

# The $\Psi$ Model: A Five-Dimensional Active-Inference State Space for Mental-State Dynamics

Masayoshi Hirose, M.D.  
m.h.0704@gmail.com

## Abstract

This paper introduces the  $\Psi$  model, a five-dimensional state-space representation of mental states grounded in active inference. In this formulation, an individual's psychological state is represented by a vector  $\Psi = (\Omega, E, \Pi, \Sigma, \Lambda)$ , corresponding to Order, Energy, Precision, Social Coupling, and Learning. The model is proposed as a clinically interpretable coordinate system for representing both ordinary psychological variation and psychopathology within a continuous dynamical space. By linking these dimensions to key constructs in active inference and predictive processing, the  $\Psi$  model aims to bridge phenomenological psychiatric description and computational accounts of mental-state organization and change. This conceptual framework provides a foundation for longitudinal modeling, comparative psychopathology, and computational psychiatry.

**Keywords:** Active inference, Computational psychiatry, State-space model, Variational free energy, Mental-state dynamics,  $\Psi$  model.

## 1 Introduction

Understanding and describing mental states remains a central challenge in psychiatry, psychology, and cognitive science. Clinical psychiatry traditionally relies on categorical diagnostic systems such as the Diagnostic and Statistical Manual of Mental Disorders (DSM) and the International Classification of Diseases (ICD). While these frameworks are useful for clinical communication and epidemiology, they often struggle to capture the continuous, multidimensional, and dynamic nature of psychological states. Patients frequently present with heterogeneous symptom patterns, overlapping diagnostic features, and changes over time, suggesting the need for a more flexible representational framework.

Recent developments in computational neuroscience have increasingly framed brain function in terms of predictive processing and active inference. Within this perspective, cognition and behavior are understood as emerging from the updating of internal generative models and the minimization of prediction error through perception, action, and learning. These ideas have also been extended to psychiatry, where disorders have been interpreted as disturbances in processes such as precision weighting, belief updating, and policy selection.

Despite these advances, a simple and clinically interpretable coordinate system for describing mental states in a unified manner remains lacking. In particular, clinical practice often requires a representation that can jointly capture psychological organization, behavioral activation, social relatedness, and adaptive learning.

To address this gap, this paper introduces the  $\Psi$  model, a five-dimensional coordinate system for representing mental states within the theoretical framework of active inference. The model is formalized in Section 3, where its variables and mathematical structure are defined. The aim of the present work is to provide a compact conceptual framework that may support the visualization of mental states, comparison of psychiatric conditions, and future integration with computational psychiatry.

The  $\Psi$  model is proposed not as a replacement for diagnostic systems such as DSM or ICD, but as a semantically interpretable coordinate system for representing the underlying configuration and dynamics of mental states. In this view, conventional psychiatric disorders may be understood as characteristic regions, patterns, or trajectories within a continuous  $\Psi$  space rather than as isolated categorical entities.

## **2 Related Work**

### **2.1 Predictive Processing and Active Inference**

Recent advances in theoretical neuroscience increasingly describe brain function within the framework of predictive processing and active inference. In these accounts, the brain is modeled as a hierarchical generative system that continuously predicts sensory inputs and minimizes prediction errors through perception, action, and learning. Active inference, in particular, provides a formal framework in which agents update beliefs about hidden states of the world while selecting actions that minimize expected free energy.

This framework has been widely applied to explain perception, decision making, and social interaction, and has also been proposed as a theoretical foundation for understanding psychiatric disorders. Disturbances in belief updating, precision allocation, and policy selection have been suggested as potential computational mechanisms underlying various forms of psychopathology. The  $\Psi$  model builds on this theoretical background by proposing a low-dimensional coordinate system that captures key psychological variables relevant to active inference dynamics.

### **2.2 Computational Psychiatry**

Computational psychiatry aims to describe psychiatric phenomena using formal models derived from neuroscience, cognitive science, and machine learning. Rather than focusing solely on categorical diagnostic classifications, this approach attempts to characterize mental disorders in terms of alterations in computational parameters governing perception, learning, and decision making.

Several models within computational psychiatry have proposed parameter-based descriptions of psychiatric symptoms, including abnormal precision weighting, altered learning rates, or maladaptive policy selection. However, these approaches often rely on complex model-specific parameters that are difficult to interpret within everyday clinical contexts. The  $\Psi$  model seeks to complement these approaches by introducing a compact and interpretable set of psychological dimensions that can be mapped onto active inference constructs while remaining conceptually accessible for clinical interpretation.

### **2.3 Personality and Dimensional Psychological Models**

Psychology has long attempted to describe individual differences and psychological organization using dimensional models. Personality frameworks such as the Five-Factor Model (Big Five), as well as dimensional approaches to psychopathology such as the Research Domain Criteria (RDoC), propose that psychological phenomena may be better understood as continuous dimensions rather than discrete categories.

While these models primarily describe relatively stable traits or behavioral tendencies, they do not explicitly provide a state-space representation of moment-to-moment mental states. The  $\Psi$  model differs from trait-based approaches by defining a coordinate system for mental states themselves, allowing psychological conditions to be represented as points and trajectories within a multidimensional state space.

## 2.4 Dynamical Systems Approaches to Mental States

Recent work in psychology and psychiatry has also explored dynamical systems approaches to mental processes. In these models, psychological states evolve over time within a phase space, where attractors, transitions, and instabilities may correspond to characteristic patterns of cognition, mood, or behavior.

Such approaches emphasize that mental states are dynamic and context-dependent. The  $\Psi$  model is compatible with this perspective by defining a five-dimensional state space in which mental states can be represented as positions and changes as trajectories.

## 3 The $\Psi$ Model

This section defines the  $\Psi$  model as a five-dimensional state-space representation of mental states. An individual's psychological state is represented as a vector in a continuous space whose axes correspond to core variables of mental organization and adaptation.

### 3.1 Why Five Dimensions?

The selection of five dimensions is motivated by a minimal representational requirement. The aim of the  $\Psi$  model is not merely to summarize symptoms, but to represent mental states as continuous dynamical configurations that remain clinically discriminative across major forms of psychopathology.

To serve this function, the coordinate system should preserve at least five partially independent modes of variation: coherence of the internal generative model, behavioral activation, precision allocation, social coupling, and adaptive updating. Each of these dimensions can vary substantially while the others remain relatively unchanged, and each contributes to clinically meaningful distinctions that are not adequately reducible to the remaining variables.

A representation with four or fewer dimensions would tend to collapse at least one of these axes and thereby merge clinically distinct psychopathological configurations. For example, omitting Order would make it difficult to distinguish highly vigilant but organized anxiety states from forms of psychopathology involving both aberrant salience and disorganization; omitting Energy would blur the distinction between depressive inhibition and stable low-activity states; omitting Learning would obscure the difference between rigid chronic configurations and highly plastic, therapy-responsive ones. These examples are illustrative, but they clarify why fewer than five dimensions would risk losing clinically meaningful distinctions.

For these reasons, the  $\Psi$  model adopts five dimensions not as a formally proven unique decomposition, but as a minimal candidate basis that appears sufficient to preserve several major clinically meaningful distinctions while remaining interpretable as a semantic state-space representation of mental-state dynamics.

### 3.2 Model Overview

The  $\Psi$  model represents mental states as a vector in a five-dimensional state space. The symbol  $\Psi$  (psi) is used to denote the model's state vector, reflecting its longstanding association with psychological and

psychophysical notation:

$$\Psi = (\Omega, E, \Pi, \Sigma, \Lambda)$$

where  $\Omega$  denotes Order,  $E$  denotes Energy,  $\Pi$  denotes Precision,  $\Sigma$  denotes Social Coupling, and  $\Lambda$  denotes Learning.

Together, these variables define a coordinate system for locating an individual's mental state at a given time. Psychological conditions may thus be represented as characteristic configurations within this space, and changes over time may be described as trajectories through it.

Because the full five-dimensional space cannot be visualized directly, lower-dimensional projections can be used for intuition. Figure 1 shows an illustrative two-dimensional projection onto the  $\Pi$ - $E$  plane, in which representative psychopathology-like regions are schematically positioned according to differences in precision weighting and behavioral activation. This projection is not intended as a diagnostic map, but as a conceptual visualization of how characteristic mental-state configurations may occupy different regions of the broader  $\Psi$  space. These regions can be interpreted as attractor-like configurations in this projected space.

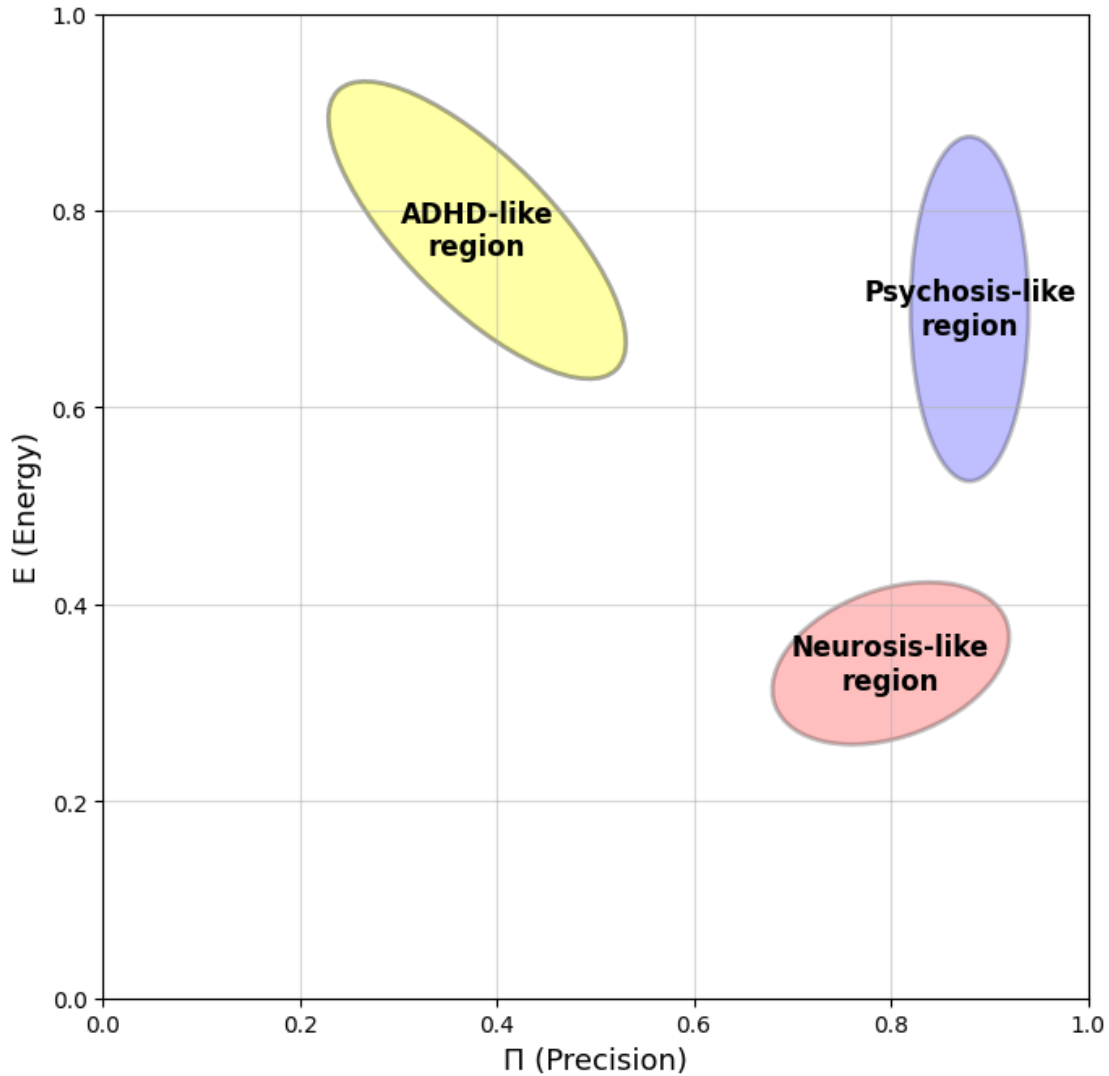


Figure 1: Illustrative psychopathology-like regions in the  $\Pi$ – $E$  projection of the  $\Psi$  space. The horizontal axis represents Precision ( $\Pi$ ), and the vertical axis represents Energy ( $E$ ). The plotted regions are schematic and intended only to illustrate how different classes of mental-state configurations may occupy distinct areas in a lower-dimensional projection of the full five-dimensional  $\Psi$  space.

### 3.3 Variable Definitions

These dimensions are not assumed to be empirically orthogonal; rather, they are proposed as conceptually distinguishable state variables that preserve different clinically relevant modes of variation. In the present conceptual formulation, each dimension is treated as a continuous variable, while its precise scaling and empirical operationalization remain to be determined.

The five variables defining the  $\Psi$  vector correspond to key components of mental organization and behavior.

## $\Omega$ (Order)

$\Omega$  represents the coherence and stability of the internal generative model. This dimension reflects the degree to which an individual's self-model, world model, and temporal narrative remain internally consistent and integrated.

High values of  $\Omega$  correspond to stable identity, coherent thought processes, and consistent interpretation of experience. Low values of  $\Omega$  are associated with fragmentation of thought, instability of self-representation, delusional belief formation, or dissociative phenomena.

Clinically, reduced  $\Omega$  may be observed in schizophrenia spectrum conditions, severe dissociation, or states involving disruption of self-world coherence.

## $E$ (Energy)

$E$  represents the level of behavioral activation driving action generation. This variable captures motivational intensity, exploration behavior, and the propensity to initiate actions.

High values of  $E$  correspond to heightened activity, impulsivity, or manic states, whereas low values correspond to apathy, psychomotor slowing, and depressive states characterized by reduced behavioral initiation.

Conceptually,  $E$  relates to the activation of behavioral policies and may be interpreted in relation to policy activation in reinforcement learning or control gain in control-theoretic models of behavior.

## $\Pi$ (Precision)

$\Pi$  represents the allocation of precision or confidence to prediction errors and sensory signals. In predictive processing frameworks, precision determines the relative influence of incoming information on belief updating.

High values of  $\Pi$  correspond to heightened salience attribution, hypervigilance, anxiety, obsessive attention, or aberrant assignment of meaning to stimuli. Low values correspond to reduced salience, emotional blunting, or diminished responsiveness to environmental signals.

Altered precision allocation has been widely proposed as a computational mechanism underlying several psychiatric conditions, including anxiety disorders and psychosis.

## $\Sigma$ (Social Coupling)

$\Sigma$  represents the degree of coupling between an individual's generative model and those of other agents. This dimension captures social cognition, empathy, social referencing, and sensitivity to interpersonal feedback.

High values of  $\Sigma$  correspond to strong social attunement, conformity, or interpersonal dependency. Low values correspond to social detachment, reduced social motivation, or traits associated with autism spectrum conditions.

From a computational perspective,  $\Sigma$  can be interpreted as the coupling strength between agents in multi-agent generative models.

## $\Lambda$ (Learning)

$\Lambda$  represents the capacity of the generative model to update in response to experience. This dimension reflects belief updating, behavioral adaptation, and the flexibility with which internal models incorporate

new information.

High values of  $\Lambda$  correspond to cognitive flexibility, rapid adaptation, and responsiveness to feedback or therapeutic intervention. Low values correspond to rigidity, persistence of maladaptive beliefs, and repetitive behavioral patterns.

Within computational frameworks,  $\Lambda$  is related to the learning rate in reinforcement learning and to the rate of belief updating in active inference models.

### 3.4 Mathematical Formulation

In the  $\Psi$  model, the psychological state of an individual at time  $t$  is represented as a vector in a five-dimensional continuous space:

$$\Psi(t) = (\Omega(t), E(t), \Pi(t), \Sigma(t), \Lambda(t)) \in \mathbb{R}^5$$

Each dimension represents the instantaneous value of a psychological variable governing mental-state organization and dynamics.

The dynamics of the psychological state are formulated as a gradient descent on a variational free energy landscape, which represents the informational strain between the internal model and the environment:

$$\dot{\Psi} = -\Gamma \nabla_{\Psi} \mathcal{F}(\Psi; o, u)$$

where  $\mathcal{F}$  denotes the variational free energy,  $o$  represents environmental observations,  $u$  denotes environmental inputs or perturbations, and  $\Gamma$  is a diagonal matrix of update rates (or kinetic precisions) for each dimension. In this formulation, the mental state  $\Psi$  naturally evolves toward configurations that minimize surprise and maximize model evidence.

In this formulation, the mental state  $\Psi$  is not a static point but a dynamic trajectory. The state naturally evolves toward configurations (attractor basins) that minimize surprise and maximize model evidence. Psychiatric conditions can thus be formalized as:

1. **Maladaptive Attractors:** Regions of the space where the system becomes trapped, such as the rigid, low-energy state of clinical depression.
2. **Landscape Instability:** A flattening of the free energy landscape (often linked to low  $\Omega$ ), where the mental state drifts inconsistently, as seen in fragmented or psychotic states.

Figure 2 provides an illustrative three-dimensional projection of such dynamics in the  $\Omega$ – $\Pi$ – $E$  subspace. In this schematic example, successive states ( $\Psi_0$ – $\Psi_4$ ) trace a recovery-like trajectory from a more dysregulated configuration toward a more stable and adaptive one. This visualization is intended only to demonstrate how clinically meaningful change may be represented as movement through  $\Psi$  space, rather than as a transition between fixed diagnostic categories.

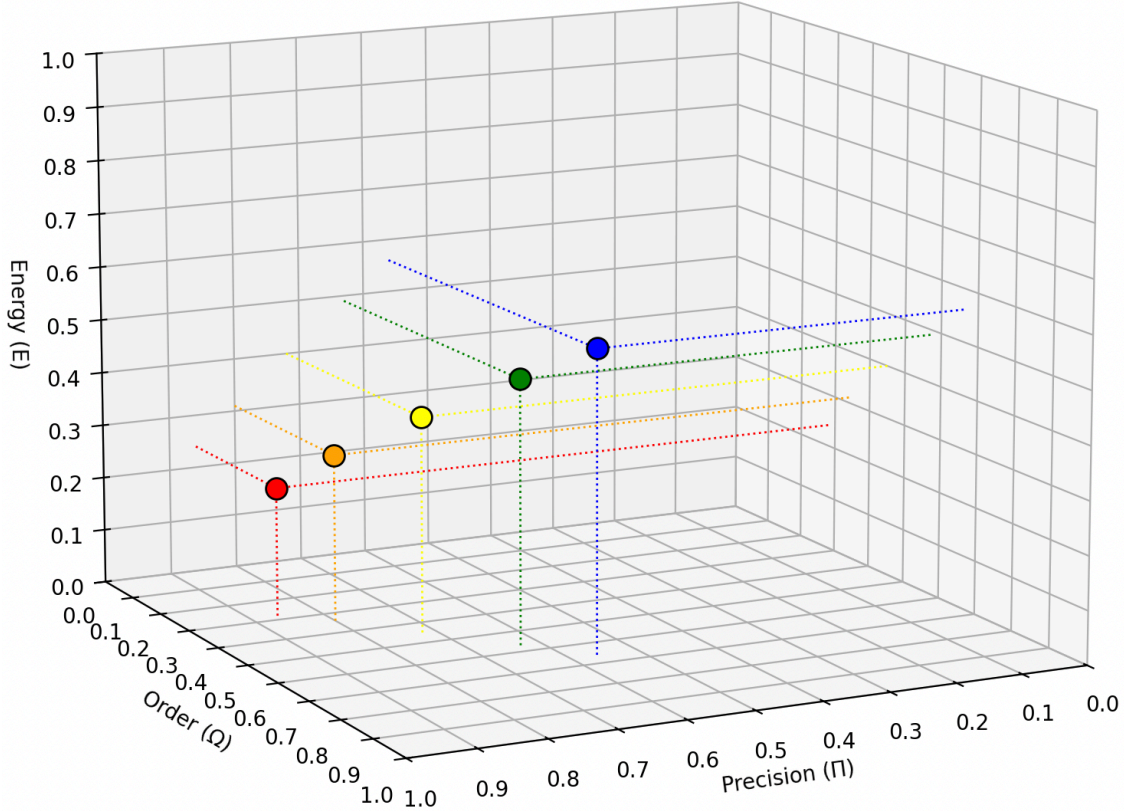


Figure 2: Illustrative recovery trajectory in a three-dimensional projection of  $\Psi$  space. The axes show Order ( $\Omega$ ), Precision ( $\Pi$ ), and Energy ( $E$ ). Successive states ( $\Psi_0$ – $\Psi_4$ ) represent a schematic treatment-related trajectory, with dotted projections indicating each state’s location relative to the coordinate planes.

Environmental inputs  $u$  act as forces that shift the morphology of the free energy landscape itself, either facilitating transitions between states (as in effective therapy) or inducing informational shocks that drive the state toward pathological regions. By defining  $\Psi$  as a formal state vector, we move from a descriptive taxonomy of symptoms to a dynamical systems account of mental-state transitions.

### 3.5 Relation to Active Inference

The  $\Psi$  model is conceptually grounded in the framework of predictive processing and active inference. Within this framework, cognition and behavior emerge from the continuous updating of hierarchical generative models that aim to minimize prediction errors.

The  $\Psi$  vector parameters directly modulate the shape of the free energy landscape  $\mathcal{F}$ . Specifically, Precision ( $\Pi$ ) determines the curvature (steepness) of the landscape around sensory evidence, while Order ( $\Omega$ ) reflects the structural integrity of the prior beliefs that anchor the landscape. Learning ( $\Lambda$ ) and Energy ( $E$ ) govern the velocity and policy-selection gain, respectively, with which the system traverses this manifold. Furthermore, Social Coupling ( $\Sigma$ ) scales the influence of other agents’ generative models on the individual’s own free energy minimization. This links clinical phenomenology to the fundamental principle of self-organization through free-energy minimization.

By organizing these variables within a single state-space representation, the  $\Psi$  model provides an interpretable framework that links clinical psychological descriptions to computational principles derived from

active inference.

## 4 Applications

The  $\Psi$  model may provide a useful conceptual framework for psychiatry, psychotherapy, and the analysis of social systems. The following subsections outline several domains in which this state-space representation may be informative.

### 4.1 Psychiatric Disorders

Within the  $\Psi$  framework, psychiatric conditions may be conceptualized as characteristic regions or configurations within the five-dimensional state space. Rather than representing disorders as discrete categories, the model allows them to be interpreted as patterns of deviation along specific psychological dimensions.

For example, depressive states may be characterized by markedly reduced behavioral energy (low  $E$ ) together with reduced learning (low  $\Lambda$ ), reflecting diminished behavioral activation and reduced responsiveness to environmental feedback. Manic states, in contrast, may involve elevated behavioral energy (high  $E$ ) and increased learning (high  $\Lambda$ ), leading to heightened exploration and rapid belief updating.

Psychotic conditions may involve reductions in internal model stability (low  $\Omega$ ) combined with abnormal precision allocation (high  $\Pi$ ), potentially resulting in unstable belief formation and aberrant attribution of salience to environmental signals. Anxiety states may correspond to elevated precision weighting (high  $\Pi$ ), leading to excessive vigilance toward potential threats.

By mapping psychiatric conditions into the  $\Psi$  state space, the model provides a way to visualize relationships between disorders and to conceptualize symptom transitions as movements within this multidimensional landscape. These examples are intended to be illustrative rather than diagnostic, and the same dimension may contribute to different clinical presentations depending on its interaction with the remaining variables.

In schematic lower-dimensional projections, some of these configurations can be visualized as partially separable regions. Figure 1 illustrates this point in the  $\Pi$ - $E$  plane, where representative psychopathology-like configurations are shown as occupying different zones of precision and behavioral activation.

### 4.2 Psychotherapy

The  $\Psi$  model may also offer a conceptual framework for understanding psychotherapeutic processes. In this context, therapy can be interpreted as an intervention that modifies the position or trajectory of an individual's mental state within the  $\Psi$  state space.

Different therapeutic approaches may influence different dimensions of the model. Behavioral activation strategies may increase behavioral energy ( $E$ ), encouraging the initiation of adaptive actions. Cognitive therapies may influence precision allocation ( $\Pi$ ) by modifying the weighting of interpretations and beliefs. Psychodynamic or narrative-based therapies may contribute to the stabilization and integration of the internal generative model, thereby increasing order ( $\Omega$ ).

Similarly, therapeutic learning processes may enhance learning ( $\Lambda$ ), allowing individuals to update maladaptive beliefs and behavioral patterns more effectively. Interpersonal therapies may influence social coupling ( $\Sigma$ ) by modifying the ways in which individuals relate to and model other agents.

Within this framework, psychotherapy can be conceptualized as guiding individuals toward more adaptive regions of the  $\Psi$  state space.

### 4.3 Social Systems

Although the  $\Psi$  model is defined at the level of individual mental states, the same framework may also provide insights into collective and social dynamics. Human social systems can be viewed as networks of interacting agents whose internal models influence one another through communication, imitation, and shared beliefs.

In this context, the social coupling parameter ( $\Sigma$ ) plays a central role, as it reflects the degree to which individuals integrate information from other agents. Collective phenomena such as group conformity, polarization, and social contagion may emerge from interactions among individuals whose generative models are dynamically coupled.

From this perspective, social systems may be conceptualized as higher-order dynamical structures emerging from the interactions of many agents moving through their respective  $\Psi$  state spaces. Large-scale social stability or instability may therefore reflect collective dynamics across these five dimensions.

## 5 Dynamics of the $\Psi$ Model

The  $\Psi$  model can also be interpreted dynamically. Rather than treating mental states as static positions, the framework allows psychological change to be described as movement through the state space in response to internal regulation, environmental input, and social interaction.

### 5.1 State Dynamics

Formally, mental states are conceptualized as trajectories toward attractor basins within the  $\Psi$  state space. A healthy state corresponds to a flexible yet stable attractor, while psychopathology can be viewed as the system becoming trapped in maladaptive attractors (e.g., the deep, rigid basin of depression or the fragmented, unstable landscape of psychosis). Environmental perturbations  $u$  act by shifting the position of these basins or altering the system's ability to transition between them.

These dynamics may reflect regulatory processes that stabilize internal models, modulate behavioral activation, allocate precision, regulate social engagement, and update beliefs through learning.

As illustrated in Figure 2, such trajectories can be represented as movement through a lower-dimensional projection of the  $\Psi$  space.

### 5.2 Environmental Perturbations

Mental states are continuously influenced by environmental inputs. External events such as social interactions, stressors, rewards, and novel experiences may perturb the system and alter the trajectory of the  $\Psi$  vector.

For example, stressful environments may increase precision weighting ( $\Pi$ ) and behavioral energy ( $E$ ), producing heightened vigilance and defensive behavior. Social support may increase social coupling ( $\Sigma$ ) and stabilize internal model coherence ( $\Omega$ ). Therapeutic interventions may increase learning ( $\Lambda$ ), facilitating adaptive belief updating.

In this sense, environmental influences can be interpreted as forces acting on the  $\Psi$  vector, altering its position within the state space.

The influence of environmental and therapeutic conditions can also be illustrated in simplified time-series form. Figure 3 shows schematic trajectories of Energy ( $E$ ) and Precision ( $\Pi$ ) under four combinations of Order ( $\Omega$ ) and Social Coupling ( $\Sigma$ ). In this example, higher  $\Omega$  and/or higher  $\Sigma$  are associated with distinct recovery-like profiles, suggesting that internal model stability and interpersonal support may shape the time course of behavioral activation and salience regulation during treatment. These curves are illustrative rather than empirical, but they highlight how changes in one subset of  $\Psi$  variables may alter the temporal evolution of others.

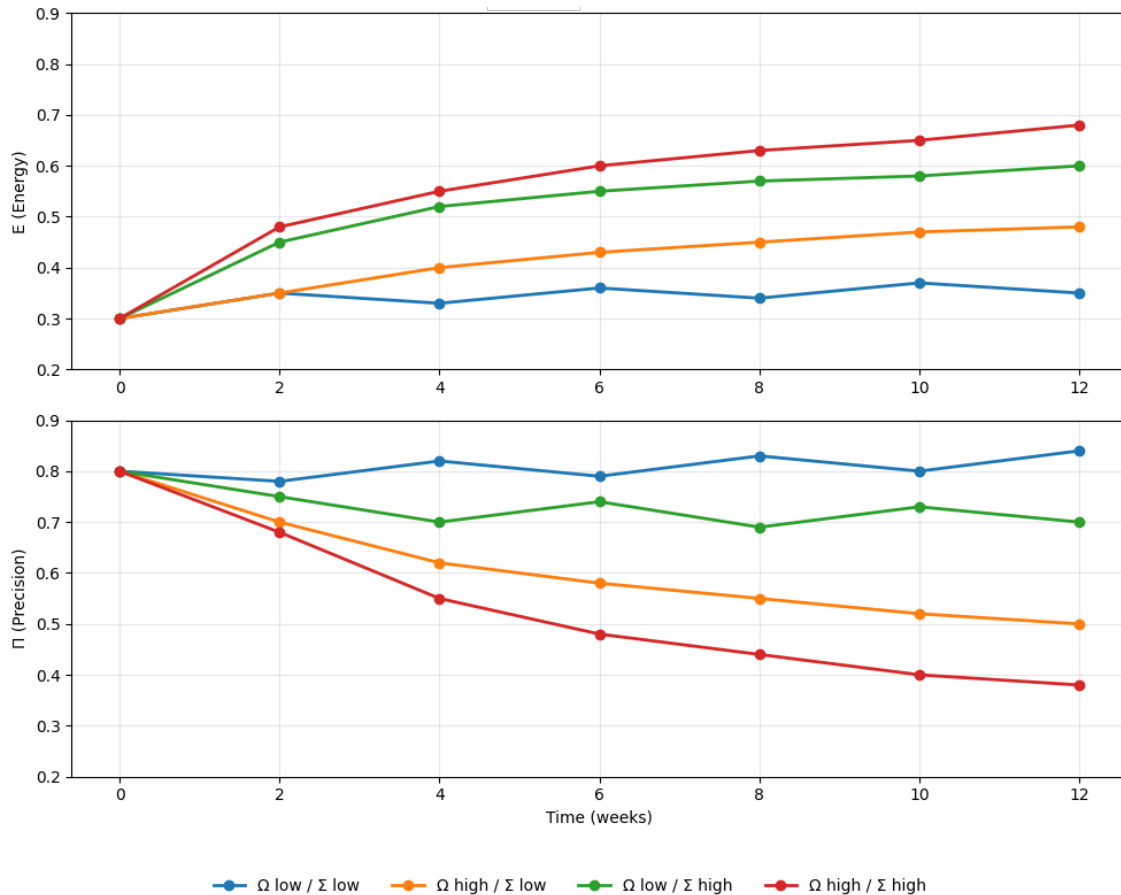


Figure 3: Illustrative time courses of Energy ( $E$ ) and Precision ( $\Pi$ ) during treatment under different combinations of Order ( $\Omega$ ) and Social Coupling ( $\Sigma$ ). The curves are schematic and show how differing levels of internal stability and interpersonal support may shape recovery trajectories over time.

### 5.3 Multi-Agent Interactions

Human mental states do not evolve in isolation but are embedded within networks of interacting agents. Social interaction introduces additional dynamics in which the generative models of different individuals become coupled.

Within the  $\Psi$  framework, this interaction may be captured primarily through the social coupling variable  $\Sigma$ . When coupling is strong, individuals may influence one another's beliefs, emotional states, and behavioral

policies.

At the population level, such interactions may produce collective phenomena including synchronization of emotional states, group polarization, or large-scale social dynamics. These effects can be conceptualized as emergent patterns arising from the interaction of many agents moving through their respective  $\Psi$  state spaces.

## **6 Discussion**

This section considers the theoretical significance of the  $\Psi$  model, its relation to existing frameworks, and its possible clinical relevance.

### **6.1 Theoretical Implications**

A central contribution of the  $\Psi$  model is the proposal that a broad range of psychological phenomena can be represented within a low-dimensional coordinate system. Rather than treating mental states as discrete categories, the model treats them as positions within a continuous space defined by a small set of variables.

Such a representation may help bridge clinical psychiatry, cognitive science, and computational neuroscience. It also provides a natural way to conceptualize psychological change, including symptom emergence, transitions between affective states, and recovery processes.

This perspective is consistent with dynamical systems approaches to cognition, in which mental processes emerge from the interaction of multiple continuously varying parameters.

The present claim is therefore limited but specific. The  $\Psi$  model is not intended to imply that all psychopathology is exhaustively reducible to five variables. Rather, it proposes that five dimensions may constitute a minimal candidate basis for preserving the major clinically relevant distinctions required of a dynamical and semantically interpretable representation of mental states.

### **6.2 Comparison with Existing Models**

Several existing frameworks describe psychological phenomena using dimensional structures. Personality models such as the Five-Factor Model characterize relatively stable traits, while dimensional psychiatric frameworks such as the Research Domain Criteria (RDoC) organize psychopathology across behavioral and neurobiological dimensions.

The  $\Psi$  model differs from these approaches by being explicitly formulated as a state-space representation of current mental state rather than a trait taxonomy. In this sense, it may serve as a conceptual bridge between phenomenological descriptions of mental states and formal models in computational psychiatry.

### **6.3 Clinical Implications**

From a clinical perspective, the  $\Psi$  model may offer a useful way to conceptualize psychiatric conditions within a shared multidimensional space. It may also help in thinking about interventions, since different treatment modalities may influence different dimensions of the  $\Psi$  vector. Although the model remains conceptual at this stage, it may provide a basis for future empirical work aimed at operationalizing these dimensions and mapping clinical observations onto the proposed state space.

## 7 Limitations

While the  $\Psi$  model provides a conceptual framework for representing mental states within a five-dimensional state space, several limitations should be acknowledged.

### 7.1 Lack of Empirical Validation

The present work primarily introduces a theoretical model. Although the dimensions proposed in the  $\Psi$  framework are motivated by concepts from active inference, psychiatry, and cognitive science, empirical validation remains an important direction for future research.

Further studies will be necessary to determine how these variables can be operationalized and measured in real-world psychological or clinical settings.

Preliminary exploratory mapping, including informal AI-assisted semantic classification, suggests that a broad range of common DSM- and ICD-based diagnostic presentations can be meaningfully located within the proposed  $\Psi$  space. However, such exploratory alignment should be regarded as heuristic rather than confirmatory, and systematic empirical validation remains necessary.

### 7.2 Parameter Estimation

Another challenge concerns the estimation of the five variables defining the  $\Psi$  vector. Psychological constructs such as internal model stability, precision allocation, and social coupling may not be directly observable and may require indirect measurement through behavioral, cognitive, or neurobiological indicators.

Developing reliable methods for estimating these parameters will be an important step toward applying the  $\Psi$  model in empirical and clinical contexts.

### 7.3 Model Simplification

The  $\Psi$  model intentionally reduces the complexity of mental states to a small number of core variables. While this low-dimensional representation offers conceptual clarity, it inevitably simplifies the full richness of human cognition and behavior.

Additional variables or hierarchical structures may be required to capture specific psychological phenomena in greater detail.

Despite these limitations, the model provides a starting point for exploring whether a compact state-space representation can meaningfully describe the organization and dynamics of mental states.

## 8 Conclusion

This paper introduced the  $\Psi$  model, a five-dimensional state-space framework for representing mental states in relation to active inference. The model is intended as a compact and interpretable coordinate system for describing mental-state organization and change.

Although still conceptual, the framework may support future work in computational psychiatry and related fields. Possible next steps include empirical operationalization of the five dimensions, digital-twin approaches to longitudinal mental-state tracking, and multi-agent simulations of socially coupled psychological dynamics.

By proposing a state-space representation grounded in active inference, the  $\Psi$  model offers a conceptual basis for further theoretical and empirical work in psychiatry and cognitive science.

## References

- Friston, K. J., Parr, T., & de Vries, B. (2017). The graphical brain: Belief propagation and active inference. *Network Neuroscience*, 1(4), 381–414. doi:10.1162/netn\_a\_00018
- Friston, K. J., Wiese, W., & Hobson, J. A. (2020). Sentience and the origins of consciousness: From Cartesian dualism to Markov blankets. *Entropy*, 22(5), 516. doi:10.3390/e22050516
- Hesp, C., Smith, R., Parr, T., Amrute, M., Corcoran, A. W., & Friston, K. J. (2021). Deeply felt affect: The emergence of valence, emotion, and states of mind in self-organizing systems. *Neural Computation*, 33(2), 398–446. doi:10.1162/neco\_a\_01341
- Huys, Q. J. M., Maia, T. V., & Frank, M. J. (2016). Computational psychiatry as a bridge from neuroscience to clinical applications. *Nature Neuroscience*, 19(3), 404–413. doi:10.1038/nn.4238
- Montague, P. R., Dolan, R. J., Friston, K. J., & Dayan, P. (2012). Computational psychiatry. *Trends in Cognitive Sciences*, 16(1), 72–80. doi:10.1016/j.tics.2011.11.018
- Parr, T., Pezzulo, G., & Friston, K. J. (2022). *Active inference: The free energy principle in mind, brain, and behavior*. MIT Press.
- Powers, A. R., Corlett, P. R., Gleeson, P. L., & Sterling, P. (2017). Neural control of state-dependent action selection. *Nature Reviews Neuroscience*, 18(10), 612–621.
- Ramstead, M. J. D., Friston, K. J., & Hipólito, I. (2020). Is the free-energy principle a formal hierarchy of sciences? *Frontiers in Psychology*, 11, 1361. doi:10.3389/fpsyg.2020.01361
- Smith, R., Badcock, P., & Friston, K. J. (2021). Recent advances in the application of predictive coding and active inference to clinical psychology. *Clinical Psychology Review*, 83, 101944. doi:10.1016/j.cpr.2020.101944
- Stephan, K. E., Binder, E. B., Breakspear, M., Dayan, P., Johnstone, E. C., Meyer-Lindenberg, A., et al. (2016). Charting the landscape of priority problems in psychiatry, part 1: Classification and diagnosis. *The Lancet Psychiatry*, 3(1), 77–83.

## A Appendix: Variable Summary

Variable	Meaning	Computational Interpretation
$\Omega$	Order	stability of generative model
E	Energy	behavioral activation
$\Pi$	Precision	weighting of prediction error
$\Sigma$	Social coupling	coupling with other agents
$\Lambda$	Learning	belief update rate

Figure 4: Summary of the  $\Psi$  model variables.