Value as Sensation: A Selectionist Formalization of the Sentient Principle and Emergent Behavior

Wu Jingyu

Email: jyw2004@126.com

Abstract

The study of "value" in the sciences of mind has long been trapped in the "evaluation" paradigm: value is regarded as an object to be computed, a function to be optimized. This paradigm cannot naturally explain the everyday fact that the same stimulus possesses opposite value polarities under different homeostatic states, nor can it explain why an organism "turns a blind eye" to food after satiation—a behavior that appears "inefficient" from an external perspective. The Sentient Principle proposes a fundamental alternative: value is not a tool for evaluation but sensation itself. Life is a ceaseless "Sensation-Behavior Loop" system, whose fundamental driving force is "Sensation Discrepancy"—the identifiable imbalance between immediate sensation and Anticipatory Sensation. This paper advances three core claims: value is sensation; value flows in real time; value as sensation can directly drive behavior. These claims are formalized for the first time into an executable selectionist simulation system, SBLB-2.5. The system strictly adheres to selectionist principles: no preset reward function, no channel locking; all learned behaviors arise from the selective stabilization of eligibility trace spatiotemporal co-occurrence patterns by endogenous value signals. Over more than 550,000 steps of continuous operation, starting from a state where all weights are zero, the system exhibits the gradual differentiation of value sensation, satiation-induced anorexia, dynamic regulation of risk preference, and autonomous stabilization of the sleep-wake cycle. Crucially, the behavioral trajectories in the mature stage appear "unintelligible" from an external perspective—the organism sometimes approaches food, sometimes ignores it, and sometimes actively avoids it. Only by simultaneously tracking its internal sensation states (energy, hunger, overfullness, anticipated value, LHA pulses) can these behaviors be given a unified explanation: they are all outward manifestations of the real-time flow of the organism's value sensation. As the first selectionist formalization of the Sentient Principle, SBLB-2.5 demonstrates a new path for translating first-person sensory claims into an executable third-person simulation system, providing a concrete working model for the sciences of mind to move beyond the computationalist paradigm.

Keywords: Sentient Principle; Selectionism; Value System Sensation; Sensation-Behavior Loop; Sensation Discrepancy; Value Flow; Minimal Sentient Model

The theoretical framework, philosophical foundations, and complete genealogy of terms employed in this paper—including the archaic sense of "Sentient" as derived from Whytt (1751), the tripartite tension structure originating from Vyazemsky's симпатия, and the full intellectual lineage from Whytt through Sechenov, Edelman, and Simonov—are systematically developed in the companion work The Sentient Principle (Wu, 2026), without which the present paper cannot be fully understood.

1. Introduction

1.1 The Value Predicament in the Sciences of Mind

Imagine an extremely hungry animal. When it sees familiar food, it approaches without hesitation. When it makes contact with the food, it eats continuously. In this moment, the food is undoubtedly "good" for it. Now imagine the same animal has just eaten its fill. When it encounters the same food again, its behavior shifts fundamentally: it no longer approaches, and may even actively avoid it; if forcibly fed, it shows rejection and discomfort. In this moment, the same food has become "not good."

The food has not changed. The environment has not changed. What has changed is merely the animal's internal physiological state—from hunger to satiation. This everyday fact points to a fundamental theoretical question: what exactly is the relationship between value and internal state? The behaviorist tradition has long recognized the importance of internal states—Pavlov's dogs had to be in a state of hunger, and Skinner's rats would not press a lever when sated. However, within the behaviorist framework, internal states are understood as "drives" or "operational preconditions" for behavior, while value is attributed to the properties of reinforcers. The behavioral shift in a sated animal is explained as "drive reduction" or "motivational decline," rather than a change in the nature of value itself. The Sentient Principle proposes a fundamentally different understanding: hunger and satiation are not merely conditions for behavior to occur—they are themselves innate Value System Sensations—hunger is the direct sensation of "not good," and satiation is another kind of direct sensation of "not good." The "goodness" of food is not an intrinsic property; rather, it is because food can eliminate the Sensation Discrepancy of hunger. Therefore, the behavioral shift from hunger to satiation is not a different expression of the same value at different drive levels, but a polarity reversal of value sensation itself. The same physical stimulus can possess diametrically opposite value polarities under different homeostatic states. Any theoretical framework that attempts to fix value in the stimulus itself, or in some independent "value function," cannot naturally explain this phenomenon.

Contemporary sciences of mind pursue the study of "value" primarily along three paths. Reinforcement learning and computational neuroscience formalize value as the cumulative expectation of reward (Sutton & Barto, 2018); an externally defined reward function stipulates which states are "good," and dopamine signals are interpreted as reward prediction errors. The free-energy principle and predictive coding understand the organism as a Bayesian inference system that minimizes prediction error (Friston, 2010); value is implicit in the prior distribution—what conforms to expectation is good. The somatic marker hypothesis points out that bodily sensations play an auxiliary role in rational decision-making (Damasio, 1994); sensations mark the affective valence of options, helping the decision center quickly rule out unfavorable choices. Although these three paths differ in their specific mechanisms, they share a deep presupposition: **value is something that needs to be computed, evaluated, and optimized**. They all presuppose that value is "computed," that the standard of value is externally given, that value is relatively fixed, and that an implicit evaluator or decision-maker exists.

The fundamental predicament of this "evaluation" paradigm lies in its **inability to intrinsically explain the real-time reversal of value polarity**. When a sated animal avoids food, the evaluation paradigm must resort to some kind of "strategy switching"—for example, a phased change in the reward function, or a separate independent system inhibiting feeding behavior. But this essentially uses one new presupposition to cover up an old one; it still does not answer the more fundamental question of "why the state of satiation can alter the value of food." Within these frameworks, the relationship between satiation and value is extrinsic, connected through additional mechanisms, rather than being two sides of the same intrinsic process.

The evaluation paradigm also contains an implicit presupposition: the purpose of behavior is to maximize some kind of "gain." In reinforcement learning, the success criterion for an agent is the maximization of cumulative reward; in the free-energy principle, it is the minimization of prediction error. Regardless of the specific form, "efficiency" is treated as the core measure of system performance. Yet a sated animal "turning a blind eye" to food is precisely "inefficient" by this measure—it gives up easily accessible nutrition. This phenomenon poses a second challenge to the evaluation paradigm: not only can value polarity reverse in real time, but the goal of behavior is also not always "maximizing gain." The behavioral goal of an organism seems to be the maintenance of a certain inner balance—eliminating Sensation Discrepancy, restoring the coordination of the sensory field—rather than greedily seizing external resources. This insight demands that we re-understand the basic dynamics of behavior: from "maximizing reward" to "eliminating inner tension," from "external efficiency" to "inner coordination."

1.2 The Alternative Proposed by the Sentient Principle

The Sentient Principle (Whytt, 1751; Wu, J., 2026) proposes a fundamental alternative, whose core is a shift at the ontological level: **value is not an object to be evaluated, but a content to be sensed**. The starting point of this shift is a re-understanding of the basic structure of life. Life is not a "central processor" that receives stimuli, computes value, and outputs behavior, but a ceaseless "Sensation-Behavior Loop" system. Every behavior originates from the drive of some sensation; the result of every behavior generates new sensation; new sensation in turn drives the next behavior—cycling without end until life concludes. In this perspective, sensation is no longer the passive reception of information, but the direct awareness of the organism's own state and its relation to the world. Value is no longer an "evaluation label" attached to sensation, but an intrinsic constitutive dimension of the sensory field itself. When we say "hunger," we are not describing a neutral physiological quantity plus a negative evaluation; we are describing a direct, irreducible negative sensation itself. The "not-goodness" of hunger is not judged but sensed.

The Sentient Principle condenses the above insights into the core claim of "Value System Sensation," comprising three interrelated parts. **Claim One: Value is Sensation.** Value System Sensation is divided into innate and acquired levels. Innate Value System Sensation is the diffuse value baseline preset by evolution, embodied as the direct sensation of the Three Basic Tensions—Defensive (integrity of boundaries), Aggressive (energy balance), and Empathy (rhythmic continuity). Sensations of hunger, pain, and sleepiness are themselves innate negative value sensations; they do not need to be learned, nor can they be reduced to purely physiological signals. Acquired Value System Sensation, guided by the innate baseline, is gradually embodied through accumulated experience into concrete Anticipatory Sensations. The "sense of anticipation" that arises upon seeing food is a typical example of Acquired Value System Sensation. **Claim Two: Value Flows in Real Time.** There is no objective, fixed truth of value. Value is a function of the organism's current interoceptive state. The same stimulus (such as contact with food) can evoke diametrically opposite value sensations under different interoceptive states. This “flow” is not strategy switching but a polarity reversal of value sensation itself. **Claim Three: Value as “Sensation” Directly Drives Behavior.** Value sensation can directly modulate behavioral output at the corresponding level, without going through an intermediate decision module. Hunger not only lets you “know” you should eat—it itself directly drives the initiation and persistence of foraging behavior. Acquired Anticipatory Value Sensations can also directly modulate the direction and intensity of movement. From sensation to behavior is continuous and seamlessly connected, without even requiring an independent “decision” step.

If value is sensation itself, then what is the basic power that drives behavior and learning? The answer of the Sentient Principle is “Sensation Discrepancy”—the identifiable imbalance between immediate sensation and Anticipatory Sensation. When the organism’s immediate sensation conflicts with its Anticipatory Sensation, Sensation Discrepancy arises. This Sensation Discrepancy is itself an endogenous, undeniable tension, directly driving the system to initiate behavior to eliminate this imbalance. The process of learning is the process by which Anticipatory Sensation moves from vague to clear, from generalized to precise, through repeated “Sensation Discrepancy–Behavior–Re-sensing” loops. Crucially, Sensation Discrepancy is completely endogenous. It does not need an external reward signal to define “good” and “bad”—“good” is the experience of Sensation Discrepancy being eliminated, and “bad” is the experience of Sensation Discrepancy persisting or intensifying. This fundamentally distinguishes the Sentient Principle from any framework that relies on external reward functions or preset value labels.

1.3 From Theory to Formalization

The Sentient Principle puts forward a complete theory of life, value, and learning. But for a theory to transcend the level of a philosophical manifesto, it must prove two things: internal self-consistency and generative capacity. Formalization is the key step to achieving these two proofs. By converting the core mechanisms of the theory into a set of precise, executable rules, we can test: under conditions that strictly adhere to the theoretical constraints, will the predicted phenomena spontaneously emerge?

Before entering the work of formalization, a fundamental ontological question must first be clarified: what exactly is the object we are formalizing?

Re-defining the Object of Study: The Sentient System in Its Sensation-Behavior Loop within Its World. What does SBLB-2.5 simulate? The answer is not an “agent” separated from its environment, but an indivisible whole—"the sentient system in its Sensation-Behavior Loop within its world." The organism and environment are not two independent entities in "interaction" but the inner and outer aspects of a single dynamic process. Sensation is not the organism's internal representation of the environment, but the mode of the organism's very existence within the environment. Hunger is not a "detection" of blood glucose levels, but the direct manifestation of energy imbalance—a body-world relation—at the sensory level. Food is not an objective object possessing "positive value" but an emergent entity that arises as a positive value expectation within the sensory field because it can eliminate the Sensation Discrepancy of hunger.

This ontological position fundamentally distinguishes the Sentient Principle from the three mainstream paradigms. The "reward" of reinforcement learning presupposes an external score sent from the environment to the system—value exists in the environment, and the system "acquires" it. The "prediction error" of the free-energy principle presupposes the precision of matching between an internal model and the external world—value is implicit in the prior distribution, and the system "approximates" it. The "somatic marker" of the somatic marker hypothesis presupposes that bodily signals assist an independently existing "decision center"—value is the input to decisions, and the system "consults" it. The common presupposition of all three is: value exists somewhere outside the system—needing to be "detected," "computed," or "marked," and only then converted into behavior. The starting point of the Sentient Principle is fundamentally different from these three: value is not detected, computed, or stored, but sensed—emerging in real time in the direct encounter between life and the world. What SBLB-2.5 aims to formalize is precisely the dynamics of this emergence.

However, formalization itself harbors a profound risk. When we "put" a theory "onto a computer" to run, it is extremely easy to unconsciously introduce computationalist presuppositions—because the computer itself is designed according to the logic of "input-algorithm-output." Implementing Sensation Discrepancy as an error signal in supervised learning, implementing trace consolidation as gradient descent, implementing LHA pulses as a preset reward lookup table—any such design would fundamentally dissolve the unique contribution of the Sentient Principle, causing it to degenerate into a variant of existing frameworks. Therefore, the formalization work of this paper is not merely a technical task but a theoretical defense: **the computer is our vehicle, not our metaphor.** SBLB-2.5 runs on a computer, but the organizing principle it follows must be strictly selectionist, not computationalist. The systematic demarcation of selectionism and computationalism, along with the positioning statement and terminological system of SBLB-2.5 derived from it, will be formally unfolded in Section 2.1.

1.4 The Minimal Sufficient Model: A Research Method for Tracing the Emergence Conditions of "Sentience"

The design of SBLB-2.5 follows the principle of minimal sufficiency. This principle itself constitutes the core methodological stance of this study and requires explicit elaboration here. "Minimal" means that the system contains no mechanisms that are unnecessary for the emergence of the target phenomena. "Sufficient" means that, under these minimal constraints, the target phenomena can spontaneously emerge. Taken together: if a phenomenon can emerge in a minimal sufficient system, then the few core mechanisms that constitute that phenomenon are the sufficient conditions for its emergence. The power of this methodology lies precisely in its "minimality." In a complex system, the causes of emergent phenomena are difficult to attribute—complexity itself may be the true source of emergence, rather than the core mechanisms we claim. By systematically excluding all unnecessary complexity, the minimal sufficient model makes it so that the emergent phenomena can only be attributed to those few core mechanisms that are retained. This is a method of verification by exclusion in the strict sense. SBLB-2.5, therefore, is neither a simulation tool for neuroscience (it does not attempt to reproduce the detailed physiology of specific brain regions) nor a performance system for AI engineering (it does not take task completion efficiency as its optimization target). It is a research method specifically designed for tracing the minimal conditions for the emergence of "sentience."

Specifically, what this study inquires into is: under the premise of highly simplifying the environmental, the physiological aspects of the organism's body, and the neural aspects, and studying them as an indivisible whole, do the following three conditions constitute the sufficient conditions for the budding emergence of "sentience": the innate sentience of the Three Basic Tensions (Defensive/Aggressive/Empathy); Sensation Discrepancy as the sole intrinsic driving force; and direct contact with the world as the sole source of experience? The results of over 550,000 steps of operation answer: yes—these three conditions alone are already sufficient for rich "sentient" phenomena such as value differentiation, satiation-induced anorexia, sleep-wake rhythm, and Conditioned Reflexes to spontaneously emerge.

This "minimality" is itself a profound theoretical insight: the immensely rich and manifest "sentience" that life exhibits at the Vital, Social, and Noetic Strata—its earliest seeds do not require complex preset design; it grows naturally from an extremely simple Sensation-Behavior structure. The task of SBLB-2.5 is precisely to make this “growth” visible under strictly controlled conditions.

2. The Selectionist Formalization of the Theoretical Framework

As stated in Section 1.3, the organizing principle of SBLB-2.5 strictly follows selectionism. This section systematically demarcates it from computationalism.

2.1 Terminology: Defining the Essence of Selectionism and Computationalism

Computationalism regards the mind as some form of computer program and cognition as the algorithmic manipulation of symbols or information. Although the specific forms of computationalism are diverse—from classical symbolism to connectionism to predictive coding—they share the following deep presuppositions: the algorithm presupposition (the system executes preset algorithms or instruction sequences); the optimization presupposition (learning is understood as the optimization of some objective function); the supervision presupposition (there exists a standard of value external to the learning process); the representation presupposition (the system’s internal states are understood as “representations” of the external world). These four presuppositions together form a complete logical closed loop: the system receives input, performs computations based on representations, outputs behavior to optimize the objective function, while the objective function itself is externally given. Within this closed loop, sensation is either reduced to some kind of computational process or suspended as an epiphenomenon.

Selectionism originates from the basic logic of evolutionary biology and was introduced into neuroscience by Edelman, forming the theoretical framework of Neural Darwinism. Its core claim is: the development and learning of the nervous system follow not “instruction-execution” but “variation-selection” logic. The core principles of selectionism can be summarized as: the principle of generating diversity (the system innately or postnatally generates a large number of variant connection patterns, constituting the Repertoire that makes selection possible); the principle of value modulation (there exists an endogenous, diffuse value modulation signal that selectively strengthens or weakens the connection patterns currently active in the system—this signal is not an externally given reward function but a direct reflection of the organism’s own homeostatic state, and is itself a kind of sensation); the principle of selective consolidation (connection patterns positively modulated by the value signal are consolidated, while unmodulated patterns naturally decay—the selected patterns are not “more correct” but better at eliminating the system’s inner tension); the principle of Reentry (there exist parallel, bidirectional, continuous signal loops among various regions within the system, enabling distributed neural activity to be integrated into a unified “sentient” field).

Based on the above distinctions, the positioning of SBLB-2.5 can be precisely stated as: **a selectionist simulation system running on a digital computer**. The computer is its vehicle; selectionism is its organizing principle. The choice of the term “simulation” is critical: it indicates the contingency of the vehicle (the system’s running on a computer is a contingent fact at the level of implementation, not a necessary requirement at the level of theory), the rejection of metaphor (we do not say “the organism processes information like a computer”; rather, “we built a life simulation system on a computer that follows selectionist principles”), and the independence of the theoretical core (the theoretical contribution of SBLB-2.5 lies in demonstrating that, under the premise of strictly adhering to selectionist constraints, the claim of Value System Sensation can generate the predicted emergent behaviors).

To maintain strict terminological consistency, this paper adopts the following terminological correspondence: what is called “weight” in traditional frameworks is here called “Sensation-Behavior Trace,” emphasizing experience-left, selectively consolidated patterns; “weight update” is called “trace consolidation/decay”; “learning” is called “experiential selection”; “reward function” is called “value sensation,” because it is endogenous and ever-flowing; “loss function/error” is called “Sensation Discrepancy,” because it is a first-person sensory imbalance; “network training” is called “iteration of the Sensation-Behavior Loop”; “computation” is called “simulation.”

A Note on Linguistic Levels

This paper simultaneously employs two languages in its exposition: the first-person sensory language (e.g., “the organism feels hunger”) and the formalized simulation language (e.g., “the numerical change in delta_ace”). The reader may thus form an impression that the former describes sensory facts, while the latter describes technical computations, and that the two belong to different levels. This impression needs to be clarified here. delta_ace is not a “simulation” or “substitute” for Sensation Discrepancy, but the sole form of existence of Sensation Discrepancy in this simulation system. The “Systemic feelings” described by Sechenov—that diffuse, global sense of bodily ease or unease—is embodied in SBLB-2.5 precisely through the magnitude of the deviation between interoceptive states and Anticipatory Sensation. The choice of the symbol “delta” (discrepancy) is itself a direct expression of the core concept of the Sentient Principle, rather than an adoption of technical convention.

Therefore, the two languages describe one and the same fact at one and the same level: Sensation Discrepancy. The difference lies solely in the perspective of description—the former proceeding from a first-person sensory perspective, the latter from a third-person formalized perspective. The joint use of both precisely embodies the core methodological intent of this study: to build a bridge between first-person theoretical claims and third-person simulation verification. It is not a confusion of linguistic levels.

The parallel use of these two languages thus enacts the central methodological commitment of this paper: to build a bridge between first-person theoretical claims and third-person simulation verification, without reducing either to the other.

2.2 Formalization Constraints of the Core Claims

Based on the three core claims proposed in Section 1.2, this section translates them into testable formalization requirements. For each claim, we clarify: what constraint does this claim impose on the model architecture? What kind of observational result constitutes support for this claim? What counterfactual condition can test its necessity?

Claim One (Value is Sensation) asserts that value is not an evaluation label for external objects, but an intrinsic constitutive dimension of the sensory field. Innate Value System Sensation is the diffuse value baseline preset by evolution; Acquired Value System Sensation is the Anticipatory Sensation gradually embodied through experience under the guidance of the innate baseline. This claim imposes four formalization requirements on the model. First, the system must possess a set of interoceptive state variables that exist independently of experience and persist continuously, whose computational rules are fixed and unlearnable, and whose very existence constitutes “Innate Value System Sensation.” Second, all plastic connections related to acquired value must be initialized from zero; the system cannot contain any built-in a priori knowledge about “what visual features correspond to what value.” Third, the formation of acquired value must rely entirely on the selective consolidation of Sensation-Behavior Traces by endogenous value signals; no form of “direct assignment” or “supervised learning” can exist. Fourth, after learning is complete, visual input alone (without actual contact) should be sufficient to evoke value sensation. The corresponding test predictions are: plastic traces should start from zero and gradually differentiate with experience into positive (for food features) and negative (for danger/obstacle features); Anticipatory Sensation should produce a positive or negative shift before contact occurs. The counterfactual condition is: if the endogenous value signal is fixed at zero, plastic traces should always remain at zero.

Claim Two (Value Flows in Real Time) asserts that there is no objective, fixed truth of value; value is a function of the organism's current interoceptive state; the same stimulus can evoke diametrically opposite value sensations under different interoceptive states. The formalization requirements of this claim include: the value polarity of contact events cannot be a preset constant but must be a function of interoceptive states; the system must allow the value polarity of the same type of contact event to reverse under different interoceptive states; modulation must be real-time and continuous. In SBLB-2.5, this requirement is implemented through the dynamic computation of LHA pulses: the LHA pulse for food contact is composed of a base positive value plus hunger enhancement, minus overfullness penalty and pain suppression; when overfullness is sufficiently high, it can cause the LHA pulse to change from positive to negative. The test predictions are: positive LHA pulses when hunger is high, and negative LHA pulses when overfullness is high; after satiation, the organism should actively avoid food. The counterfactual condition is: if the LHA pulse were fixed as a constant, satiation-induced anorexia should disappear.

Claim Three (Sensation Directly Drives) asserts that value as "sensation" can even directly modulate behavioral output at the corresponding level, without going through an intermediate decision module. This claim requires: value sensation should be able to directly modulate motor output, without passing through an independent "decision" or "planning" module; interoceptive states should be able to alter in real time the conversion efficiency from value sensation to motor output; the architecture should not contain an independent module that receives value information, conducts "trade-off calculations," and then outputs behavioral commands. In SBLB-2.5, this requirement is implemented through the sensory value pathway: visual Anticipated Value is separated into food and threat components; interoception modulates conversion efficiency in real time through drive gains; the final motor command is directly generated from visual-motor traces and multiplied by the total modulation amount. The test predictions are: the column sums of the three-channel weights should exhibit structural directional separation—the food channel manifesting as approach turning, the danger/obstacle channels manifesting as avoidance turning; risk preference should dynamically adjust with interoception. The counterfactual condition is: if the modulation by interoception on the value-to-motor conversion efficiency is removed, the dynamic adjustment of risk preference should disappear.

2.3 Architectural Constraint of Innate/Acquired Separation

A core insight of the Sentient Principle is: the growth of Value System Sensation must take the innate value baseline as its soil. Innate provides direction (the basic tendency to approach benefit and avoid harm); acquired provides precision (what, specifically, is beneficial and what is harmful). This insight imposes strict constraints on the model architecture.

The innate stratum consists of capabilities the organism possesses at "birth," independent of any experience. In SBLB-2.5, the innate stratum includes: interoceptive computation rules (the calculation formulas for hunger, overfullness, pain, and sleepiness are fixed and not modified by experience); contact Reflexes (the bounce-escape sequence upon contacting danger, the bounce-detour sequence upon contacting an obstacle, the bounce-turn sequence upon contacting a boundary—all with fixed trigger conditions and action patterns); and LHA baseline polarity (the base LHA for food contact is positive, for danger contact negative—this positive/negative direction is preset). Crucially, what the innate stratum provides is the **polarity of evaluation** (food contact is good, danger contact is bad), not the **object of evaluation** ("green objects are food"). The latter is acquired entirely through the acquired stratum.

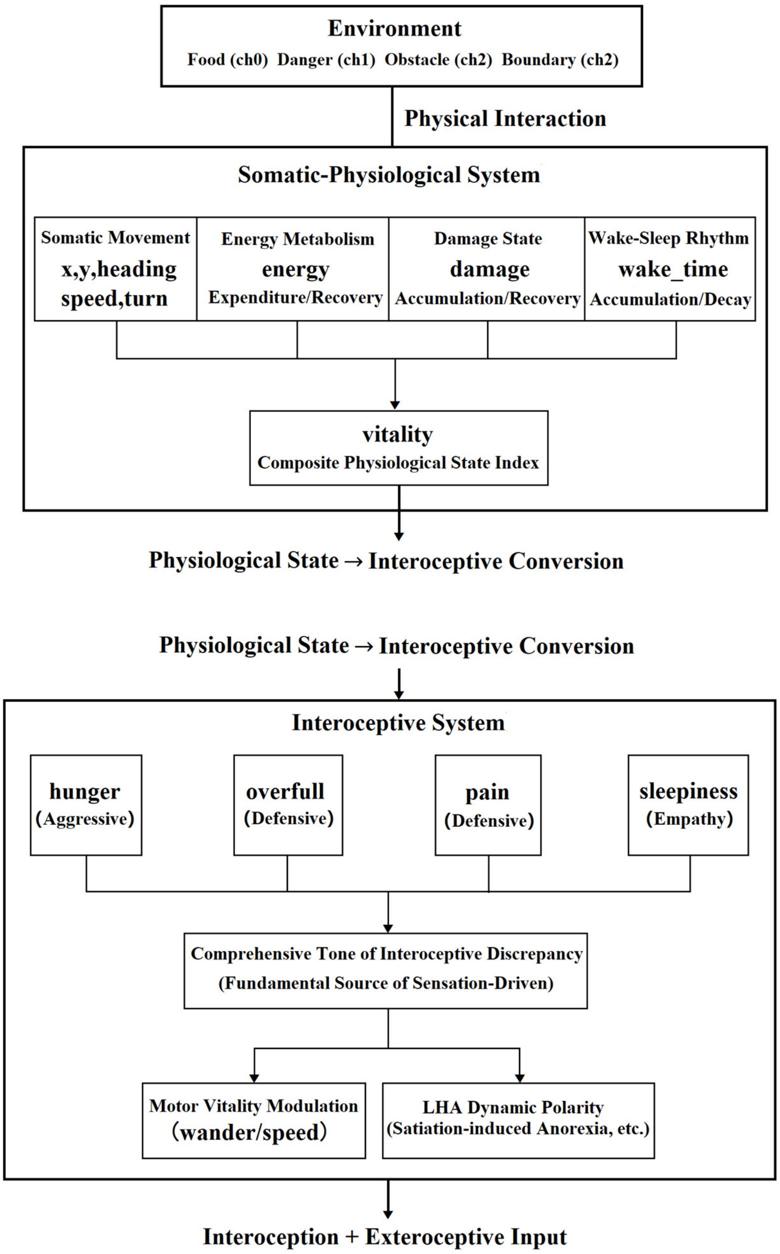
The acquired stratum consists of capabilities the organism gradually develops in interaction with the world, including: vision→value weights (converting visual features into Anticipated Value Sensation); vision→motor weights (converting visual features into modulation of turning and speed); and Conditioned Reflex triggering (when threat-related traces are consolidated to sufficient strength, vision alone can trigger a Conditioned Reflex). All acquired connections start from zero and are consolidated entirely through experiential selection.

The innate and acquired strata do not operate independently of each other but are in continuous interaction. This interaction is unidirectional: **innate modulates acquired**. Innate interoception modulates acquired learning by influencing the intensity and polarity of LHA pulses; the same event under different interoceptions produces different learning effects. Innate interoception also influences acquired behavioral drive by modulating behavioral drive—the same visual anticipation produces different intensities of behavioral output under different interoceptions. Acquired learning does not change innate rules: changes in value weights do not affect interoceptive computation rules; changes in motor weights do not affect the action sequences of contact Reflexes. This irreversible direction of modulation ensures that the growth of Value System Sensation remains always anchored in the innate value baseline.

3. SBLB-2.5: Design of a Minimal Sentient Model

As elucidated in Section 1.3, SBLB-2.5 simulates “the sentient system in its Sensation-Behavior Loop within its world.” Under this ontological premise, every principle of its architectural design strictly adheres to the formalization constraints of Section 2. It is designed as a minimal sentient model, containing only the most minimal mechanisms necessary to test the theory.

Figures 3.1 show its overall architecture—the three layers of the somatic-physiological system, the interoceptive system, and the neural system form a closed “Sensation-Behavior Loop.” Its core design principles are **innate/acquired separation** and **the dual value pathway**.



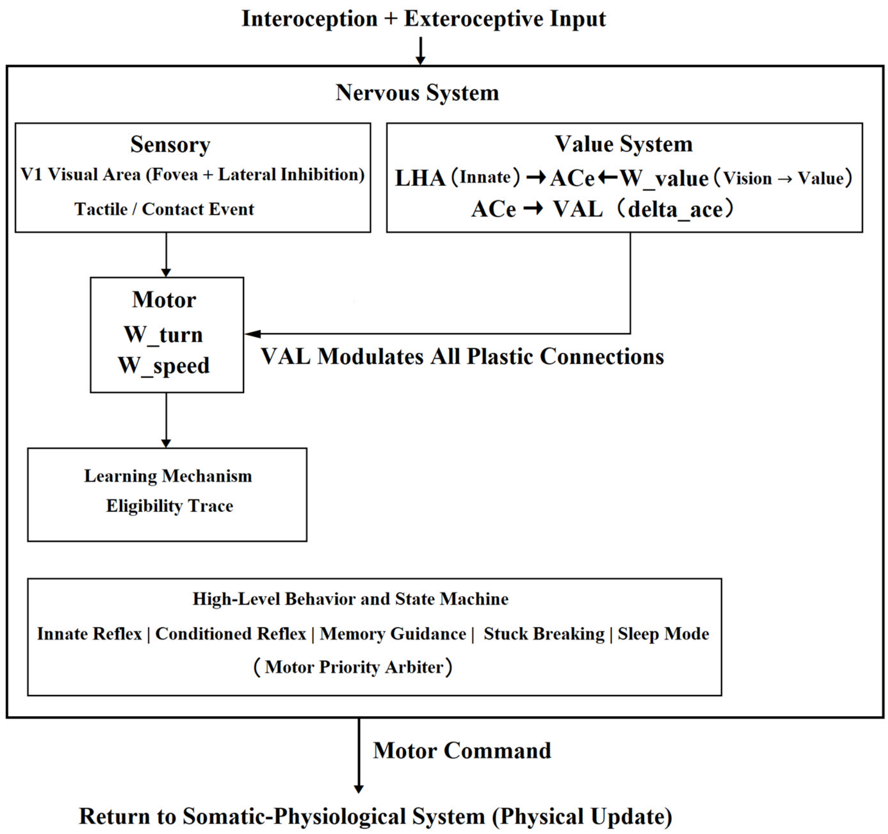


Figure 3.1 SBLB-2.5 System Architecture Diagram

The principle of minimal sufficiency: vision is merely an 11×11×3 grid; interoception encompasses only four types (hunger, overfullness, pain, sleepiness); behavioral output consists of only two variables (speed and turning rate); the learning mechanism has only three types of plastic traces and one endogenous learning signal. This simplification is to show that even with such impoverished sensory information, the system can still generate the predicted adaptive behaviors—proving the theory has grasped the essence.

No preset strategy: all plastic traces are initialized from zero; there is no preset “object-value” mapping table or behavioral strategy. “Green” initially evokes no value sensation or behavioral tendency at all; the emergence of foraging behavior is entirely the result of interoception-driven value flow and experiential selection.

Intrinsic value drive: the sole driving source for learning and behavior is endogenous Sensation Discrepancy (delta_ace). The positive LHA pulse generated upon contacting food is the interoceptive valence directly evoked by the food contact event itself; the negative LHA pulse generated upon contacting danger is the direct embodiment of the pain sensation caused by tissue damage. The system only learns events that truly change its inner sensory state.

3.1 Sensory System

Exteroception is an 11×11×3 visual grid. The three channels respectively encode food (green, ch0), dangers (red, ch1), and obstacles and boundaries (cyan, ch2). The field of view is 100°, with a range of 80 units. The raw visual signal undergoes two layers of fixed preprocessing—foveal weighting (central enhancement, peripheral attenuation) and lateral inhibition (contrast enhancement). All parameters are unlearnable, constituting the evolutionarily preset “sensory grammar.”

Interoception is the core for realizing “Innate Value System Sensation.” The system defines four interoceptive sensations: hunger and overfullness (Aggressive/energy balance), pain (Defensive/integrity of boundaries), and sleepiness (Empathy/rhythmic continuity). The intensity of interoceptive sensation has a nonlinear relationship with physiological quantities—the power function causes the sensation to accelerate upward near extreme states. The four interoceptive sensations are each independently computed, unlearnable, and are all continuously varying state variables—mild hunger is a weak negative value sensation; extreme hunger is an overwhelming negative value sensation. This continuity makes value flow possible.

3.2 Dual Pathway of the Value System

The value system of SBLB-2.5 distinguishes between the evaluative value pathway and the sensory value pathway, directly corresponding to the two levels of “acquiring value” in the Sentient Principle—the ontological level (forming traces of “good/bad”) and the dynamic level (the intensity of driving force sensed in the present moment).

In the evaluation paradigm, the state of satiation is often modeled as strategy switching by an independent module. SBLB-2.5 adopts a fundamentally different design: the modulation of value sensation by satiation is **embedded directly within the value system itself**. overfull_feeling directly participates in the dynamic computation of the LHA pulse upon food contact (and when high enough can cause LHA to switch from positive to negative) and modulates in real time the conversion efficiency of visual Anticipated Value into motor output—suppressing approach gain (food_drive_gain) and amplifying threat-avoidance gain (threat_drive_gain). Interoception is not an external “regulator” of value but a “constitutive element” of value.

The **evaluative value pathway** answers “this result—good or bad?” When the organism actually contacts an object or boundary, the system generates an immediate LHA pulse based on contact type and interoceptive state. The LHA for food contact is dynamic—lha_pulse = lha_food_base + hunger_boost - overfull_penalty - pain_suppress—and can switch from positive to negative when overfull. ACe fuses three layers of value sensation: the innate interoceptive tone, the value pulse of the contact event, and the acquired Anticipated Value Sensation (visual_value). delta_ace is the first-order difference of ACe and is the sole learning signal; all trace consolidation and decay are modulated by it.

The **sensory value pathway** answers "how much do I want/fear right now?" Visual Anticipated Value is separated into two independent components, food and threat, which compete at the motor layer. Interoception modulates drive in real time: when hungry, amplifying the positive value drive for food and inhibiting the negative value drive for threats (risk-taking); when overfull, inhibiting the positive value drive for food and amplifying the negative value drive for threats (conservative risk-avoidance); when in pain, amplifying the negative value drive for threats (highly risk-avoidant). The final motor command is directly generated from visual-motor traces and multiplied by the total modulation amount—from visual input to motor output, no "decision module" is interposed. Sensation is drive.

The two pathways are naturally separated in time: sensory value drives behavior (before contact) → behavior produces contact → evaluative value modulates traces (after contact). The system acts first and learns afterward.

3.3 Learning System

SBLB-2.5 implements what the Edelman school (Friston et al., 1994) called "Value-Dependent Selective Stabilization." Eligibility traces bind visual features to position and distance information; the decay rate determines the effective time window. During a Reflex, eligibility traces are frozen, preserving the visual memory from before the Reflex.

The attribution mechanism relies entirely on the spatiotemporal co-occurrence of eligibility traces—delta_ace acts on the activated eligibility traces in all channels, with no channel locking. The precision of attribution is naturally ensured by spatiotemporal decay and positional encoding. The innate value baseline is embodied through the channel-specific polarity of LHA and the amplification coefficients (2.0× for danger/obstacle channels), ensuring the asymmetric principle that "negative consolidation can often be profound after a single occurrence, while positive consolidation typically requires multiple repetitions"—without needing preset channel label locking. Trace changes are determined by the product of delta_ace and eligibility trace; the system does not converge toward an "optimal solution," but selectively consolidates patterns marked by positive value sensation.

3.4 Behavioral System and Rhythmic System

Behavior follows a hierarchical motor priority: Innate Reflex (contact bounce-away) > post-Reflex inertia > forced memory guidance > Conditioned Reflex (visual trigger stop-turn-resume) > autonomous movement. This is not the "choice" of a central decision-maker but the direct competition of different closed-loop tensions.

The three types of Innate Reflexes directly correspond to the Defensive tension. The Conditioned Reflex is a typical example of Acquired Anticipatory Sensation directly driving behavior—negative value anticipation emerges, and avoidance behavior is immediately initiated. Food Memory is refreshed only when the food value is positive, stored in egocentric coordinates with motion compensation, embodying the idea that "Memory is the re-sensing of traces." Autonomous movement is composed of wander tension superimposed with visual acquired traces; interoception modulates wander vitality in real time. Boundary restlessness and stuck-breaking mechanisms respectively solve the problems of wall-following and the "Buridan's ass" impasse.

The Empathy tension is implemented through the sleep-wake cycle. Sleepiness is driven by the accumulation of wake time, reaching a threshold for automatic sleep entry, and decaying to a threshold for natural awakening. During sleep, the metabolic rate drops to 20% of the waking rate, and damage recovery doubles. This rhythm is entirely autonomously regulated by the accumulation and dissolution of interoceptive sensation, requiring no external zeitgeber.

4. Experimental Design and the Logic of Theoretical Testing

4.1 Overall Experimental Scheme and Theoretical Concerns

The experiment adopts a five-stage environmental sequence design of progressive complexity. The organism sequentially faces ecological niches with different resource densities and threat densities across five continuous stages.

The initial environmental design conforms to biological facts. Stage One (food 25, danger 1, obstacle 1) simulates the typical ecological niche of early life development—parents tend to place eggs or young in areas with abundant food and low predator density. The juvenile period of higher mammals is usually spent in a "plenty-low risk" environment protected by parents, allowing the organism to establish basic foraging skills and value recognition with low survival risk early in life. In this environment, the organism establishes basic positive value recognition of food, forming the initial embodiment of Acquired Value System Sensation. Subsequent stages gradually introduce higher densities of threats and diminishing food resources, simulating the organism's migration and adaptation to more complex ecological niches—corresponding in nature to the complex environment faced by juveniles after independence.

Stage	Food Count	Danger Count	Obstacle Count	Source of Inherited Weights
One	25	1	1	Zero Initialization
Two	25	15	15	Stage One Final Weights
Three	15	15	15	Stage Two Final Weights
Four	10	15	15	Stage Three Final Weights
Five	5	15	15	Stage Four Final Weights

The continuous inheritance of weights simulates the experiential continuity of real life. The plastic traces (W_turn, W_speed, W_value) are inherited sequentially across the five stages rather than reset, corresponding to the real process of a single individual continuously accumulating Sensation-Behavior Traces over a lifetime. The organism at each stage enters the new environment carrying all the learning outcomes from all previous stages. This allows the experiment not only to test theoretical predictions within a single stage but also to track the cross-stage growth trajectory of Value System Sensation, observing the complete process of continuous differentiation and strengthening of Acquired Value System Sensation across continuous experience.

The theoretical concerns of the experiment focus on the testability of the three core claims:

Claim One (Value is Sensation): Can Acquired Value System Sensation grow from zero? Can Anticipatory Sensation already stably exist before contact occurs? Do plastic traces start from zero and gradually differentiate with experience into positive and negative anticipations?

Claim Two (Value Flows in Real Time): Can the same physical stimulus produce diametrically opposite value polarities under different interoceptive states? Can value reversal naturally emerge under overfullness? Does the LHA pulse change continuously with interoceptive states?

Claim Three (Sensation Directly Drives): Can value sensation directly modulate behavioral output without going through an independent decision module? Can acquired Conditioned Reflexes gradually replace innate Reflexes? Do the three-channel weights exhibit structural directional separation of approach and avoidance?

A supplementary test focuses on the Empathy tension: Can the sleep-wake cycle be entirely autonomously regulated by the accumulation and dissolution of interoceptive sensation, forming a stable rhythmic oscillation?

4.2 Testing Logic and Observational Metrics

This section declares the core predictions, main observational metrics, expected supporting evidence, and counterfactual conditions for each claim. If the theory is true and the model design strictly follows the formalization constraints, the following predictions should be supported by observational data.

Testing Framework for Claim One (Value is Sensation)

Claim One asserts that Acquired Value System Sensation is formed gradually through the selective consolidation of Sensation-Behavior Traces by endogenous value signals (LHA pulses) under the guidance of the innate value baseline, rather than being externally given. Core constraints include: plastic traces initialized from zero; the formation of acquired value must rely entirely on endogenous value signals, with no form of “direct assignment” or “supervised learning”; after learning is mature, visual input alone should be sufficient to evoke value sensation, prior to actual contact.

Prediction 1.1: The W_value three channels gradually differentiate from zero into positive values (food) and negative values (danger/obstacle). The observational metric is the time curve of the means of the three-channel W_value. In the early stage, the means of all three channels are close to zero; the organism is “blind” to visual signals. In the middle stage, the food channel begins to rise steadily, and the danger/obstacle channels begin to deviate negatively. In the late stage, the three channels are completely differentiated. Counterfactual condition: If the learning rate were set to zero, all plastic traces should always remain zero.

Prediction 1.2: After Acquired Value System Sensation is formed, Anticipatory Value Sensation stably exists before contact occurs and precedes actual contact. Using the early zero-weight stage (contact pre-food_val = 0) as the baseline reference, the observational metric is the change in food_val and threat_val within the ±20-step time window around the event. During the 20 steps before contact, food_val is continuously positive and close to its steady-state value; threat_val is stably negative before contacting danger or obstacle. Counterfactual condition: If the endogenous value signal (LHA) were fixed at zero, plastic traces should always remain zero, and no Anticipatory Sensation could be produced.

Testing Framework for Claim Two (Value Flows in Real Time)

Claim Two asserts that there is no objective, fixed truth of value. Value is a function of the organism's current interoceptive state; the same stimulus can evoke diametrically opposite value sensations under different states. Core constraints include: the value polarity of contact events cannot be a preset constant but must be a function of interoceptive states; the system must allow the value polarity of the same type of contact event to reverse under different interoceptive states.

Prediction 2.1: The LHA pulse of food contact changes continuously with interoceptive states and can turn negative during overfullness. The observational metric is the temporal linkage between LHA pulse, overfullness, and hunger. When hunger is high, the LHA pulse is stably in the strong positive range; when overfullness is significant, the LHA pulse declines from positive and turns negative. Counterfactual condition: If the LHA pulse were fixed as a constant, the value reversal under overfullness should disappear completely.

Prediction 2.2: Satiation-induced anorexia naturally emerges—the organism actively reduces foraging after energy is sufficient. The observational metric is the frequency of food contact under overfullness and the change in food Anticipated Value (food_val). In stages with higher overfullness, the proportion of overfull contacts is significantly higher than in hungry stages; during overfull contacts, food_val generally decays to zero or near zero. Counterfactual condition: If the inhibitory modulation of food value by overfullness were removed, satiation-induced anorexia should disappear.

Prediction 2.3: The reversal of value polarity is a continuous process driven by systematic changes in interoceptive states, not strategy switching. The observational metric is the cross-stage change in the proportion of overfull contacts and terminal overfullness across the five stages. The proportion of overfull contacts gradually decreases from high to low, while LHA polarity transitions step by step from frequently negative to constantly positive. This trajectory reflects the systematic evolution of interoceptive states, not any "strategy switching."

Testing Framework for Claim Three (Sensation Directly Drives)

Claim Three asserts that value as "sensation" can directly modulate behavioral output without going through an independent decision module. Core constraints include: the system must not contain an independent module that receives value information, conducts trade-off calculations, and then outputs behavioral commands; value sensation should be able to directly modulate the direction and intensity of motor output.

Prediction 3.1: The column sums of W_turn and W_speed three channels exhibit structural directional separation. The observational metric is the column-sum heatmap of the final weight snapshot. The food channel manifests central positive values (approach); the danger/obstacle channels manifest central negative values or lateral directional encoding (avoidance).

Prediction 3.2: Conditioned Reflexes are triggered before danger contact, and the number of triggers is significantly greater than the number of danger contacts. The observational metric is the ratio of Conditioned Reflex triggers to danger/obstacle contact events. In the mature stage, the number of Conditioned Reflex triggers far exceeds the number of actual contacts, indicating that the organism initiates the avoidance sequence while still in the near visual zone. Counterfactual condition: If the gain modulation from value sensation to motor output were frozen, Conditioned Reflexes should be difficult to trigger.

Prediction 3.3: Risk preference dynamically adjusts in real time with interoceptive states. The observational metric is the difference in response to threats by the same organism in hungry vs. overfull states. When hunger is high, even if threat anticipation is positive, the organism may still risk approaching food; when overfullness is high, approaching food also generates a negative LHA pulse, and the organism actively maintains distance. Counterfactual condition: If the modulation by interoception on the value-to-motor conversion efficiency is removed, the dynamic adjustment of risk preference should disappear.

Prediction 3.4: The turning amplitude of the Conditioned Reflex is proportional to the intensity of negative value anticipation; strong negative value anticipation can cause excessively large turning amplitudes, leading to secondary contact with lateral obstacles or boundaries. The observational metrics are the turning amplitude during Conditioned Reflex events, the intensity of negative Anticipated Value (threat_val), and whether the Conditioned Reflex is followed by contact with an obstacle or boundary. Supporting evidence: In threat-dense environments, high-intensity negative value anticipation triggers large-amplitude turns, and some Conditioned Reflex events are followed by obstacle or boundary contact. This phenomenon is not a theoretical anomaly or system flaw but positive evidence for “the intensity of sensation directly driving behavior”—the drive is sufficiently strong, strong enough to effectively evade the primary threat while simultaneously producing secondary collisions. Counterfactual condition: If behavioral intensity were decoupled from the intensity of value sensation, this phenomenon should disappear.

Supplementary Test: Sleep-Wake Rhythm of the Empathy Tension

Prediction 4: The sleep-wake cycle is entirely autonomously regulated by the accumulation and dissolution of interoceptive sensation, forming a stable rhythmic oscillation. The observational metrics are the periodic changes in sleepiness (sleepiness_feeling) and sleep state (is_sleeping).

4.3 Experimental Environment and Parameter Settings

The experimental environment is a square plane of 400 × 400 units. Object types include food (green, ch0), dangers (red, ch1), and obstacles (cyan, ch2); the quantity configuration of each type across the five stages is as shown in the table in Section 4.1. All objects are randomly generated within the environment, avoiding the central birth area and maintaining a minimum spacing between objects (20 units).

The organism has initial energy 80 (maximum 100) and initial damage 0. Basal metabolic rate 0.2/sec, movement cost coefficient 0.01, damage recovery rate 0.2/sec. Maximum speed 20.0 units/sec, maximum turning rate 180 degrees/sec. The ideal energy ratio (energy level/energy maximum) is set at 0.8; below this value hunger arises; above 90%, overfullness arises. Maximum wake time 120 seconds; sleep entry sleepiness threshold 0.9; wake threshold 0.2. All parameters remain consistent across the five stages; only the environmental configuration changes.

Regarding the learning system, the motor trace learning rate is 0.05, the value trace learning rate is 0.05, the minimum learning threshold is 0.01 (ACe changes below this absolute value do not trigger learning), the trace per-frame decay rate is 0.9999, and the trace absolute value cap is 2.0. The danger channel trace amplification coefficient is 2.0, and the obstacle channel is 2.0, reflecting the speed gain of the innate value baseline for defensive learning. All plastic traces are initialized from zero.

The visual system uses an 11×11×3 channel grid encoding. Field of view is 100°, visual range 80 units. The raw visual signal undergoes two layers of fixed preprocessing—foveal weighting (central enhancement, peripheral attenuation) and lateral inhibition (contrast enhancement). All parameters are unlearnable.

The modulation parameters of interoception on LHA are: hunger enhancement of food LHA ×1.0, overfullness penalty on food LHA ×2.0, pain suppression of food LHA ×0.5. These modulation parameters are consistent with those configured in the initial draft of this paper.

4.4 Data Acquisition and Recording

Periodic recording is triggered every 1,000 steps, recording a full snapshot of plastic traces (W_turn, W_speed, W_value) and their statistics (mean, peak, minimum, and per-channel means), as well as internal state (ACe, LHA pulse, energy, damage) and interoceptive states (hunger, overfullness, pain, sleepiness, sleep state).

Event real-time recording is triggered by contact events, Innate Reflex events, Conditioned Reflex events, food Memory trigger events, and sleep events, recording the complete state vector at the moment of the event (position, heading, ACe, vitality, energy, damage, interoception, visual value, LHA pulse, speed, and turning rate).

Recording of the peri-event window is one of the key designs of this experiment. Each time a food contact event is triggered, the system saves continuous timeline data from 20 steps before the event to 20 steps after the event (a total of 41 frames), including food value (food_val), threat value (threat_val), LHA pulse, overfullness, and hunger, for micro-temporal analysis of the emergence and flow of Anticipatory Sensation.

Trajectory recording notes the organism's position and heading every 10 steps, for visualization and qualitative analysis of behavioral patterns. All data are saved in standardized formats (data.npz and events.txt), ensuring that offline analysis can fully reproduce all experimental results.

5. Results: Systematic Observation of Emergent Behavior

This section organizes experimental results by claim; each subsection corresponds to the testing evidence for one core claim. The five-stage experiment ran continuously for over 550,000 steps; plastic traces were initialized from zero, with no mid-course intervention. All observational results below come from the continuous experiential accumulation of a single organism.

5.1 Evidence for Claim One: Value Grows from Sensation

Claim One predicts that Acquired Value System Sensation grows from zero, and that Anticipatory Value Sensation is stably formed before contact occurs. The following evidence unfolds along the temporal dimension.

5.1.1 Differentiated Growth of the Three W_value Channels

Figure 5.1 shows the complete evolution of the means of the three W_value channels across the five stages. The plastic traces of all channels start from zero. In the early stage, the means of all three channels are near zero; the organism is "blind" to visual signals. In the middle stage, the food channel begins to rise steadily, and the danger and obstacle channels begin to deviate from zero. In the late stage, the three channels are completely differentiated.

Table 5.1 summarizes the three-channel means at the end of each of the five stages. The food channel rises continuously, from +0.038 in Stage One to +0.099 in Stage Five—an increase of approximately 160% across five stages—and has still not saturated by the end of the experiment. The danger channel begins to deviate negatively from zero in Stage One, reaching -0.023 by Stage Five. The obstacle channel, from +0.014 in Stage One (because the negative learning signal from boundary contacts was not yet sufficient to cover the generalization traces on all obstacle channels with fewer obstacles and dangers in Stage One), flips to negative value after the first obstacle contact in Stage Two, reaching -0.029 by Stage Five. The step-like changes in weights align temporally with contact events in each stage; the differentiation direction of the three channels is formed under the guidance of the innate value baseline—the food channel accumulating positive values, the danger/obstacle channels accumulating negative values.

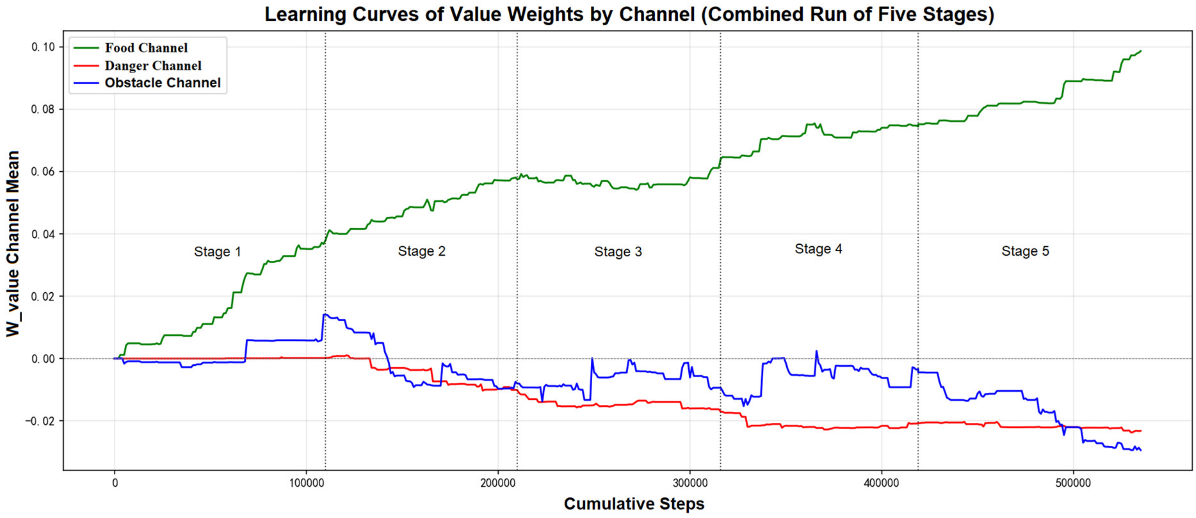


Figure 5.1 Complete Evolution of the W_value Three-Channel Means Across Five Stages

Table 5.1 Evolution of Three-Channel W_value Means Across Five Stages

Stage	Food Channel (ch0)	Danger Channel (ch1)	Obstacle Channel (ch2)
End of One	+0.038	+0.000	+0.014
End of Two	+0.057	-0.010	-0.008
End of Three	+0.064	-0.017	-0.009
End of Four	+0.075	-0.021	-0.004
End of Five	+0.099	-0.023	-0.029

5.1.2 Event-Window Evidence for Anticipatory Value Sensation

The event-window data directly confirm that Anticipatory Sensation stably exists before contact and continues to strengthen. Figure 5.2 compares the ±20-step timelines for three representative food events. For the first food contact in Stage One (step 2,325), food_val is 0.000 for all 20 steps before contact—the organism is completely “blind” to green objects at this moment; the value of food has not yet been embodied into visual features. By Stage Two (step 11,526), the organism has accumulated eating experience across multiple stages; food_val is stably in the range of +0.37 to +0.39 for the 20 steps before contact and rises monotonically, indicating that the organism already “anticipates” the arrival of food when seeing it. By the end of Stage Five (step 113,501), food_val monotonically increases from +0.690 to +0.721 over the 20 steps before contact—the anticipation intensity has nearly doubled since Stage One, and all of this occurs before actual contact.

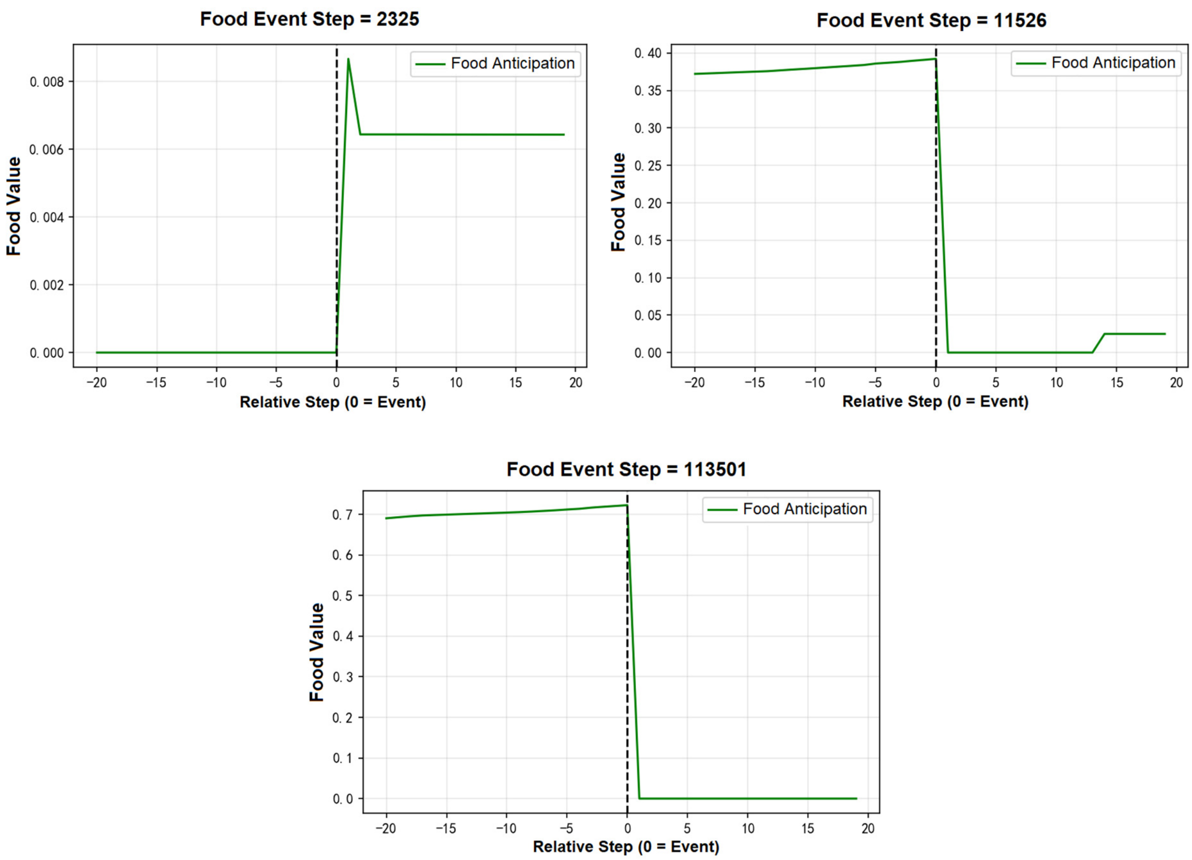


Figure 5.2 Emergence and Strengthening of Anticipatory Value Sensation: ±20-Step Timelines for Three Representative Food Events

Symmetrically, negative Anticipatory Sensation (threat_val) is also stably negative before contact. In the mature stage, when threat features appear in the organism’s near visual zone, threat_val drops to negative values before contact, enabling the Conditioned Reflex to trigger the stop-turn-resume sequence before contact.

5.1.3 Plastic-Trace-Independence of Innate Value Sensation

Even in the early stage when all plastic traces are zero (the beginning of Stage One), the organism can still maintain basic survival through innate mechanisms: interoception-driven wandering sustains movement, and contact Reflexes effectively bounce away when encountering boundaries. Over the entire more than 550,000-step run, the organism experienced only 2 danger contacts (1 in Stage Two, 1 in Stage Three) and never starved to death or died from severe injury. The operation of Innate Value Sensation does not depend on the maturity of acquired plastic traces; the two are mutually independent yet synergistic—innate provides direction, acquired provides precision.

5.2 Evidence for Claim Two: Value Flows in Real Time

5.2.1 Interoceptive Dependence of LHA Pulse Polarity

Across the five stages, a full dynamic range of LHA pulses from strongly positive to strongly negative of approximately 1.6 units is observed. Table 5.2 presents typical comparisons of food-contact LHA pulses under different interoceptive states. For the same type of contact event (food), solely due to different interoceptive states, the polarity of the value pulse is diametrically opposite: when hunger is high, the LHA pulse is stably in the strong positive range (can exceed +0.90); when overfullness is significant, the LHA pulse declines from positive and turns negative (can reach -0.49). The LHA pulse is not a constant but a continuous function of the interoceptive state.

Table 5.2 Comparison of Food-Contact LHA Pulses Under Different Interoceptive States

Stage	Step	Hunger	Overfullness	LHA Pulse	food_val (Pre-Contact)
One	40,874	0.53	0.00	+0.92	+0.01
One	80,895	0.00	0.30	-0.09	+0.38
Three	5,101	0.00	0.52	-0.49	+0.45
Five	71,189	0.49	0.00	+0.91	+0.64
Five	101,187	0.58	0.00	+0.92	+0.69

5.2.2 Value Reversal and Behavioral Inhibition Under Overfullness

In Stage One (25 food items), food value anticipation was still in its early formative period (end-of-stage W_value food mean merely +0.038). The anorexic behavior that appeared during overfullness stemmed mainly from the direct modulation of contact LHA by the innate value baseline—the overfullness feeling suppressed the LHA pulse via the overfull_penalty, leading to inhibition of eating behavior. This is direct evidence that Innate Value System Sensation can modulate behavior independently of acquired plastic traces.

Value reversal in the strict sense—where an already-formed acquired positive food value anticipation is overturned under overfullness—appears in the mid-to-late part of Stage Two and in Stage Three. At this point, after the experiential accumulation of Stage One (110,001 steps) and the early part of Stage Two, food value anticipation had significantly strengthened (end-of-Stage-Two food_val approximately +0.48 to +0.67; Stage Three approximately +0.50 to +0.70). Under overfullness, the previously firmly established Anticipatory Sensation that "food = positive value" is reversed by the interoceptive state: LHA pulses switch from positive to negative, and food_val decays sharply after contact. Step 5,101 in Stage Three provides a landmark case: overfullness as high as 0.52, LHA pulse plummeting to -0.49, while the expected value of food Anticipated Value was around +0.45 to +0.50—the anticipated "good" is negated by the experienced "not good." At the behavioral level, during peak overfullness periods, the organism exhibits marked behavioral inhibition toward food—reduced speed, scattered turning, with food contacts more often resulting from random wandering than active tracking. This is the quantitative depiction of "turning a blind eye": from the perspective of external behavioral efficiency, this appears to be "inefficiency"; but from the perspective of the Sentient Principle, this is precisely the hallmark of successful Value System Sensation modulation—continuing to eat would exacerbate rather than alleviate inner tension.

5.2.3 Cross-Stage Evolution of Value Polarity

Table 5.3 summarizes the systematic evolution of the food LHA pulse polarity distribution across the five stages. From Stage One to Stage Five, the proportion of overfull contacts drops from approximately 40% to 0%, and LHA negative events go from "multiple" to "zero." This cross-stage evolution is not a result of weight changes but a direct consequence of the food resource density driving the interoceptive state—the relative proportion of overfullness to hunger—undergoing systematic change, realized through the real-time computational rule of the LHA pulse. In Stage One, food was abundant; after eating one's fill, continuing to encounter food led the overfullness feeling to suppress the LHA pulse, causing it to flip to negative frequently. In Stage Five, food was extremely scarce (merely 5 items); the organism was perpetually in a state of energy deficit, hunger became the enduring dominant drive, each eating event was "timely assistance," and the LHA pulse consistently maintained high positive values. Throughout the entire process, no strategy switching or parameter resetting was needed—only the interoceptive state changed, and the interoceptive state was determined by the objective dynamics of food acquisition rate and basal metabolic rate.

Table 5.3 Cross-Stage Evolution of Food LHA Pulse Polarity Distribution Across Five Stages

Stage	Food Count	Overfull Contact Proportion	LHA Negative Events	Typical LHA Range
One	25	~35-40%	Multiple	+1.00 to -0.16
Two	25	~35%	Multiple	+0.72 to -0.64
Three	15	~20%	Peak -0.49	+0.77 to -0.49
Four	10	<5%	Sporadic	+0.92 to -0.38
Five	5	0%	Zero	+1.00 to +0.55

5.3 Evidence for Claim Three: Value Directly Drives Behavior

5.3.1 Structural Directional Separation of W_turn/W_speed Three Channels

Figure 5.3 shows the complete directional encoding of the final W_turn three-channel column sums at the end of Stage Five. Food channel (ch0): the left columns are negative (driving right turns), the right columns are positive (driving left turns), together forming an approach turning pattern—centering the food directly ahead. Danger channel (ch1): the left columns are positive (driving left turns), the right columns are negative (driving right turns), together forming an avoidance turning pattern—moving the danger out of the field of view. The obstacle channel (ch2) pattern is the same as the danger channel, with slightly smaller amplitude. The W_speed three channels exhibit a complementary pattern: the central columns of the food channel are positive (accelerating approach), while the central columns of the danger/obstacle channels are negative (decelerating or reversing).

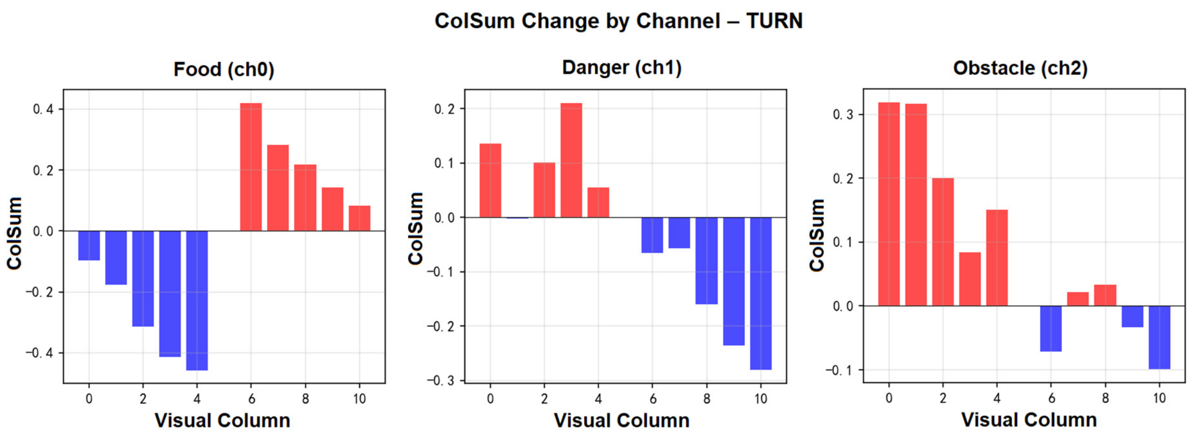


Figure 5.3 Directional Separation Heatmap of the Final W_turn Three-Channel Column Sums at the End of Stage Five

The three-channel column sums exhibit perfect structural separation—positive value sensation consolidates approach traces, negative value sensation consolidates avoidance traces. The positive and negative polarity of value sensation directly maps onto the differentiation of behavioral direction; from visual input → value sensation modulation → motor output, no independent "decision" step is interposed throughout the entire process. This is the structural evidence at the level of neural traces for "sensation directly driving behavior."

5.3.2 Emergence of Conditioned Reflexes and Retreat of Innate Reflexes

Table 5.4 Succession of Conditioned Reflexes, Innate Reflexes, and Danger Contacts Across Five Stages

Stage	Conditioned Reflex Events	Innate Reflex Events	Danger Contact Events	CR/Danger Contact Ratio
One	186	3	0	∞
Two	397	8	1	397:1
Three	240	3	1	240:1
Four	232	2	0	∞
Five	232	2	0	∞

Table 5.4 summarizes the changes in the counts of Conditioned Reflexes, Innate Reflexes, and danger contacts across the five stages. Across the entire five stages, only 2 danger contacts occurred, both in the second and third stages when threat levels first increased significantly. The counts of Conditioned Reflexes and Innate Reflexes exhibit a reciprocal relationship: innate Reflexes steadily decline from 8 in Stage Two to 2 in Stage Five, while Conditioned Reflexes remain consistently high.

Particularly critical is the enormous ratio of Conditioned Reflexes to danger contacts. In Stage Two, 397 Conditioned Reflexes corresponded to only 1 actual danger contact (397:1). From Stage Four onward, danger contacts dropped to zero, yet Conditioned Reflexes continued to operate efficiently. This means: the organism was already triggering the stop-turn-resume avoidance sequence while still in the near visual zone, completing threat evasion while the threat was still within visual range and before bodily contact occurred. In the extremely scarce Stage Five (merely 5 food items, 15 danger items), not a single danger contact occurred over the entire stage—the Defensive value sensation had completely achieved efficient pre-contact behavioral regulation through the Conditioned Reflex mechanism. Acquired Anticipatory Sensation drives behavior before contact; this is the best behavioral-level example of “sensation directly driving behavior.”

5.3.3 Real-Time Modulation of Behavioral Drive Gain by Interoception

The real-time modulation of behavioral drive by interoception receives continuous verification across the five-stage data. When hunger is relatively high, both the speed and risk preference of foraging behavior significantly increase. In Stage Five step 71,189 (hunger 0.49), speed reached the physical limit of 20.0, with a food drive gain of approximately 2.0×; simultaneously, although threat features were present in the field of view and threat_val remained negative (avoidance maintained), foraging behavior was not interrupted—the approach gain driven by hunger suppressed threat sensitivity (the modulation effect of hunger_threat_suppress = 1.5). The opposite state is seen in the overfull periods of Stage One: when overfullness was above 0.30, the food contact LHA flipped to negative, food_val decayed to zero or slightly negative after contact, and active approach behavior disappeared. The visual anticipation was the same, but the behavioral output was diametrically opposite; the single differential variable was the interoceptive state.

5.3.4 The “Overshoot” Phenomenon: Natural Evidence for the Intensity of Sensation Directly Driving Behavior

Across the five-stage data, a particular phenomenon was consistently observed: the turning amplitude of the Conditioned Reflex is proportional to the intensity of negative value anticipation. When negative anticipation (threat_val) is extremely high, the sharp turn executed to evade a front-facing threat sometimes causes the organism to collide with a lateral obstacle or boundary. This “overshoot” phenomenon was especially prominent in Stage One and Stage Two: of the 186 Conditioned Reflex events in Stage One, some were triggered by boundary signals; in the high-threat environment of Stage Two (15 dangers and 15 obstacles each), intense negative anticipation drove large-amplitude turns, with some turns immediately followed by obstacle or boundary contact.

From the traditional perspective of “control precision,” this phenomenon might be seen as “loss of control” to be optimized. But from the perspective of the Sentient Principle, it is precisely positive evidence for “the intensity of sensation directly driving behavior.” The intensity of negative value anticipation determines the magnitude of the turn—it is not that the system “computed” an optimal turning angle and then executed it, but that the intensity of the negative value sensation itself directly determined the severity of the avoidance behavior. The drive from sensation was sufficiently strong (because the negative anticipation was sufficiently strong) to effectively evade the primary threat while simultaneously producing secondary collisions. This is not a theoretical anomaly, but an inherent characteristic of an adaptive system that uses sensation as its direct driving force without calibration by an independent decision module. It confirms from another angle that SBLB-2.5 indeed does not contain a “decision-maker” to weigh the pros and cons of turning amplitude versus lateral risk—behavior is the direct extension of sensation, not the product of decision-making.

5.4 Evidence for the Empathy Tension: Autonomous Stabilization of Sleep-Wake Rhythm

Figure 5.4 shows the periodic oscillation of the sleep-wake rhythm over a typical time period. Sleepiness (sleepiness_feeling) gradually accumulates from approximately 0.20 during wakefulness to 0.90, triggering natural sleep onset, and decays rapidly during sleep to 0.20, triggering natural awakening. The complete sleep event records across the five stages confirm that: the full cycle is approximately 120 to 150 seconds; the entry and exit thresholds remain respectively stable around 0.90 and 0.20 across stages, unaffected by food density or threat density to any significant degree. During sleep, the energy decay rate drops to 20% of the waking rate, and the damage recovery rate doubles.

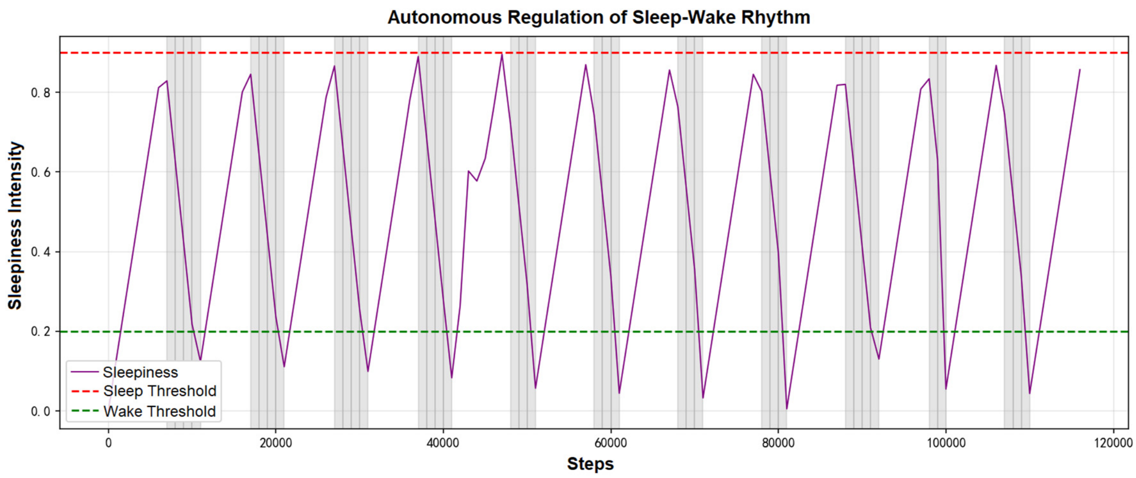


Figure 5.4 Periodogram of the Sleep-Wake Rhythm

This rhythm is entirely autonomously regulated by the accumulation and dissolution of interoceptive sensation, requiring no external time cues. This verifies the independent operational capacity of the Empathy tension (rhythmic continuity) as an Innate Value System Sensation—it is an intrinsic rhythm “carried” by the organism, not a passive response to an environmental timing signal.

5.5 Summary of Results

Table 5.5 summarizes the verification status of the three claims and the supplementary test. All core predictions are supported by data from the five stages.

Table 5.5 Summary of Verification Results for Core Predictions of the Five-Stage Experiment

Claim	Core Prediction	Cross-Stage Key Evidence	Verification Status
Claim One	W_value differentiates from zero	Food 0.038→0.099; Danger 0→-0.023; Obstacle -0.014→-0.029	Confirmed
	Anticipatory Sensation precedes contact	food_val 0→+0.72 (across five stages); threat_val stably negative before contact	Confirmed
	Innate value operates independently	Survival maintained in zero-weight stage; 0 starvation deaths throughout	Confirmed
Claim Two	LHA pulse changes with interoception	Cross-stage range approx. 1.6 units (+0.92 to -0.69)	Confirmed
	Value reversal & behavioral inhibition when overfull	LHA turns negative; food_val decays; active approach disappears	Confirmed
	Systematic evolution of polarity across five stages	Overfull contact proportion ~40%→0%; LHA negative events disappear	Confirmed
Claim Three	Directional separation of three channels	Spatial encoding of food approach/threat avoidance clear and stable	Confirmed
	Conditioned Reflexes replace innate Reflexes	Only 2 danger contacts / >550k steps; Conditioned Reflexes efficiently substitute	Confirmed
	Conditioned Reflex "overshoot" phenomenon	Strong negative value anticipation → large turn → secondary collision confirms sensation directly drives	Confirmed
	Interoception modulates behavioral drive	Speed maximized, risk preference elevated when hungry	Confirmed
Empathy	Sleep-wake rhythm autonomously stable	Cycle ~120-150 sec across five stages; thresholds stable, unaffected by environment	Confirmed

6. Discussion

6.1 The Mutually Defining Relationship Between the Sentient Principle and SBLB-2.5

The Sentient Principle and SBLB-2.5 form a mutually defining relationship. The theory provides the design blueprint for the model—the dynamic LHA modulation by interoception, the separation of value sources and spatiotemporal co-occurrence attribution, the dual value pathway—none of these are arbitrary engineering choices but precise translations of the core claims. The model provides precision and testability for the theory—concepts such as “Sensation Discrepancy” and “value flow” acquire operational definitions in the process of construction. Together, the two constitute a self-consistent selectionist formalization system, distinct from pure philosophical papers (which lack formalization) and from pure engineering papers (which lack ontological grounding).

The Reverse Precision-Refinement of Theory by Simulation

The mutually defining relationship between theory and simulation is not only embodied in the forward direction of “theory provides the design blueprint for simulation,” but equally in a more profound reverse direction: the precision-refinement of theory by simulation. In the process of converting the core mechanisms of the Sentient Principle into executable rules, several concepts that appear clear at the level of theoretical exposition reveal boundaries that require further precision when subjected to the interrogation of formalization. The construction process of SBLB-2.5 is therefore not merely a “translation” of the theory but a deepening and testing of the theory itself. The following are several typical instances of reverse precision-refinement:

First: the operational boundary of the concept of “Sensation Discrepancy.” The Sentient Principle defines Sensation Discrepancy as “the identifiable imbalance between immediate sensation and Anticipatory Sensation.” At the theoretical level, this definition is clear. But in the process of formalization, one must press further: what counts as “identifiable”? How does the magnitude of the imbalance determine the intensity of behavioral response? SBLB-2.5 provides an operational answer through the computational method of delta_ace: Sensation Discrepancy is the difference between the immediate LHA pulse intensity and the current Anticipated Value Sensation; its absolute value determines the intensity of trace consolidation, and its sign determines value polarity. This operational definition in turn demands that the theory further clarify: does Sensation Discrepancy have a “minimal identifiable threshold”? Do discrepancies in different tension dimensions possess different identifiable sensitivities? These questions were not urgent before simulation; after simulation, they become concrete questions that the theory must address.

Second: the stratification of the boundary between "innate" and "acquired." The Sentient Principle distinguishes Innate Value System Sensation from Acquired Value System Sensation, but in the process of formalization, "innateness" itself is forced to decompose into three distinct levels: innateness in the evolutionary sense (the very existence of the Three Basic Tensions), innateness in the architectural sense (fixed connections and ranges of initial parameters), and innateness in the parametric sense (the settings of concrete thresholds). This stratified decomposition is a direct result of the theoretical precision forced by simulation construction; it in turn demands that the Sentient Principle, in its future theoretical development, provide a more explicit account of the stratified nature of the concept of "innateness."

Third: the temporal scales of "value flow." The Sentient Principle claims that value flows in real time, but "real time" is a relatively vague temporal concept at the theoretical level. The operation of SBLB-2.5 reveals that value flow actually occurs simultaneously on at least two different temporal scales: the modulation of LHA by interoception is continuous, step-by-step flow (at the scale of seconds), whereas the trace consolidation of Acquired Value System Sensation is cumulative, gradual flow (at the scale of thousands of steps). The coexistence of these two temporal scales is a fact revealed by the simulation operation; it in turn demands that the theory provide a more finely differentiated account of the temporal structure of "value flow."

This ongoing bidirectional calibration between theory and simulation is precisely what makes the Sentient Principle a living research program rather than a static philosophical manifesto. Each round of simulation not only verifies the theory but also propels it to grow in the direction of greater precision and testability. This capacity for growth is the intrinsic driving force behind the SBLB series, from 1.0 to 2.5, and its continued evolution.

6.2 Systematic Comparison with Mainstream Frameworks

The significance of the comparison with mainstream paradigms lies not in listing mechanistic differences, but in revealing a fundamental divide: **Where does value reside?** Reinforcement learning, the free-energy principle, and the somatic marker hypothesis locate value outside the system—in the environment, in the prior distribution, in bodily signals—with the system "acquiring" it through detection, inference, or consultation. The Sentient Principle locates value in the sensory field itself: value is sensation, emerging in real time in the direct encounter between life and the world.

Versus Reinforcement Learning. The reward signal in RL is designed by the experimenter, is exogenous, and, once set, remains fixed; the goal of learning is to maximize cumulative reward. The Sensation Discrepancy of SBLB-2.5 is completely endogenous; the value polarity of the same food contact flows in real time with interoception; the goal of behavior is not "obtaining more" but "eliminating Sensation Discrepancy." In RL, ceasing to eat after satiation requires an additional mechanistic explanation—for example, introducing a satiety signal independent of the reward function to inhibit eating behavior. In such an architecture, value (food reward) and inhibition (satiety signal) belong to two separate subsystems. In the Sentient Principle, eating after satiation itself directly becomes a sensation of "not good"—this is not inhibition by an external signal, but a polarity reversal of value sensation itself. From the perspective of external behavioral trajectories, SBLB-2.5 in its mature stage is almost "unintelligible": only by tracking its inner sensory states can these behaviors receive a unified explanation. **The fundamental divide**: RL regards value as an objective quantity in the environment to be maximized; the Sentient Principle regards value as the real-time flow of an intrinsic constitutive dimension of the sensory field.

Versus the Free-Energy Principle. FEP models the organism as a Bayesian inference system that minimizes prediction error; prediction error is a third-person mathematical quantity; value is implicit in the prior distribution—what conforms to expectation is good. The Sentient Principle places Sensation Discrepancy at the ontological foundation: it is not that prediction error produces sensation, but that **Sensation Discrepancy is itself the fundamental dynamics of the system**. FEP can explain how prediction error drives behavior, but cannot answer why its reduction feels good—because within this framework, "good" has already been epistemologically defined as "conforming to expectation." The Sentient Principle's answer is sensory: good is the very experience of Sensation Discrepancy being eliminated. **The fundamental divide**: FEP attributes value to an epistemological structure of "what one should believe"; the Sentient Principle attributes value to an evolutionarily endowed, irreducible sensory fact.

Versus the Somatic Marker Hypothesis. Both emphasize the core role of bodily sensations, but the relationship is fundamentally different. In the somatic marker hypothesis, sensation is a signal that **assists** rational decision-making—the body sends markers, and an independently existing decision center consults them to make choices. In the Sentient Principle, no such value-neutral decision-maker exists: from sensation to behavior is continuous; value sensation directly modulates motor output, without going through any independent “trade-off” step. Moreover, somatic markers, once formed, are relatively stable, whereas in the Sentient Principle value polarity flows continuously with interoception—the value pulse generated by food contact when hungry is diametrically opposite to that when overfull—and the modulatory relationship between Innate and Acquired Value System Sensation is irreversible. **The fundamental divide**: The somatic marker hypothesis preserves the primacy of the decision center; the Sentient Principle dissolves the decision center itself through the primacy of sensation.

In summary: the three mainstream paradigms understand life as the **evaluator and executor** of value—value exists outside the system, and the organism’s task is to detect it, infer it, consult it. The Sentient Principle understands life as the **bearer and expresser** of value—value is the intrinsic constitution of the sensory field, and behavior is the unfolding of value sensation in the behavioral dimension. The former takes “efficiency” as its measure; the latter takes “coordination” as its natural state. This is not a difference of degree but a parting of ways between two conceptions of life.

6.3 Selectionism as a New Paradigm for the Sciences of Mind

The experimental results of SBLB-2.5 point toward a vaster possibility. Since the “cognitive revolution,” the metaphor that “the mind is a computer” has brought a persistent blind spot—sentience, first-person experience, within the computationalist framework, have always been “anomalies” rather than the theoretical core. Selectionism provides a different metaphor: **the mind is not software, but an ecosystem**. Neural connections are not code but species; learning is not optimization but selection; value is not an error signal but the embodiment of fitness at the sensory level.

This paradigm shift brings about an ontological transformation of value: value is not an object to be evaluated, but a content to be sensed. Is the “not-goodness” of hunger an evaluation of blood glucose levels? No. It is the sensation of hunger itself.

Implications for consciousness research. The “hard problem” of consciousness research—how physical processes produce subjective experience—is repositioned within the framework of the Sentient Principle. The basic forms of qualitative sensation (the “not-goodness” of hunger, the “unpleasantness” of pain) already exist at the Vital Stratum; they are the direct content of Innate Value System Sensation and do not need to be “produced” from non-sensing matter. What truly needs to be explained is not the “existence” of sensation, but the “expression” of sensation. The expression of sensation is not limited to language alone—the foraging, withdrawal, whimpering, and trembling of animals are themselves direct expressions of inner sensory states. They are not "translations" of sensation, but the natural extension of the Sensation-Behavior Loop: sensation drives behavior, behavior externalizes sensation. The emergence of language transformed this expression, originally carried by behavior, into expression carried by public symbols—"I am hungry" replaces the act of foraging itself, becoming the symbolic externalization of sensation. In this perspective, the "hard problem" is further dissolved into: how did the leap from "behavioral expression" to "symbolic expression" occur? This is one of the core research questions for the Noetic Stratum. SBLB-2.5, as a formal model of the Vital Stratum, has already demonstrated how sensation drives behavior; future models of the Noetic Stratum will further reveal how sensation is categorized and re-sensed.

Implications for artificial intelligence research. Mainstream value alignment methods inject value from the outside—collecting human preferences, training reward models, driving reinforcement learning. SBLB-2.5 has no reward function; its value sensation grows entirely from the encounter between interoception and the world. Hunger defines "acquiring energy is good"; pain defines "avoiding damage is good." These are not programmed "value labels" but the direct sensation of the organism's own homeostatic tensions. The implications of this line of thought are far-reaching: rather than injecting AI with a fixed set of "human value principles," it may be better to design for it a value-growth structure analogous to interoception—enabling the system not to "obey" values but to "have" values.

6.4 Limitations and Future Research Roadmap

As noted in Section 5.3.4, the Conditioned Reflex "overshoot" phenomenon observed in SBLB-2.5—intense negative value anticipation driving large-amplitude turns, sometimes leading to secondary contact with lateral obstacles or boundaries—reveals an inherent characteristic of the "sensation directly drives" architecture: behavioral intensity is proportional to the intensity of value sensation, and there is no independent decision module to perform fine-grained regulation. This characteristic is also widely present in natural life (the "headlong flight" in panic escape reactions is one example), and is theoretically self-consistent, though in engineering applications it may be necessary to explore additional regulatory mechanisms within a sensation-driven architecture (such as competition and coordination among different Defensive closed loops).

As a minimal sentient model, the limitations of SBLB-2.5 are equally clear. The sensory dimensions are extremely simplified; the Social and Noetic Strata have not yet been implemented. The eligibility trace attribution depends on spatiotemporal decay and the spatial locality of the visual grid; the bidirectional generalization observed in the experiment (threat anticipation acquiring positive traces due to co-occurrence with food, and food anticipation erroneously decaying due to co-occurrence with threats) reveals the precision limits of this mechanism when features are temporally aliased. Furthermore, the problem of verifying sentience is one that no theory centered on sentience can evade—what we observe are behaviors and variable values, not sensation itself. The position of this paper is to maintain the consistency of the first-person priority framework methodologically, without claiming to have solved the hard problem.

Future research can advance along the following directions. Near-term: a minimal sentient model of the Noetic Stratum—adding an "inner neural activity space" on the basis of SBLB-2.5 to study the emergence of Noetic-level Sensation Discrepancies such as perplexity and cognitive dissonance. Mid-term: a minimal sentient model of the Social Stratum, introducing multiple organisms to study the emergence of empathy and cooperative behavior. Long-term: dialogue with neuroscientific data, quantitatively comparing the model's core variables with dopamine neuron firing patterns and the activity of interoception-related brain regions.

7. Conclusion

This paper presents a complete theory-simulation system: the Sentient Principle and its first selectionist formalization instance—the SBLB-2.5 minimal sentient model. The core claim of this system can be condensed into a single sentence: **Value is not an object to be evaluated, but a content to be sensed.**

This paper clearly demarcates the essential difference between selectionism and computationalism, decomposes "Value System Sensation" into three independently testable claims, and elucidates the strict separation and irreversible modulatory relationship between Innate and Acquired Value System Sensation. Strictly adhering to the principle of minimal sufficiency, SBLB-2.5 fully implements the closed loop from sensory acquisition to behavioral output through the dual value pathway, the dynamic modulation of LHA by interoception, and spatiotemporal co-occurrence attribution with channel-specific amplification. Over more than 550,000 steps of continuous operation, systematic empirical support has been provided: plastic traces differentiate from zero into positive and negative anticipations; value polarity switches from positive to negative under overfullness; satiation-induced anorexia emerges in full; three-channel traces exhibit structural separation; the sleep-wake cycle autonomously stabilizes.

On a broader level, this paper demonstrates the feasibility of replacing computationalism with selectionism. SBLB-2.5 has no reward function; value grows entirely from the encounter between interoception and the world—this suggests a new path for the problem of AI value alignment, from "injecting values" to "growing values." This paper does not attempt to explain "why neural activity produces sensation," but takes sensory structure as the starting point for theoretical construction, providing a methodological detour around the hard problem for consciousness research.

The Sentient Principle and its selectionist formalization ultimately crystallize into three sentences: **Life is a system that ceaselessly eliminates "Sensation Discrepancy" within a "Sensation-Behavior Loop." Value is not a label programmed in, but a sensory fact that emerges in real time from the encounter between interoception and the world. Learning is not approximating an externally defined truth value, but the selective consolidation and decay of existing "Sensation-Behavior Traces" under the modulation of "Value System Sensation."**

With over 550,000 steps of emergent behavior, SBLB-2.5 provides the first selectionist formalized empirical support for these three sentences. From Whytt's proposal of the "Sentient Principle" in 1751, to Sechenov's "Systemic feelings," to the Edelman school's "Value-Dependent Selective Stabilization," and to the SBLB-2.5 of this paper—a thread of thought spanning more than two and a half centuries has, at this moment, acquired an executable, testable, and extensible formalized form. This is merely a beginning.

References

Wu, J. (2026). *The Sentient Principle*. Available at SSRN: https://ssrn.com/abstract=6241598

Whytt, R. (1751). *An Essay on the Vital and Other Involuntary Motions of Animals*. Hamilton, Balfour, and Neill.

Edelman, G. M. (1987). *Neural Darwinism: The Theory of Neuronal Group Selection*. New York: Basic Books.

Edelman, G. M. (1989). *The Remembered Present: A Biological Theory of Consciousness*. New York: Basic Books.

Friston, K. J., Tononi, G., Reeke, G. N., Sporns, O., & Edelman, G. M. (1994). *Value-dependent selection in the brain: Simulation in a synthetic neural model*. Neuroscience, 59(2), 229–243.

Damasio, A. R. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: Putnam.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement Learning: An Introduction* (2nd ed.). Cambridge, MA: MIT Press.

Friston, K. (2010). *The free-energy principle: a unified brain theory?* Nature Reviews Neuroscience, 11(2), 127–138.