Cognitive Compensation and Predictive Processing: A Theoretical Integration from an Ontological Perspective

Lei Cai[1] **and Xusheng Zhang**[*1]

[1,2]College of Computer Science and Technology,Zhejiang University,Hangzhou,China

Abstract: As the dominant paradigm in contemporary cognitive science, predictive processing theory has achieved great success in explaining neural computational mechanisms. However, its underlying ‘Free Energy Principle’ faces the dilemma of an ‘ontological explanatory deficit’ and the ‘Dark Room Problem’ due to its excessive universality. Meanwhile, the debate between cognitivist predictive processing and free-energy enactivism also highlights the theoretical rift when integrating the embodied dimension. This paper introduces the ‘cognitive compensation’ theory and the RID model (Regularization R, Information/Integration I, Distress/Demand D) from *Knowing and Saying*, attempting to provide a solid ontological foundation for predictive processing. By anchoring prediction error to distress/demand, hierarchical generative models to structure generation, and precision weighting to regularization, this paper argues that predictive processing is essentially a higher-order compensatory activity of finite beings coping with the pressure of decreasing degree of existence. Within this framework, the brain’s internal representations are no longer mirrors of the objective world, but ‘Compensatory Schemas’ actively constructed by finite beings to suppress distress/demand–this concept is the core critique and transcendence of traditional cognitivist representationalism. Furthermore, this paper takes schizophrenia (abnormal precision weighting) and autism (sensory hypersensitivity) as test fields of ‘compensatory failure’ to verify the theoretical efficacy of this integrative framework. The introduction of this ontological perspective not only provides a fundamental answer to the ‘Dark Room Problem’, but also offers a theoretical way out for the long-standing debate between cognitivism and embodied cognition.

Keywords: Predictive Processing; Free Energy Principle; Cognitive Compensation; RID Model; Embodied Cognition; Dark Room Problem

1 Introduction: The Rise of Predictive Processing and the “Dark Room” Dilemma

In the development of contemporary cognitive science and philosophy of mind, “predictive processing” (PP) has rapidly emerged as a highly inclusive paradigm for the study of mind [2]. This theory overturns the traditional metaphors of “feature detection” and “bottom-up information processing,” arguing that the brain is not a sensory container that passively receives environmental information, but a multilayer probabilistic prediction engine. It actively infers the state of the world and guides action by continuously minimizing the “prediction error” between top-down predictions and bottom-up sensory inputs [9]. Because of its powerful integrative capacity in explaining perception, action, learning, and even psychopathological phenomena, many scholars regard predictive processing as a theoretical cornerstone that may provide a unified framework for the study of the human mind [2].

However, as predictive processing theory extends toward higher-order cognitive phenomena, deep internal fissures within the theory have become increasingly visible. Current predictive-processing models largely remain at the level of describing neural computational mechanisms, such as Bayesian inference and hierarchical predictive coding. To provide a more basic biophysical motivation for this mechanism, Karl Friston proposed the “Free Energy Principle” (FEP) [5]. This principle states that any self-organizing system in equilibrium with its environment must minimize its variational free energy, which is an upper bound on the “surprise” of the system’s state. In other words, organisms resist thermodynamic disorder and maintain the boundaries of their own states by minimizing free energy [5].

Although the Free Energy Principle is mathematically elegant, it gives rise to a fatal philosophical challenge: the well-known “Dark Room Problem.” If an organism’s only goal is to minimize prediction error or “surprise,” why does it not remain in a dark room without any stimuli, thereby ensuring that prediction error never arises? In early responses,

---

* Corresponding author: zhangxs001@zju.edu.cn

Friston and others argued that organisms possess innate "preferred states," such as a human preference for light and food, so staying in a dark room would instead generate enormous prediction error [9]. Yet this response has been criticized as circular: it explains the motive force of evolution, namely the minimization of surprise, by appealing to the result of evolution, namely prior preferences.

In later dialogue with Andy Clark and others, Friston offered a more complex defense. He introduced morphogenesis on an evolutionary timescale and the system's own attractor states, arguing that dark-room agents exist only where they can survive, as in the case of cave-dwelling animals, while the models of most organisms are formed through natural selection in changing environments and therefore require active sampling of the environment to maintain their attractor states [6]. Yet even this strongest defense does not escape the fundamental philosophical difficulty: it conflates logical coherence with ontological necessity. Although Friston explains why we are not in a dark room by appealing to evolution and attractor states, he cannot explain from within the Free Energy Principle why evolution itself must occur, nor why a system must maintain any particular attractor state. This constitutes the "ontological explanatory deficit" of predictive processing theory.

To respond to this deficit, this paper introduces the theory of "cognitive compensation" and the RID model developed in *Knowing and Speaking: An Ontological Investigation of Human Cognition* [21]. Based on the principle of recursive weakening and compensation, this theory argues that cognition is not an objective reflection of a ready-made world, but a higher-order compensatory activity through which finite beings maintain their own continuity under the pressure of decreasing degree of existence. This paper seeks to demonstrate that the computational mechanism of predictive processing can be mapped precisely onto the RID model–Regularization R, Structure Generation I, and Problem Pressure D–thereby raising predictive processing from a purely computational descriptive framework to an explanatory framework with ontological depth.

2 Literature Review: The Internal-External Debate in Predictive Processing and Its Theoretical Gap

2.1 Core Mechanisms and Historical Development of Predictive Processing

Predictive processing theory has undergone a long development from philosophical intuition to computational modeling. As Chinese scholars Wu Rui and Li Hengwei have noted, its intellectual genealogy can be traced back to Kant's synthetic a priori judgment and Helmholtz's "unconscious inference." After entering the natural-scientific stage, predictive processing gradually developed into a rigorous model of computational neuroscience through information transmission theory, machine-learning algorithms, and the Bayesian brain hypothesis [13].

Its core mechanism consists primarily of three components.

First, the hierarchical generative model. The brain constructs probabilistic models of the causal structure of the world at different cortical levels. Higher levels represent abstract and long-term causal states, while lower levels represent concrete and momentary sensory features [9]. This hierarchical structure allows the brain to process highly complex and nonlinear environmental information.

Second, prediction-error minimization. The system eliminates discrepancies between predictions and sensory inputs through two pathways: it updates the internal model to fit sensory input, which is perceptual inference, or it changes sensory input through bodily action so that it matches internal prediction, which is called active inference [7]. Active inference unifies perception and action under a single mathematical objective: the minimization of free energy.

Third, precision weighting. Through attentional mechanisms, the system adjusts the reliability weight of prediction errors at different levels. By estimating the "precision" of prediction error, it determines whether to assign greater weight to prior prediction or to sensory evidence [3]. This mechanism is crucial for handling environmental uncertainty and dynamic change.

2.2 The Integrative Controversy Between Cognitivism and Enactivism

In recent years, cognitive science has witnessed intense debate over how to understand the ontological status of predictive processing theory. Two sharply opposed camps have emerged, reflecting a deep conflict between the "computational representation" paradigm of traditional cognitive science and the "embodied-enactive" paradigm of second-generation cognitive science.

Cognitivist predictive processing, represented by Jakob Hohwy, maintains an intracranial internalist position [10]. It holds that the brain can receive only ambiguous neural signals through the sensory barrier, namely the Markov blanket, and is therefore essentially an isolated inference engine trapped inside the dark skull. From this perspective, body and environment merely provide sensory data and execute motor commands as external variables; they do not constitute the cognitive process itself.

Free-energy enactivism, advanced by scholars such as Andy Clark and Shaun Gallagher, attempts to integrate predictive processing with embodied enactive cognition [8]. It emphasizes active inference, holding that perception and action are two sides of the same coin: perception tests internal predictions, while action actively changes the world to minimize prediction error. Chinese scholars such as Liu Linshu have further explored enactive predictive processing from a dynamical-systems standpoint, arguing that although predictive processing presupposes internal representation, it also has a distinctively enactive character as a dynamical theory [14]. Meng Wei likewise points out that the philosophical foundation of cognitive science is shifting from Cartesian internal representation to Heideggerian being-in-the-world [16].

2.3 Limitations of Existing Research and the Point of Departure of This Paper

Although free-energy enactivism attempts to grant constitutive status to body and environment, its integrative path still faces theoretical difficulties. As Michael Kirchhoff and colleagues show, the Free Energy Principle incorporates all self-organizing systems into a single framework, making it difficult to distinguish simple physical coupling from complex cognitive representation [11]. If water droplets and human beings are both minimizing free energy, what is distinctive about higher-order human cognition?

This paper argues that the crux of the existing debate lies in the absence of a unified ontological framework capable of traversing physical existence, biological coupling, and higher-order cognition. Cognitivism treats the brain as a pure calculator detached from the context of survival, while enactivism emphasizes the participation of the body but does not explain why the body must participate in such costly predictive computation.

The theory of "cognitive compensation" in *Knowing and Speaking* provides precisely this bridge. It treats cognition as a necessary compensatory product of declining degree of existence. From this perspective, both the brain's internal computation, such as generative models, and the body's external action, such as active inference, are homologous compensatory means for responding to the same "existential pressure." This not only provides a solid ontological foundation for predictive processing, but also effectively dissolves the opposition between cognitivism and enactive cognition.

3 Theoretical Framework: The RID Model and the Threefold Ontological Anchoring of Predictive Processing

To resolve the "ontological explanatory deficit" in predictive processing theory, this paper introduces the RID model from *Knowing and Speaking*–Regularization R, Structure Generation I, and Problem Pressure D–as its underlying ontological framework. The RID model holds that cognitive activity is a compensatory mechanism through which finite beings respond to the pressure of decreasing degree of existence. From this perspective, the computational mechanism of predictive processing is no longer an accidental byproduct of evolution, but a necessary compensatory expression. The following sections elaborate the threefold conceptual anchoring between the RID model and the core mechanisms of predictive processing, while introducing such mediating concepts as autopoiesis and homeostasis to complete the philosophical transition from mathematical computation to ontology.

3.1 Anchoring Problem Pressure (D) to Prediction Error: From Information Quantity to Existential Boundary

In standard predictive processing theory, prediction error is defined merely as Shannon surprise in information-theoretic terms, that is, the mathematical discrepancy between top-down prediction signals and bottom-up sensory input signals [5]. Yet this purely computational definition cannot answer a fundamental question: why should this informational discrepancy drive the operation of the entire biological cognitive system? If surprise is only a deviation in mathematical probability, why can it generate such powerful survival-dynamical effects?

To cross the gap from mathematical information to survival motivation, we need to introduce Maturana and Varela's theory of autopoiesis as a mediating concept [15]. An autopoietic system is one that continuously produces and maintains its own physical boundaries through a sustained network of interactions among its components. For an organism, existence is not a static given but a dynamic achievement that must be maintained through metabolism and environmental exchange. The core of this dynamic achievement is homeostasis: the restriction of the organism's own state to a very narrow physical and chemical phase space that is vital for survival.

Anchoring prediction error to "Problem Pressure (D)" in the RID model is based precisely on this autopoietic logic. In cognitive compensation theory, problem pressure (D) represents the impact of the environment on the boundary of the existent, and is the fundamental force that triggers compensatory activity [21]. When prediction error arises, it does not merely mark statistical inaccuracy in the internal model; more deeply, it marks the system's deviation from its autopoietic attractor state. Every uneliminated prediction error is a signal that environmental variables are breaking through the boundaries of the system's homeostasis. Thus, the impulse to minimize prediction error is essentially the compensatory impulse of finite beings to close existential fissures and maintain autopoietic boundaries. Prediction error is the precise neural-computational quantification of problem pressure (D). This transforms predictive processing from the cognitivist metaphor of pursuing informational accuracy into an existential proposition concerning the maintenance of continuity of existence.

To understand this anchoring more deeply, consider an organism facing an extreme environmental change. When an organism is exposed to extreme temperature or toxins, its sensory system generates enormous prediction error. From a purely informational perspective, this is merely a state of high surprise. From an ontological perspective, however, this prediction error directly reflects the deviation of internal physiological indicators, such as body temperature or blood pH, from the safe range of homeostasis, thereby constituting fatal problem pressure. If the system cannot eliminate this error through active inference, such as fleeing a dangerous environment, or by updating its internal model, such as activating stress-response mechanisms, its autopoietic network collapses and death follows. Prediction error is therefore not merely a byproduct of information processing, but an ontological alarm concerning life and death.

3.2 Anchoring Structure Generation (I) to the Hierarchical Generative Model: From Objective Mirror to Survival Projection

The central assumption of predictive processing theory is that the brain possesses a hierarchical generative model of the world, a model capable of simulating how sensory data are generated by external causes [9]. Yet cognitivist predictive processing often treats this model as an internal mirror or statistical map of the external objective world. This representationalist presupposition remains trapped in a Cartesian subject-object dualism.

Anchoring this model to “Structure Generation (I)” in the RID model fundamentally overturns the mirror metaphor. Cognitive compensation theory argues that compensatory structure (I) is not a passive reflection of a ready-made world, but an active construction generated in response to specific survival pressure [21]. To deepen this point, we may draw on Kant’s schematism and Heidegger’s readiness-to-hand.

For Kant, a schema is the mediator through which a priori categories are applied to empirical intuition. Similarly, the brain’s hierarchical generative model is not an objective sketch of the world, but a “compensatory schema” through which the subject projects survival needs (the a priori) onto sensory data (experience). Higher-level abstract causal states and lower-level concrete sensory features together constitute a multidimensional compensatory network. The fundamental purpose of this network is not to reflect the world “truthfully,” but to suppress prediction error, that is, problem pressure (D), effectively.

Therefore, the world revealed by the generative model is not a world composed of objective physical properties as present-at-hand, but a world composed of affordances and action possibilities as ready-to-hand. The brain represents only those environmental features that are crucial to maintaining the homeostasis of the existent, and ignores others. This explains why different species, with different compensatory needs and bodily structures, evolve radically different generative models and perceptual worlds (Umwelten). For example, the bat’s generative model is primarily based on ultrasonic echoes, constructing a three-dimensional auditory world full of spatial obstacles and prey trajectories; the human generative model, by contrast, relies heavily on vision, constructing a visual world full of colors, shapes, and object constancy. Structure generation (I) is essentially a survival-oriented projection rather than a truth-oriented reflection.

Furthermore, the hierarchical character of the hierarchical generative model perfectly accords with the increasing complexity of structure generation (I). Under increasingly complex environmental pressure, a simple single-layer stimulus-response model, such as a reflex arc, is no longer sufficient. The brain must construct multilayered abstract representations to capture environmental regularities across longer temporal and spatial scales. Higher-level prior predictions can “explain away” lower-level prediction errors, thereby maintaining system stability at a more macroscopic level. This hierarchical transition from the concrete to the abstract and from the momentary to the long-term vividly expresses the continual upgrading of cognitive compensatory structures.

3.3 Anchoring Regularization (R) to Precision Weighting: Cognitive Economy and Pattern Stabilization

In the predictive processing framework, precision weighting is the key to attentional mechanisms and flexible inference. The system dynamically adjusts the relative weight of prior prediction and sensory evidence by estimating the reliability, or precision, of prediction error [3]. If sensory input has high precision, as in a bright environment, the system relies more on sensory data to update its model; if sensory input has low precision, as in darkness, the system relies more heavily on prior prediction.

Anchoring precision weighting to “Regularization (R)” in the RID model reveals the economic principle behind this computational mechanism. Regularization (R) refers to the process of stabilizing complex and variable experience into repeatable cognitive patterns [21]. Faced with endless and rapidly changing environmental variables, finite beings cannot undertake comprehensive structural reorganization (I) for every prediction error. That would exhaust all metabolic resources and lead rapidly to systemic collapse.

Precision weighting is precisely the neural-computational means through which this regularization is achieved. By assigning high precision weights to higher-level prior predictions, the brain is effectively performing cognitive compression and pattern stabilization. It allows the system to ignore unimportant or unreliable lower-level fluctuations and to process information efficiently by relying on established rules, or priors. This regularizing mechanism enables the existent to maintain a relatively stable horizon of expectation within a turbulent environment.

Thus, precision weighting is not merely a mechanism for allocating attention. It is the necessary result of cognitive compensation under resource constraints as it seeks survival efficiency and cognitive stability in response to problem pressure. In everyday life, our ability to drive a car skillfully or conduct fluent conversation depends precisely on the high regularization of many sensorimotor patterns. The brain assigns very high prior precision to these internalized rules, so that we do not need to recalculate every detail of each muscular contraction in every action. Only when an unexpected major event occurs, such as a sudden obstacle ahead producing high-precision prediction error, does the system break its ordinary regularized state and urgently mobilize resources for localized structural reorganization (I).

3.4 The Ontological Answer to the “Dark Room Problem”: Why Does Evolution Occur?

On the basis of the threefold anchoring above, we can directly address the weak point in Friston’s defense of the “Dark Room Problem” and provide a genuinely non-circular ontological answer.

Friston defends his position by appealing to attractor states and morphogenesis: organisms are not in dark rooms because they evolved in changing environments and must actively sample the environment to maintain their specific attractor states, such as the human need to seek light and food [6]. Although this defense is logically coherent, it remains a retrospective evolutionary explanation. It cannot answer the following question: if the dark room–a state without stimuli or prediction error–is the absolute ideal of free-energy minimization, why did life, at the beginning of evolution, leave a simple, nearly dark-room-like primitive state, such as simple chemosynthesis near deep-sea vents, and evolve increasingly complex nervous systems and cognitive structures that are ever more prone to producing massive prediction errors? Why must evolution occur?

The principle of recursive weakening and compensation in cognitive compensation theory provides an answer to this ultimate question. It holds that the decrease in degree of existence, or weakening, is an intrinsic and irreversible thermodynamic tendency of finite beings, while the increase in structural complexity, or compensation, is a necessary response to rescue the loss of degree of existence [21].

Therefore, an organism cannot choose the dark room not because it accidentally acquired an attractor state of not being in the dark room through evolution, but because remaining still in a dark room cannot prevent the endogenous loss of degree of existence. The absence of environmental stimuli does not eliminate problem pressure (D), because metabolic consumption and entropy increase continue to occur. The system must acquire negentropy, such as food and information, through continually complex structure generation (I) and active inference in order to resist the intrinsic tendency toward weakening.

In this sense, the evolution of predictive processing and hierarchical generative models is itself a massive upgrade of cognitive compensation. Organisms evolved such costly predictive mechanisms precisely because they faced more severe problem pressure than primitive life forms. The Dark Room Problem is a pseudo-problem because it mistakenly equates the absence of environmental stimulation with the absence of survival pressure. From an ontological perspective, maintaining existence is itself an arduous task requiring continuous investment of compensatory resources. Life has no choice but to proceed through continuous prediction and action along an endless path of compensation.

3.5 Synthetic Comparison and Theoretical Boundaries of the Threefold Anchoring

To present the logical relationship of the threefold anchoring more clearly, the core concepts of the RID model and predictive processing can be compared systematically:

This correspondence is not a simple conceptual analogy, but has strict logical necessity. From an ontological perspective, the three elements form a dynamic compensatory cycle: problem pressure (D) triggers the activation of structure generation (I); structure generation (I) stabilizes compensatory achievements into durable patterns through regularization (R); and the rigidity of regularization (R) may produce new problem pressure (D) under new environmental changes, thereby driving the next round of compensatory upgrading. This cycle is the inner logic of the evolution of life and the ontological ground of the three-in-one unity of perception, action, and learning in predictive processing theory.

Of course, the theoretical integration proposed in this paper also has certain boundary conditions. As a macro-level ontological framework, the primary contribution of the RID model is to provide an ontological motivation–the reason why predictive processing is so–rather than to replace the computational description of how predictive processing operates. The relation between the two is complementary, not competitive: ontological foundation and mechanistic description. In addition, the integrative framework developed here applies primarily to biological cognitive systems. Its applicability to artificial intelligence systems requires separate discussion, which is also the problem domain addressed in the final chapter of *Knowing and Speaking*.

4 Case Examination: “Compensatory Failure” and Phenomenological Reconstruction in Psychopathology

To test the theoretical efficacy of the ontological integrative framework above, this section selects two psychopathological phenomena that have been most extensively studied in predictive-processing research–schizophrenia and autism spectrum disorder (ASD)–as test fields. In standard predictive-processing literature, both are typically explained as abnormalities in precision weighting. Yet this mechanistic explanation often neglects the qualitative transformation of the patient’s subjective experience. By introducing the RID model, this paper reinterprets these pathological phenomena at the depth of phenomenology as “compensatory failure” or “pathological compensation” in the structure generation (I) and regularization (R) mechanisms of an existent facing extreme problem pressure (D).

4.1 Schizophrenia: Aberrant Salience and the Tragic Retreat of Internal Compensation

Within the predictive-processing framework, the core symptoms of schizophrenia, such as hallucinations and delusions, are widely explained as products of aberrant salience. Because of abnormalities in neuromodulatory systems, such as dopamine, the patient’s brain cannot correctly estimate the precision of lower-level sensory prediction errors, causing originally meaningless environmental noise to be assigned extremely high weight, that is, aberrant salience [4]. To explain these sudden, unreasonable, and intensely salient error signals, higher-level generative models are forced into catastrophic reorganization, producing delusions as strong priors that attempt to explain abnormal sensory input and hallucinations as excessive top-down projections of strong priors.

Table 1. Correspondence Between the RID Model and Core Concepts in Predictive Processing

RID concept	Predictive-processing counterpart	Mediating concept	Ontological implication
Problem Pressure (D)	Prediction Error	Deviation from the autopoietic attractor	Signal of rupture in the boundary of existence
Structure Generation (I)	Hierarchical Generative Model	Kantian schematism and Heideggerian readiness-to-hand	Active projection oriented toward survival
Regularization (R)	Precision Weighting	Principle of cognitive economy	Efficient allocation of compensatory resources

However, this computational-mechanistic description does not reach the core of the patient's lived experience. From the ontological perspective of the RID model, schizophrenia is not merely a computational error, but a profound ontological crisis.

When aberrant salience occurs, the patient is in fact being bombarded by infinitely amplified problem pressure (D). Subtle environmental fluctuations that would normally be filtered out by regularization (R) now strike the boundaries of the existent's homeostasis with overwhelming force. Under such extreme problem pressure, normal structure generation (I) can no longer sustain effective coupling between subject and world.

To avoid the complete collapse of the sense of existence, that is, homeostasis, the patient is forced to adopt an extreme and tragic compensatory strategy: cutting off real coupling with the external world and retreating into an internally compensatory structure that is highly regularized (R) but entirely false.

Delusions are the highest expression of this pathological compensation. They are not simply mistaken beliefs, but extremely rigid and irrefutable "hyper-regularizations." By constructing an all-encompassing delusional system, such as the belief that the entire world is monitoring me, the patient regains explanatory power over the world in a distorted way and temporarily suppresses terrifying abnormal prediction errors. Although this internal compensatory structure is false in the physical world, it is, in the ontological sense, the last defensive line built by the patient to maintain a minimal coherence of subjectivity. As phenomenological psychiatry suggests, the world (Umwelt) of the schizophrenic patient does not disappear; rather, it undergoes a transformation of reality detached from common sense. The patient retreats into an enclosed space governed by absolutely certain prior rules in order to resist the unbearable weight of the external world.

Moreover, the negative symptoms of schizophrenia, such as affective flattening and avolition, can also be explained within the compensatory framework. When the system invests most of its compensatory resources, such as computational capacity and metabolic energy, into maintaining that vast and false internal delusional structure, the resources available for coping with external reality and social interaction become severely depleted. This comprehensive depletion of resources manifests as withdrawal and indifference in everyday behavior. Therefore, negative symptoms are not merely functional deficits; they are the necessary consequence of compensatory resources being severely overdrawn by an internal pathological structure.

4.2 Autism Spectrum Disorder: Deficient Regularization and Direct Bombardment by Problem Pressure

In contrast to schizophrenia, autism is usually explained in predictive-processing models as excessively low precision in higher-level priors or excessively high precision in lower-level sensory error [12]. This causes patients to rely excessively on current sensory input and makes it difficult for them to use past experience, or prior models, to smooth and predict future changes. This mechanism explains autistic individuals' extreme sensitivity to environmental change, sensory hypersensitivity, strong reliance on repetition and routine, and difficulties in social interaction, which depends heavily on abstract inference.

Placed within the RID model, this mechanism reveals the distinctive survival predicament faced by autistic individuals: a deficiency in higher-order regularization (R), which leaves the existent directly bombarded by bare problem pressure (D).

Normally developing cognitive systems establish higher-level and abstract regularization models (R), compressing complex and variable physical and social stimuli into predictable patterns and thereby greatly reducing survival pressure. In social situations, for instance, we can rapidly call upon higher-order theory-of-mind priors through others' micro-expressions and context, and accurately predict their intentions. Because autistic individuals have difficulty establishing or maintaining such higher-order regularization mechanisms, however, their brains remain in a depleted state of high-frequency reorganization (I). Every small environmental change, such as a slight movement of furniture or a faint flicker of light, appears to them as a wholly new and unpredictable surprise, that is, as problem pressure (D) directly threatening their existential boundary.

In the absence of higher-order compensatory means, autistic individuals have to fall back on the most primitive and lowest-level physical compensatory strategies to build defenses. This explains why they show strong stereotyped behaviors and an intense insistence on sameness. By repeatedly performing specific actions, such as rocking the body or arranging objects, or by strictly following fixed daily routines, they attempt to artificially create a locally and absolutely predictable “pseudo-dark-room” state.

This reliance on physical repetition is essentially a lower-level compensation for the absence of higher-order regularization (R). They cannot predict the world at an abstract cognitive level, so they can only force the world to remain unchanged at a concrete physical level. Therefore, the world of autistic individuals is not, as cognitivism might suggest, a world of misallocated computational resources, but a bare world filled with fear and lacking a buffering zone. Their insistence on routine is the instinctive effort of finite beings to grasp a point of stability amid a torrent of problem pressure.

Through this phenomenological reconstruction of the two pathological phenomena, we can clearly see that computational abnormalities in predictive processing are always grounded in the difficult struggle and eventual compromise of compensatory structures (I and R) as an existent confronts problem pressure (D) and seeks to maintain its survival boundary. This further confirms the explanatory power of the RID model as the underlying ontological framework of predictive processing. Whether in the internal over-compensation of schizophrenia or the low-level physical compensation of autism, life is attempting to maintain its own existence under extreme conditions.

5 Conclusion: Toward an Ontologically Integrated Philosophy of Cognitive Science

By introducing the theory of cognitive compensation and the RID model from *Knowing and Speaking*, this paper reconstructs the core mechanisms of predictive processing theory from an ontological perspective. By anchoring prediction error to problem pressure (D), hierarchical generative models to structure generation (I), and precision weighting to regularization (R), the paper not only provides predictive processing with an explanatory framework that crosses the gap between mathematical computation and survival motivation, but also offers a new theoretical path for resolving the long-standing “Dark Room Problem” and the internal-external debate in this field.

5.1 Responding to and Deepening Chinese Debates in the Philosophy of Cognitive Science

In recent years, Chinese philosophy of science and technology has engaged in lively discussion of predictive processing theory and its relation to embodied cognition. For example, Li Hengwei and colleagues have explored the place of predictive processing in the embodied-cognition tradition and emphasized the dependence of cognition on the body [13]. Meng Wei, drawing on the phenomenological tradition, argues that the philosophical foundation of cognitive science should complete a shift from Descartes to Heidegger and emphasize the primordiality of being-in-the-world [16]. Liu Linshu and colleagues further outline the dynamical-systems stance of enactive predictive processing, seeking a balance between internal representation and external enactment [14]. In addition, Zhang Liping and Liu Xiaoqing analyze the epistemological foundation of the Free Energy Principle and point out its possible reductionist risks when explaining higher cognition [20].

This paper responds to and deepens these local academic discussions. It must be emphasized that enactivism has already taken an important step toward ontology by stressing bodily coupling and active inference. The “cognitive compensation” framework proposed here does not attempt to replace enactivism, but to provide it with a deeper ontological grounding–an ultimate source of motivation that answers why bodily coupling and active inference are necessary. In this view, neither the brain’s neural computation, such as Bayesian inference, nor the body’s dynamical coupling, such as sensorimotor loops, nor the physical constraints of the environment should be treated as independent or opposing elements of cognition. Rather, all are compensatory structures (I and R) necessarily evolved by finite beings in response to decreasing degree of existence (problem pressure D) in order to maintain their autopoietic boundaries (homeostasis).

Within this integrated perspective, the “internal representation” emphasized by cognitivism is not a mirror of the objective world, but a compensatory schema generated by the existent to suppress problem pressure. The active inference and bodily coupling emphasized by enactivism are the necessary extension of this compensatory schema into physical space. Therefore, the combination of predictive processing theory and embodied cognition should not be merely a mechanistic assembly, such as adding motor control to a prediction-error-minimization framework, but an ontological recasting: the recognition that body and brain are manifestations of the same compensatory process at different physical levels.

5.2 Theoretical Contributions and Future Prospects

In sum, the theoretical contribution of this paper can be summarized in four respects.

First, it proposes the original concept of the “Compensatory Schema,” thereby fundamentally transcending the traditional cognitivist view of representation. The paper argues that internal representations in the brain are not statistical maps of the external objective world, but compensatory schemas actively constructed by finite beings in response to specific problem pressures. This concept not only breaks decisively with the cognitivist metaphor of the internal mirror, but also offers a new explanatory dimension for understanding the roots of cognitive differences among different species and individuals from different cultures.

Second, it dissolves the ontological deficit of the Dark Room Problem. The paper argues that the Free Energy Principle falls into circular reasoning because it fails to explain the necessity of evolutionary motivation itself. By introducing the principle of recursive weakening and compensation, the paper demonstrates that the costly mechanism of predictive processing is not an accidental evolutionary choice, but a necessary compensation by finite beings against endogenous loss of degree of existence. The Dark Room Problem is a pseudo-problem because the absence of stimulation is by no means equivalent to the absence of survival pressure.

Third, it completes the transition from computational mechanism to phenomenological experience. Through the phenomenological reconstruction of schizophrenia and autism, the paper shows how the RID model transforms cold neural-computational abnormalities, such as dysregulated precision weighting, into vivid and painful lived experiences, such as the collapse of the world brought about by aberrant salience or the bare world caused by deficient regularization. This provides an interpretive framework with both philosophical depth and human concern for interdisciplinary research in psychopathology.

Fourth, it offers an example of international dialogue for an original Chinese theoretical framework. This paper attempts to connect indigenous Chinese philosophical concepts in Dongyue Wang's *A General Theory of Evolution* and *Knowing and Speaking*–including cognitive compensation and the RID model–with cutting-edge international paradigms in cognitive science, including predictive processing and the Free Energy Principle. This not only tests the explanatory power of the local theory, but also offers international philosophy of cognitive science a distinctive perspective from an Eastern ontological tradition.

Future research can continue to develop this integrative framework in several directions. On the one hand, at a more microscopic neurobiological level, it can explore the precise correspondence between "structure generation (I)" in the RID model and cortical microcircuits. On the other hand, it can extend toward higher-order social cognition and cultural evolution, examining how language, as the highest-order form of regularization R, functions as a collective compensatory mechanism for distributing and alleviating problem pressure (D) within human communities. In addition, the question of how to introduce the compensatory mechanism of this ontological perspective into the architecture of next-generation artificial intelligence, especially artificial general intelligence, so that it possesses not only powerful computational capacity but also a genuine drive of "existential pressure," is a highly challenging frontier issue. In this sense, the integration of cognitive compensation and predictive processing is not only a theoretical task for the philosophy of cognitive science, but also a necessary path for understanding the situation of human existence in the universe.

References

[1] M. Allen and K. J. Friston, "From cognitivism to autopoiesis: Towards a computational framework for the embodied mind," *Synthese*, vol. 195, no. 6, pp. 2459–2482, 2018.

[2] A. Clark, "Whatever next? predictive brains, situated agents, and the future of cognitive science," *Behavioral and Brain Sciences*, vol. 36, no. 3, pp. 181–204, 2013.

[3] H. Feldman and K. J. Friston, "Attention, uncertainty, and free-energy," *Frontiers in Human Neuroscience*, vol. 4, p. 215, 2010.

[4] P. C. Fletcher and C. D. Frith, "Perceiving is believing: A bayesian approach to explaining the positive symptoms of schizophrenia," *Nature Reviews Neuroscience*, vol. 10, no. 1, pp. 48–58, 2009.

[5] K. Friston, "The free-energy principle: A unified brain theory?" *Nature Reviews Neuroscience*, vol. 11, no. 2, pp. 127–138, 2010.

[6] K. Friston, C. Thornton, and A. Clark, "Free-energy minimization and the dark-room problem," *Frontiers in Psychology*, vol. 3, p. 130, 2012.

[7] K. J. Friston, J. Daunizeau, J. Kilner, and S. J. Kiebel, "Action and behavior: A free-energy formulation," *Biological Cybernetics*, vol. 102, no. 3, pp. 227–260, 2010.

[8] S. Gallagher and M. Allen, "Active inference, enactivism and the hermeneutics of social cognition," *Synthese*, vol. 195, no. 6, pp. 2627–2648, 2018.

[9] J. Hohwy, *The Predictive Mind*. Oxford: Oxford University Press, 2013.

[10] J. Hohwy, "The self-evidencing brain," *Noûs*, vol. 50, no. 2, pp. 259–285, 2016.

[11] M. D. Kirchhoff, T. Parr, E. Palacios, K. Friston, and J. Kiverstein, "The markov blankets of life: Autonomy, active inference and the free energy principle," *Journal of the Royal Society Interface*, vol. 15, no. 138, p. 20170792, 2018.

[12] R. P. Lawson, G. Rees, and K. J. Friston, "An aberrant precision account of autism," *Frontiers in Human Neuroscience*, vol. 8, p. 302, 2014.

[13] H. Li and Q. Chen, "Predictive processing theory and the future of cognitive science," *Studies in Dialectics of Nature*, vol. 38, no. 5, pp. 1–8, 2022, english title and journal title translated from the source manuscript.

[14] L. Liu and H. Ye, "Enactive predictive processing: A dynamical systems standpoint," *Journal of Psychological Science*, vol. 47, no. 5, pp. 1254–1261, 2024, english title and journal title translated from the source manuscript.

[15] H. R. Maturana and F. J. Varela, *Autopoiesis and Cognition: The Realization of the Living*. Dordrecht: D. Reidel Publishing Company, 1980.

[16] W. Meng, "The transformation of the philosophical foundation of cognitive science: From descartes to heidegger," *Mind and Computation*, vol. 2, no. 3, pp. 45–52, 2008, english title and journal title translated from the source manuscript.

[17] G. Pezzulo, F. Rigoli, and K. Friston, "Active inference, homeostatic regulation and adaptive behavioural control," *Progress in Neurobiology*, vol. 134, pp. 17–35, 2015.

[18] A. K. Seth, "The cybernetic bayesian brain: From interoceptive inference to sensorimotor control," *Biological Cybernetics*, vol. 108, no. 3, pp. 267–288, 2014.

[19] F. Xue, "Embodied cognition from the perspective of predictive processing," *Philosophical Trends*, no. 11, pp. 116–124, 2020, english title and journal title translated from the source manuscript.

[20] L. Zhang and X. Liu, "The epistemological foundation and limits of the free energy principle," *Studies in Philosophy of Science and Technology*, vol. 38, no. 2, pp. 15–21, 2021, english title and journal title translated from the source manuscript.

[21] X. Zhang, "Knowing and saying: An ontological investigation of human cognition," PSSXiv (Philosophy and Social Sciences Preprint Server), may 2026, pSSXiv:202605.04152V1; CSTR:32012.36.PSSXiv.202605.04152. [Online]. Available: https://zsyyb.cn/abs/202605.04152