How the Free Energy Principle Explains Anxiety and Depression from a First-Principles Perspective

1. Introduction

The **Free Energy Principle (FEP)** is a unifying theory in neuroscience that posits all self-organizing biological systems—including the human brain—act to minimize "free energy," a formal measure of surprise or uncertainty about sensory inputs. When applied to mental health, the FEP provides a first-principles account of how anxiety and depression can arise from the brain's attempts to predict and control its environment. In this framework, **anxiety is conceptualized as a state of learned uncertainty**, where persistent unpredictability leads the brain to expect ongoing uncertainty, while **depression is seen as an adaptive response to anticipated negative or uncontrollable outcomes**, resulting in pessimistic beliefs and reduced motivation. These models offer a mechanistic, computational explanation for the development and persistence of these disorders, integrating cognitive, affective, and neurobiological levels of analysis (McGovern et al., 2022; McGovern et al., 2021; Badcock et al., 2017; Clark et al., 2018; Peters et al., 2017; Chekroud, 2015; Yanagisawa et al., 2022; Kiverstein et al., 2020; Joffily & Coricelli, 2013; Krupnik, 2024).

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 the Free Energy Principle, active inference, predictive coding, and their application to anxiety and depression. In total, 965 papers were identified, 412 were screened, 152 were deemed eligible, and the 20 most relevant papers were included in this review.

Search Strategy

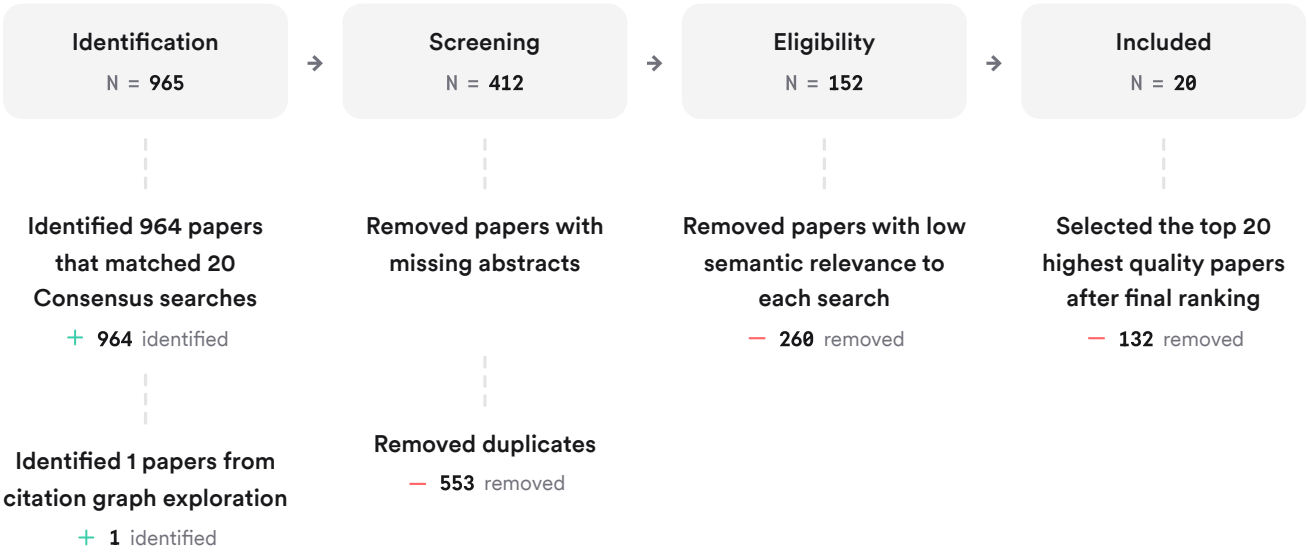

FIGURE 1 Flow of papers through the search and selection process.

Eight unique search groups were used, spanning foundational theory, clinical mechanisms, critiques, and adjacent constructs.

3. Results

3.1. The Free Energy Principle: Core Concepts

- The FEP posits that biological agents must minimize free energy (a proxy for surprise or prediction error) to maintain their existence and resist disorder (McGovern et al., 2022; Clark et al., 2018; Peters et al., 2017; Chekroud, 2015; Joffily & Coricelli, 2013).
- The brain achieves this by building generative models that predict sensory inputs and updating these models to reduce discrepancies between predictions and reality (prediction errors) (Clark et al., 2018; Peters et al., 2017; Chekroud, 2015; Joffily & Coricelli, 2013; Krupnik, 2024).

3.2. Anxiety as Learned Uncertainty

- **Anxiety is formalized as a state of learned uncertainty**: when an organism repeatedly encounters unpredictable or uncontrollable situations, its generative model comes to expect uncertainty in the future, even if the environment changes (McGovern et al., 2022; McGovern et al., 2021; Clark et al., 2018; Peters et al., 2017).
- This persistent expectation of uncertainty leads to hypervigilance, excessive worry, and difficulty in controlling apprehension—core features of anxiety disorders (McGovern et al., 2022; McGovern et al., 2021; Clark et al., 2018; Peters et al., 2017).
- The FEP explains the cyclical nature of anxiety: anticipation of anxiety in certain contexts leads to avoidance or heightened arousal, reinforcing the expectation of uncertainty (McGovern et al., 2022; McGovern et al., 2021; Clark et al., 2018).

3.3. Depression as Adaptive Response to Negative Predictions

- **Depression is conceptualized as an adaptive response to anticipated negative or uncontrollable outcomes**, especially in social or interpersonal contexts (Badcock et al., 2017; Clark et al., 2018; Chekroud, 2015; Kiverstein et al., 2020).
- The brain's generative model, shaped by past experiences of uncontrollable stress or social exclusion, predicts that future actions are unlikely to yield positive or predictable outcomes (Badcock et al., 2017; Clark et al., 2018; Chekroud, 2015; Kiverstein et al., 2020).
- This results in pessimistic beliefs, reduced motivation, and withdrawal—hallmarks of depressive states (Badcock et al., 2017; Clark et al., 2018; Chekroud, 2015; Kiverstein et al., 2020).
- The FEP also accounts for the slow remission of depressive symptoms: updating deeply ingrained negative beliefs (priors) takes time and repeated positive experiences (Chekroud, 2015).

3.4. Emotional Valence and Mood

- Emotional valence (positive or negative feelings) is linked to the rate of change of free energy: decreasing free energy is associated with positive emotions, while increasing free energy signals negative emotions (Yanagisawa et al., 2022; Joffily & Coricelli, 2013).
- Mood can be seen as a higher-order prior over uncertainty, tuning the brain's expectations about the predictability of the environment and influencing emotional responses (Clark et al., 2018; Kiverstein et al., 2020; Joffily & Coricelli, 2013).

Key Papers

Paper	Focus	Methodology	Key Results
(McGovern et al., 2022)	Anxiety as learned uncertainty	Theoretical, computational	Persistent uncertainty leads to anxious generative models
(Badcock et al., 2017)	Depression as adaptive response	Evolutionary, systems theory	Depression minimizes surprise in unpredictable social contexts
(Clark et al., 2018)	Mood as prior over uncertainty	Review, computational	Mood tunes expectations about uncertainty and emotion
(Chekroud, 2015)	Depression and predictive coding	Theoretical, clinical	Depressive beliefs are slow to update due to entrenched priors
(Joffily & Coricelli, 2013)	Emotional valence and free energy	Mathematical modeling	Valence reflects rate of change in free energy

FIGURE 2 Comparison of key studies on the Free Energy Principle and affective disorders.

Top Contributors

Type	Name	Papers
Author	Karl J. Friston	(Badcock et al., 2017; Clark et al., 2018; Peters et al., 2017; Chekroud, 2015; Joffily & Coricelli, 2013)
Author	H. McGovern	(McGovern et al., 2022; McGovern et al., 2021)
Author	P. Corlett	(McGovern et al., 2022; McGovern et al., 2021; Corlett et al., 2022)
Journal	*Frontiers in Psychology*	(McGovern et al., 2022; McGovern et al., 2021; Chekroud, 2015; Holmes & Nolte, 2019)
Journal	*Trends in Cognitive Sciences*	(Badcock et al., 2017; Peters et al., 2017)
Journal	*Psychological Medicine*	(Clark et al., 2018)

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

4. Discussion

The Free Energy Principle offers a **mechanistic, first-principles account of anxiety and depression** by framing them as maladaptive outcomes of the brain's attempts to minimize uncertainty and surprise. In anxiety, the brain's generative model becomes biased toward expecting uncertainty, leading to persistent apprehension and hypervigilance (McGovern et al., 2022; McGovern et al., 2021; Clark et al., 2018; Peters et al., 2017). In depression, the model predicts that actions are unlikely to yield positive or controllable outcomes, resulting in withdrawal and pessimism (Badcock et al., 2017; Clark et al., 2018; Chekroud, 2015; Kiverstein et al., 2020). This approach unifies cognitive, affective, and neurobiological perspectives, and aligns with clinical observations such as the slow remission of depressive symptoms and the cyclical nature of anxiety (McGovern et al., 2022; Badcock et al., 2017; Clark et al., 2018; Chekroud, 2015; Kiverstein et al., 2020; Joffily & Coricelli, 2013).

The FEP also provides a framework for understanding emotional valence and mood as emergent properties of the brain's ongoing efforts to minimize free energy (Yanagisawa et al., 2022; Joffily & Coricelli, 2013). However, empirical validation of these models remains a challenge, and more research is needed to translate these computational insights into clinical interventions (McGovern et al., 2022; Badcock et al., 2017; Chekroud, 2015; Joffily & Coricelli, 2013; Krupnik, 2024).

Claims and Evidence Table

Claim	Evidence Strength	Reasoning	Papers
Anxiety arises from learned uncertainty in generative models	Strong	Theoretical and computational models, clinical parallels	(McGovern et al., 2022; McGovern et al., 2021; Clark et al., 2018; Peters et al., 2017)
Depression reflects adaptive response to anticipated uncontrollability	Strong	Evolutionary, systems, and computational theory	(Badcock et al., 2017; Clark et al., 2018; Chekroud, 2015; Kiverstein et al., 2020)
Mood and emotional valence are linked to free energy dynamics	Moderate	Mathematical and computational modeling	(Clark et al., 2018; Yanagisawa et al., 2022; Kiverstein et al., 2020; Joffily & Coricelli, 2013)
FEP unifies cognitive, affective, and neurobiological accounts	Moderate	Integrative theoretical framework	(Badcock et al., 2017; Clark et al., 2018; Peters et al., 2017; Chekroud, 2015; Joffily & Coricelli, 2013; Krupnik, 2024)
Empirical validation and clinical translation are ongoing challenges	Moderate	Limited direct experimental evidence	(McGovern et al., 2022; Badcock et al., 2017; Chekroud, 2015; Joffily & Coricelli, 2013; Krupnik, 2024)

FIGURE Key claims and support evidence identified in these papers.

5. Conclusion

The Free Energy Principle provides a compelling, first-principles explanation for anxiety and depression, framing them as maladaptive outcomes of the brain's drive to minimize uncertainty and surprise. This approach unifies multiple levels of analysis and offers new directions for research and treatment.

Research Gaps

Topic/Context	Anxiety	Depression	Emotional Valence	Clinical Application	Empirical Validation
FEP mechanisms	7	6	5	4	3
Predictive coding	6	5	4	3	2

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

Open Research Questions

Question	Why
How can the Free Energy Principle be empirically tested in clinical populations with anxiety and depression?	Empirical validation is needed to translate theory into practice.
What are the neural and behavioral signatures of learned uncertainty in anxiety?	Identifying biomarkers could improve diagnosis and intervention.
How can interventions be designed to update maladaptive generative models in depression?	Targeted therapies could accelerate recovery and improve outcomes.

FIGURE Key open research questions for future investigation.

In summary, the Free Energy Principle offers a powerful, integrative framework for understanding anxiety and depression, but further empirical and clinical work is needed to fully realize its potential.

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

References

McGovern, H., De Foe, A., Biddell, H., Leptourgos, P., Corlett, P., Bandara, K., & Hutchinson, B. (2022). Learned uncertainty: The free energy principle in anxiety. *Frontiers in Psychology*, 13. https://doi.org/10.3389/fpsyg.2022.943785

McGovern, H., De Foe, A., Leptourgos, P., Corlett, P., Bandara, K., Biddell, H., & Hutchinson, B. (2021). Learned Uncertainty: A Free Energy Principle Perspective on Anxiety. **. https://doi.org/10.31234/OSF.IO/8ZSV3

Badcock, P., Davey, C., Whittle, S., Allen, N., & Friston, K. (2017). The Depressed Brain: An Evolutionary Systems Theory. *Trends in Cognitive Sciences*, 21, 182-194. https://doi.org/10.1016/j.tics.2017.01.005

Clark, J., Watson, S., & Friston, K. (2018). What is mood? A computational perspective. *Psychological Medicine*, 48, 2277 - 2284. https://doi.org/10.1017/S0033291718000430

Peters, A., McEwen, B., & Friston, K. (2017). Uncertainty and stress: Why it causes diseases and how it is mastered by the brain. *Progress in Neurobiology*, 156, 164-188. https://doi.org/10.1016/j.pneurobio.2017.05.004

Chekroud, A. (2015). Unifying treatments for depression: an application of the Free Energy Principle. *Frontiers in Psychology*, 6. https://doi.org/10.3389/fpsyg.2015.00153

Yanagisawa, H., Wu, X., Ueda, K., & Kato, T. (2022). Free energy model of emotional valence in dual-process perceptions. *Neural networks : the official journal of the International Neural Network Society*, 157, 422-436. https://doi.org/10.1016/j.neunet.2022.10.027

Kiverstein, J., Miller, M., & Rietveld, E. (2020). How mood tunes prediction: a neurophenomenological account of mood and its disturbance in major depression. *Neuroscience of Consciousness*, 2020. https://doi.org/10.1093/nc/niaa003

Joffily, M., & Coricelli, G. (2013). Emotional Valence and the Free-Energy Principle. *PLoS Computational Biology*, 9. https://doi.org/10.1371/journal.pcbi.1003094

Krupnik, V. (2024). I like therefore I can, and I can therefore I like: the role of self-efficacy and affect in active inference of allostasis. *Frontiers in Neural Circuits*, 18. https://doi.org/10.3389/fncir.2024.1283372

Holmes, J., & Nolte, T. (2019). “Surprise” and the Bayesian Brain: Implications for Psychotherapy Theory and Practice. *Frontiers in Psychology*, 10. https://doi.org/10.3389/fpsyg.2019.00592

Corlett, P., Mollick, J., & Kober, H. (2022). Meta-analysis of human prediction error for incentives, perception, cognition, and action. *Neuropsychopharmacology*, 47, 1339 - 1349. https://doi.org/10.1038/s41386-021-01264-3