	4	3	2
Computational model validation	6	5	3	2	2

FIGURE Matrix of research topics and study attributes, highlighting areas with fewer studies.

Open Research Questions

Question	Why
How can predictive coding models be empirically constrained and validated in clinical populations?	Empirical validation is needed for clinical translation and mechanistic specificity.
What are the disorder-specific versus transdiagnostic features of predictive coding deficits?	Clarifying specificity will improve diagnosis and personalized intervention.
How can computational models inform the development of targeted treatments for psychological distress?	Bridging theory and practice could enhance treatment efficacy and precision.

FIGURE Key open research questions for future investigation.

In summary, Bayesian brain and predictive coding models are highly influential in understanding psychological distress, offering a unifying and mechanistic account, but require further empirical and clinical development for full translational impact.

These papers were sourced and synthesized using Consensus, an AI-powered search engine for research. Try it at https://consensus.app

References

Sterzer, P., Adams, R., Fletcher, P., Frith, C., Lawrie, S., Muckli, L., Petrovic, P., Uhlhaas, P., Voss, M., & Corlett, P. (2018). The Predictive Coding Account of Psychosis. *Biological Psychiatry*, 84, 634 - 643. https://doi.org/10.1016/j.biopsych.2018.05.015

Heinz, A., Murray, G., Schlagenhauf, F., Sterzer, P., Grace, A., & Waltz, J. (2018). Towards a Unifying Cognitive, Neurophysiological, and Computational Neuroscience Account of Schizophrenia.. *Schizophrenia bulletin*. https://doi.org/10.1093/schbul/sby154

Adams, R., Stephan, K., Brown, H., Frith, C., & Friston, K. (2013). The Computational Anatomy of Psychosis. *Frontiers in Psychiatry*, 4. https://doi.org/10.3389/fpsyt.2013.00047

Adams, R., Brown, H., & Friston, K. (2014). Bayesian inference, predictive coding and delusions. *Avant: Journal of the Philosophical-Interdisciplinary Vanguard*, 51-88. https://doi.org/10.26913/50302014.0112.0004

Gilbert, J., Wusinich, C., & Zarate, C. (2022). A Predictive Coding Framework for Understanding Major Depression. *Frontiers in Human Neuroscience*, 16. https://doi.org/10.3389/fnhum.2022.787495

Griffin, J., & Fletcher, P. (2017). Predictive Processing, Source Monitoring, and Psychosis.. *Annual review of clinical psychology*, 13, 265-289. https://doi.org/10.1146/annurev-clinpsy-032816-045145

Corlett, P., Corlett, P., Frith, C., & Fletcher, P. (2009). From drugs to deprivation: a Bayesian framework for understanding models of psychosis. *Psychopharmacology*, 206, 515 - 530. https://doi.org/10.1007/s00213-009-1561-0

Adams, R., Huys, Q., & Roiser, J. (2015). Computational Psychiatry: towards a mathematically informed understanding of mental illness. *Journal of Neurology, Neurosurgery, and Psychiatry*, 87, 53 - 63. https://doi.org/10.1136/jnnp-2015-310737

Barrett, L., & Simmons, K. (2015). Interoceptive predictions in the brain. *Nature Reviews Neuroscience*, 16, 419-429. https://doi.org/10.1038/nrn3950

Smith, R., Badcock, P., & Friston, K. (2020). Recent advances in the application of predictive coding and active inference models within clinical neuroscience. *Psychiatry and Clinical Neurosciences*, 75. https://doi.org/10.1111/pcn.13138

Harding, J., Wolpe, N., Brugger, S., Navarro, V., Teufel, C., & Fletcher, P. (2024). A new predictive coding model for a more comprehensive account of delusions. *The lancet. Psychiatry*, 11, 295 - 302. https://doi.org/10.1016/S2215-0366(23)00411-X

Qela, B., Damiani, S., De Santis, S., Groppi, F., Pichiecchio, A., Asteggiano, C., Brondino, N., Monteleone, A., Grassi, L., Politi, P., Fusar-Poli, P., & Fusar-Poli, L. (2025). Predictive coding in neuropsychiatric disorders: A systematic transdiagnostic review. *Neuroscience & Biobehavioral Reviews*, 169. https://doi.org/10.1016/j.neubiorev.2025.106020

Huys, Q., Browning, M., Paulus, M., & Frank, M. (2020). Advances in the computational understanding of mental illness. *Neuropsychopharmacology*, 46, 3-19. https://doi.org/10.1038/s41386-020-0746-4

Haarsma, J., Kok, P., & Browning, M. (2020). The promise of layer-specific neuroimaging for testing predictive coding theories of psychosis. *Schizophrenia Research*, 245, 68 - 76. https://doi.org/10.1016/j.schres.2020.10.009

Fradkin, I., Adams, R., Parr, T., Roiser, J., & Huppert, J. (2020). Searching for an anchor in an unpredictable world: A computational model of obsessive compulsive disorder.. *Psychological review*. https://doi.org/10.1037/rev0000188

Van De Cruys, S., & Van Dessel, P. (2021). Mental distress through the prism of predictive processing theory.. *Current opinion in psychology*, 41, 107-112. https://doi.org/10.1016/j.copsyc.2021.07.006

Bolis, D., Balsters, J., Wenderoth, N., Becchio, C., & Schilbach, L. (2017). Beyond Autism: Introducing the Dialectical Misattunement Hypothesis and a Bayesian Account of Intersubjectivity. *Psychopathology*, 50, 355 - 372. https://doi.org/10.1159/000484353

Petzschner, F., Weber, L., Gard, T., & Stephan, K. (2017). Computational Psychosomatics and Computational Psychiatry: Toward a Joint Framework for Differential Diagnosis. *Biological Psychiatry*, 82, 421-430. https://doi.org/10.1016/j.biopsych.2017.05.012

Katthagen, T., Fromm, S., Wieland, L., & Schlagenhauf, F. (2022). Models of Dynamic Belief Updating in Psychosis—A Review Across Different Computational Approaches. *Frontiers in Psychiatry*, 13. https://doi.org/10.3389/fpsyt.2022.814111

Weber, L., Diaconescu, A., Mathys, C., Schmidt, A., Kometer, M., Vollenweider, F., & Stephan, K. (2019). Ketamine Affects Prediction Errors about Statistical Regularities: A Computational Single-Trial Analysis of the Mismatch Negativity. *The Journal of Neuroscience*, 40, 5658 - 5668. https://doi.org/10.1101/528372

Manjaly, Z., & Iglesias, S. (2019). A Computational Theory of Mindfulness Based Cognitive Therapy from the "Bayesian Brain" Perspective. *Frontiers in Psychiatry*, 11. https://doi.org/10.3389/fpsyt.2020.00404

Fong, C., Law, W., Uka, T., & Koike, S. (2020). Auditory Mismatch Negativity Under Predictive Coding Framework and Its Role in Psychotic Disorders. *Frontiers in Psychiatry*, 11. https://doi.org/10.3389/fpsyt.2020.557932

Huys, Q., Maia, T., & Frank, M. (2016). Computational psychiatry as a bridge from neuroscience to clinical applications. *Nature Neuroscience*, 19, 404-413. https://doi.org/10.1038/nn.4238

Van Den Bergh, O., Brosschot, J., Critchley, H., Thayer, J., & Ottaviani, C. (2020). Better Safe Than Sorry: A Common Signature of General Vulnerability for Psychopathology. *Perspectives on Psychological Science*, 16, 225 - 246. https://doi.org/10.1177/1745691620950690

FitzGerald, T., Schwartenbeck, P., Moutoussis, M., Dolan, R., & Friston, K. (2015). Active Inference, Evidence Accumulation, and the Urn Task. *Neural Computation*, 27, 306-328. https://doi.org/10.1162/NECO_a_00699

Yamashita, Y. (2021). Psychiatric disorders as failures in the prediction machine. *Psychiatry and Clinical Neurosciences*, 75, 1 - 2. https://doi.org/10.1111/pcn.13173

Sterzer, P., Mishara, A., Voss, M., & Heinz, A. (2016). Thought Insertion as a Self-Disturbance: An Integration of Predictive Coding and Phenomenological Approaches. *Frontiers in Human Neuroscience*, 10. https://doi.org/10.3389/fnhum.2016.00502

Smith, R., Parr, T., & Friston, K. (2019). Simulating Emotions: An Active Inference Model of Emotional State Inference and Emotion Concept Learning. *Frontiers in Psychology*, 10. https://doi.org/10.3389/fpsyg.2019.02844