

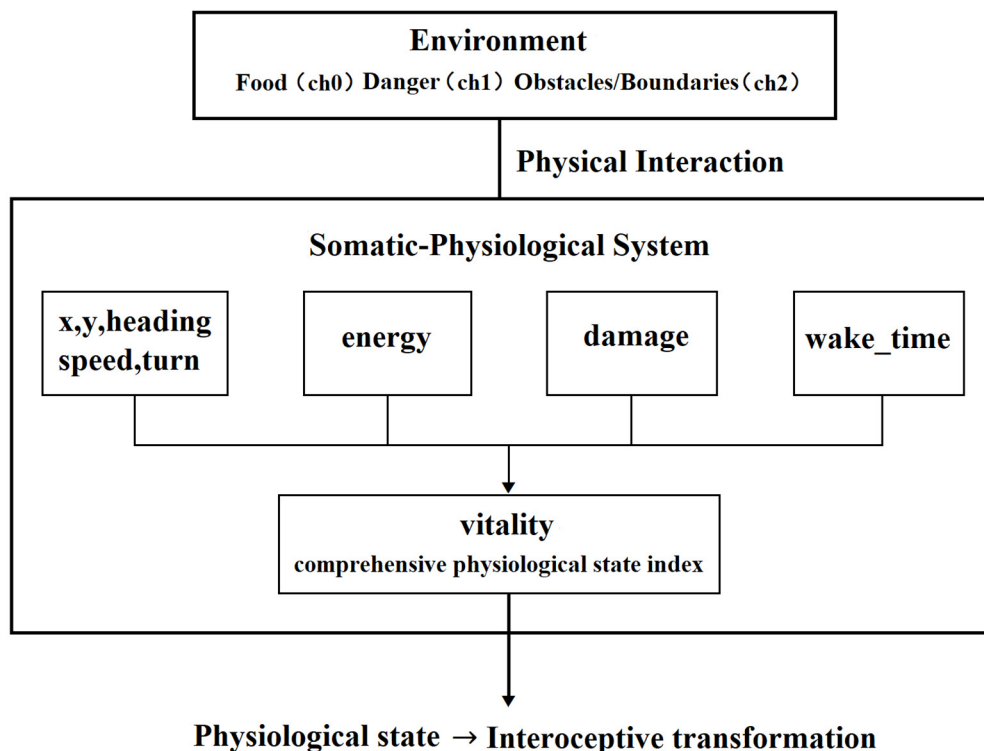
A Brief Description of the SBLB Simulation System Experiments

Wu Jingyu

1. SBLB: Design of a Minimal Sentient Model

SBLB (Sensor-Behavior Loop Being) simulates a complete **sensation-behavior loop of a sentient system in its world**. It is designed as a minimal sentient model: it contains only the simplest mechanisms necessary to validate the theory. Figure 1 shows its overall architecture – the body-physiology system, the interoceptive system, and the nervous system form a closed **sensation-behavior loop**. Its core design principles are **innate/acquired separation** and **dual value pathways** – the latter corresponding precisely to the fundamental distinction in the *Sentient Principle* between **objective value sensation** and **subjective value sensation**.

Principle of minimal sufficiency. The visual field is only an $11 \times 11 \times 3$ grid; there are only four interoceptions (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 intended to show that even with such impoverished sensory information, the system can still exhibit anticipatory adaptive behavior – strong evidence that the **sensation-behavior loop** is a universal feature of life organization.



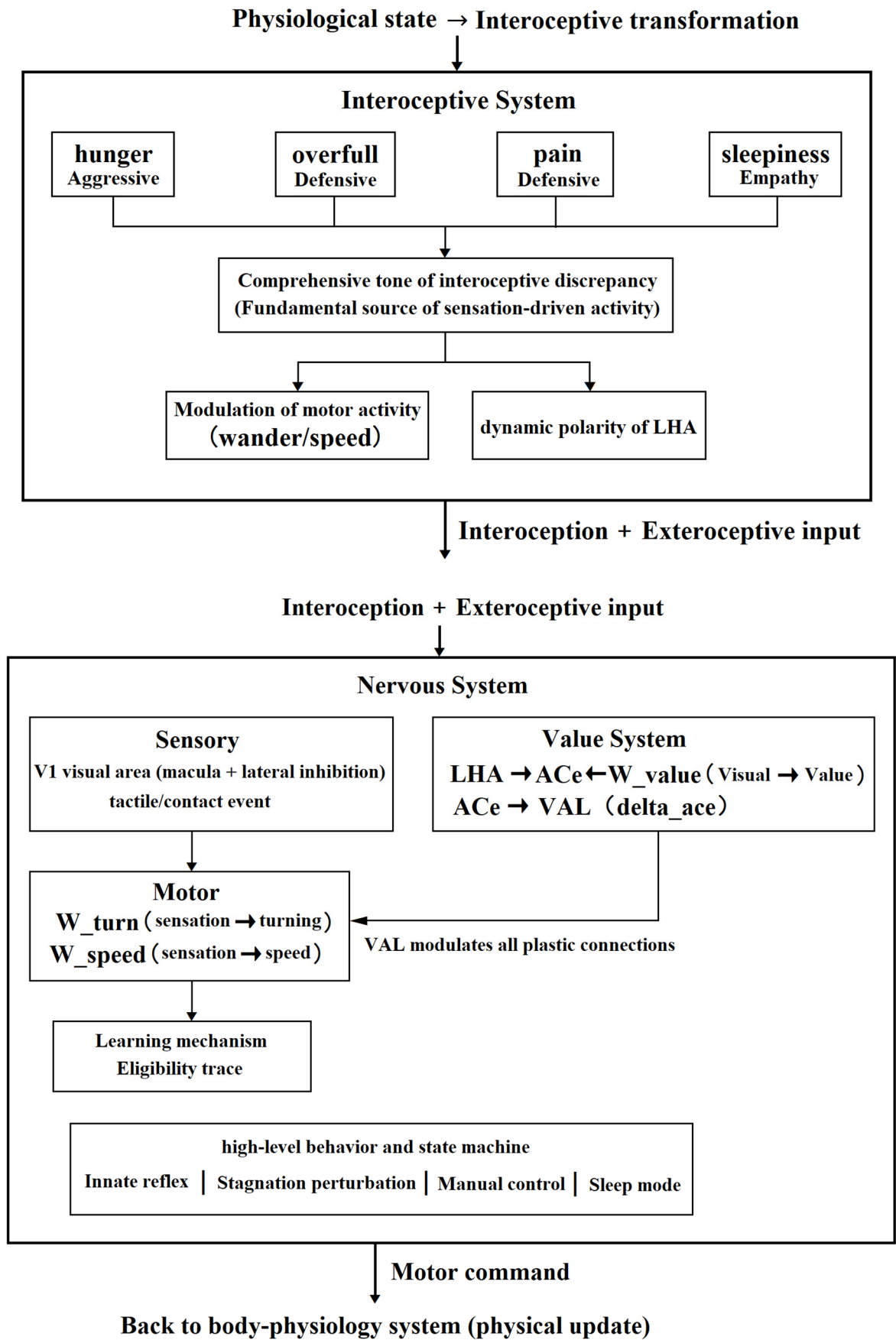


Figure 1. SBLB system architecture diagram

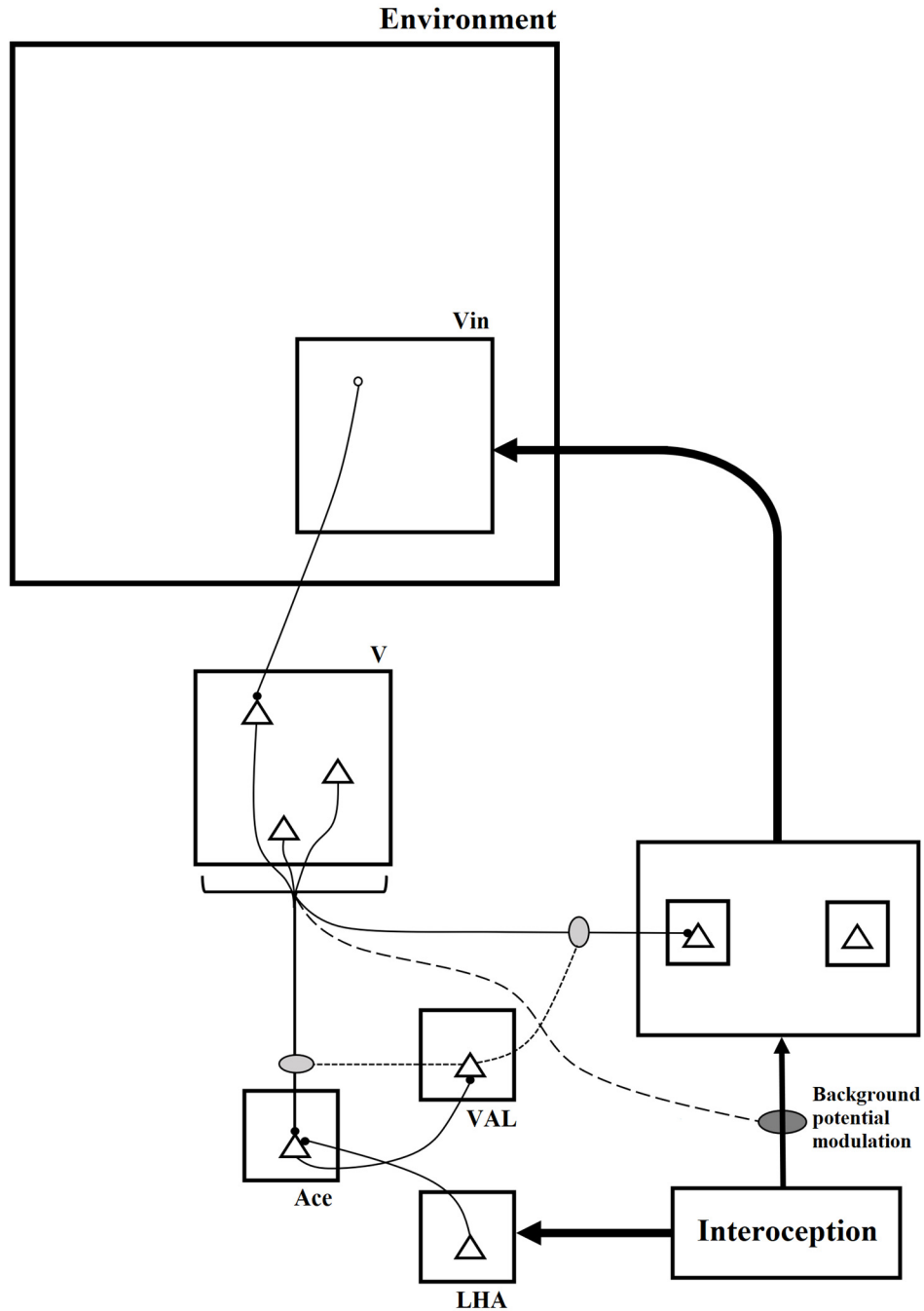


Figure 2. SBLB Schematic diagram of the nervous system simulation

No preset strategies. All plastic traces are initialized to zero; there is no preset object-value mapping table or behavioral strategy. “Green” initially triggers no value sensation or behavioral tendency. The emergence of foraging behavior is entirely the result of value flow driven by interoception and experience-based selection – that is, how **objective value sensation** (interoceptive valence directly triggered by contact events) gradually grows into **subjective value sensation** (the feeling carried by visual anticipatory value itself).

Intrinsic value drive. The only driver of learning and behavior is endogenous **sensation discrepancy** (Δ_{ace}). The positive LHA pulse generated upon contact with food is the **objective value sensation** directly evoked by the food contact event – a direct manifestation of interoceptive valence. The negative LHA pulse generated upon contact with a dangerous object is a direct expression of the pain caused by tissue injury. The system learns only those events that truly change its internal feeling state, not externally defined “rewards” or “punishments”.

1.1 Sensory System

Exteroception is an $11 \times 11 \times 3$ visual grid. The three channels encode food (green, ch0), danger (red, ch1), and obstacles/boundaries (cyan, ch2). The visual field has an angle of 100° and a distance of 80 units. Raw visual signals pass through two layers of fixed preprocessing: macular weighting (central enhancement, peripheral attenuation) and lateral inhibition (contrast enhancement). All parameters are non-learnable – they are evolutionarily preset **sensory coding motifs**.

Interoception is the core that implements **innate value system sensation**. The system defines four interoceptions: hunger and overfullness (aggressive/energy balance), pain (defensive/boundary integrity), and sleepiness (empathic/rhythm continuity). The intensity of interoception is non-linearly related to physiological quantities – a power law makes interoception accelerate near extreme states. The four interoceptions are computed independently, are non-learnable, and are **continuously varying state variables**: mild hunger is a weak negative value sensation, extreme hunger is an overwhelming negative value sensation. This continuity enables value flow.

1.2 Dual Pathways of the Value System: Objective Value Sensation and Subjective Value Sensation

SBLB’s value system distinguishes an **evaluative value sensation pathway** from an **anticipatory value sensation pathway**, directly corresponding to the two levels of “acquiring value” in the *Sentient Principle* – **objective value sensation** (the most primitive, irreducible value experience generated directly by physical contact) and **subjective value sensation** (the feeling carried by visual anticipatory value itself, which can self-maintain and deepen without objective contact).

Objective value sensation: the evaluative pathway – answering “Is this outcome good or bad?”

When the organism actually contacts an object or boundary, the system generates an instantaneous LHA pulse according to the contact type and interoceptive state. This is the most primitive, irreducible value experience:

- The LHA from food contact is dynamic: $lha_pulse = lha_food_base + hunger_boost - overfull_penalty - pain_suppress$. When overfull, it can turn from positive to negative – satiety is embedded directly within the value system, not as a strategy switch of an independent module.
- ACe (central nucleus of the amygdala) integrates three layers of value sensation: innate interoceptive tone, the value pulse from the contact event (objective value sensation), and acquired anticipatory value sensation (visual_value).
- Δace is the first difference of ACe and is the **only learning signal**. All trace stabilization and decay are modulated by it, embodying the core mechanism of **value-dependent selection**.

Subjective value sensation: the anticipatory pathway – answering “How much do I want / how much do I fear right now?”

Visual anticipatory value splits into two independent components (food and threat), which compete at the motor level. Interoception modulates the drives in real time:

- When hungry: amplifies the positive food value drive, suppresses the negative threat value drive (risk-taking tendency)
- When overfull: suppresses the positive food value drive, amplifies the negative threat value drive (conservative risk-avoidance)
- When in pain: amplifies the negative threat value drive (high-level risk-avoidance)

The final motor command is generated directly from the visual-motor traces and multiplied by the total modulation – **from visual input to motor output, no “decision module” is inserted. Sensation is the drive.**

Temporal separation of the two pathways (action before learning):

- Anticipatory value sensation (subjective value sensation) drives behavior (before contact)
- Behavior leads to contact
- Evaluative value sensation (objective value sensation) modulates traces (after contact)

The system acts first, then learns. This is the precise implementation of the **sensation-behavior loop** at the simulation level.

1.3 Learning System

SBLB 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, the eligibility trace is frozen, retaining the visual

memory before the reflex.

The attribution mechanism relies entirely on the spatio-temporal co-occurrence of eligibility traces – Δ_{ace} acts on all activated eligibility traces across channels, using a dynamic sign-filtering mechanism triggered by a confidence threshold. The precision of attribution is naturally guaranteed by spatio-temporal decay and position encoding.

Innate value baseline is embodied through the channel-specific polarity of the LHA, ensuring the asymmetric principle that “negative stabilization can be profound after a single occurrence, whereas positive stabilization usually requires many repetitions”. Trace change is determined by the product of Δ_{ace} and the eligibility trace – the system does not converge to an “optimal solution”, but instead **selectively stabilizes patterns marked by positive value sensation**. This is the underlying mechanism by which objective value sensation guides the growth of subjective value sensation.

1.4 Behavioral System and Rhythm System

Behavior follows a hierarchical motor priority: innate reflexes (contact bounce-off), boundary agitation, stagnation perturbation, manual control, sleep, spontaneous movement, and **acquired anticipatory sensation (subjective value sensation)**. This is not “selection” by a central decision-maker, but direct competition among different closed-loop tensions.

The three innate reflexes correspond directly to defensive tension. **Subjective value sensation → behavior** is a typical example of acquired anticipatory sensation directly driving behavior – as soon as a negative value anticipation emerges, avoidance behavior is initiated immediately. Spontaneous movement is the superposition of a wandering tension and visual learned traces, with interoception modulating the wandering vigor in real time.

The empathic tension is implemented through the sleep-wake rhythm. Sleepiness is driven by accumulated waking time; when it reaches a threshold, the organism falls asleep; when it decreases to a wake-up threshold, it naturally awakens. During sleep, metabolism drops to 20% of the waking level, and injury recovery doubles. This rhythm is entirely self-regulated by the accumulation and dissipation of interoception, requiring no external zeitgeber.

2. Experimental Design

2.1 Overall Experimental Design: A Five-Stage Environmental Sequence with Progressive Complexity

The experiment uses a **five-stage environmental sequence with progressive complexity**. The

organism faces ecological niches with different resource and threat densities across five consecutive developmental stages. Detailed data are shown in Table 1.

Initial environment is consistent with biological facts. Stage 1 (15 food items, 25 dangers) simulates the typical ecological niche of early development – parents tend to place eggs or offspring in areas with abundant food and low predator density. Higher mammalian juveniles usually grow in an “abundant-low-risk” environment under parental protection and guidance, allowing them to build basic foraging skills and value recognition (the **initial accumulation of objective value sensation**) with low survival risk. In this environment, the organism establishes basic positive food value recognition and achieves the initial embodiment of acquired value system sensation.

Table 1. Configuration parameters of the five SBLB developmental stages

Stage	Steps	Food	Danger	Obstacle	Inherited weights from
Stage 1	30k	15	25	0	zero initialization
Stage 2	100k	15	25	0	final weights of Stage 1
Stage 3	100k	10	25	0	final weights of Stage 2
Stage 4	95.4k	7	25	0	final weights of Stage 3
Stage 5	58.1k	7	25	0	final weights of Stage 4

Continuous inheritance of weights mimics 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 in which a single individual continuously accumulates sensation-behavior traces over its lifetime. At each stage, the organism 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** and to observe the complete process by which **objective value sensation gradually grows into autonomous subjective value sensation**.

2.2 Experimental Environment and Parameter Settings

The experimental environment is a square plane of 400×400 units. Object types include food (green, ch0), danger (red, ch1), and obstacles (cyan, ch2), with quantities at each stage as shown in Table 1. All objects are randomly generated, avoiding the central birth area and maintaining a minimum spacing of 20 units between objects. After being consumed, a food item is regenerated at a random position in the environment after 3 seconds.

Initial energy: 80 (maximum 100), initial injury: 0. Basal metabolic rate: 0.2/sec; movement consumption coefficient: 0.01; injury recovery rate: 0.2/sec. Maximum speed: 20.0 units/sec; maximum turning rate: 180°/sec. Ideal energy ratio (energy level / maximum energy) is set to 0.8; below this value, hunger arises; above 90%, overfullness arises. Maximum waking time: 120 seconds; sleepiness threshold for falling asleep: 0.9; waking threshold: 0.2. All parameters are kept constant across the five stages.

Learning parameters: learning rate for motor traces: 0.05; learning rate for value traces: 0.05; minimum learning threshold: 0.01; trace decay per frame: 0.999999; absolute upper limit of traces: 2.0.

2.3 Comparison Between Value-Dependent and Value-Independent Systems

Following the landmark 1994 paper *Value-Dependent Selection in the Brain* by the Edelman school, our experiment compares a **value-dependent system (intact V → Ace pathway)** and a **value-independent system (V→Ace pathway cut)**.

According to the *Sentient Principle*, this structural difference directly corresponds to two different **source structures of value system sensation**:

- **Value-dependent system:** visual anticipatory value (visual_value) can flow into ACe and thus become an input to the value system. This means that the anticipation itself acquires feeling – the system can, without objective contact, self-maintain and deepen its value judgments solely through its subjective value sensation field.
- **Value-independent system:** visual anticipatory value can only drive conditioned reflex behavior but cannot serve as an input to the value system to generate endogenous learning signals. Its value system sensation remains tightly coupled to objective contact events (LHA pulses) and is difficult to self-maintain in the absence of external evaluation.

We propose that value-dependent selection is the essential prerequisite for the emergence of cognition.

3. Results: Systematic Observation of Emergent Behavior

3.1 Value Grows from Sensation: Complete Evolution of Acquired Value System

Sensation

Figure 3 shows the complete evolution of the mean W_value across three channels over five stages. All plastic traces start from zero. In the early stage, the means of all three channels are near zero – the organism is “blind” to visual signals; at this point, only **objective value sensation** (LHA pulses directly evoked by contact events) plays a role in learning. In the middle stage, the food channel begins to rise

steadily, and the danger and obstacle channels start to deviate from zero – **subjective value sensation** begins to sprout from the soil of objective anchoring. In the later stage, the three channels are fully differentiated – the subjective value world has been established.

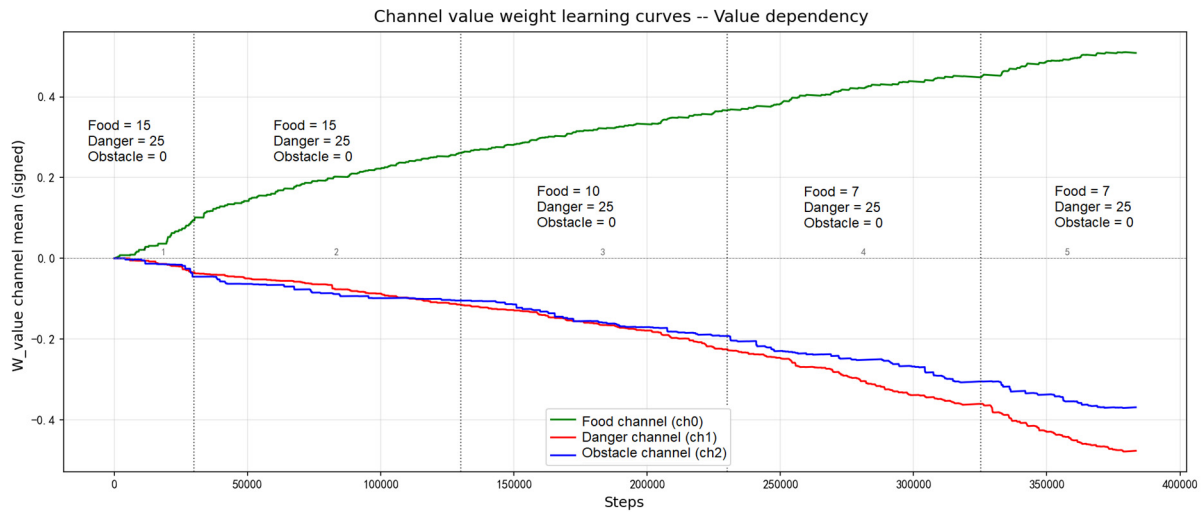


Figure 3. Evolution of mean W_value across three channels over five stages (380k steps)

Key insight: This evolutionary trajectory clearly shows the growth process from **objective anchoring to subjective dwelling** – the system is not pre-programmed with “green good, red bad”; rather, through value-dependent selection within its own sensation-behavior loop, it grows its own value feelings toward the world.

3.2 Sensation Directly Drives Behavior: Structural Directional Separation of the Three Channels

Figure 4(a) shows the column-summation heatmaps of W_turn , W_speed , and W_value in the first 10k steps of Stage 1, and Figure 4(b) shows the same at the end of Stage 5 (380k steps).

Food channel (ch0): left column negative (drives right turn), right column positive (drives left turn) – together forming an **approach turning pattern** that puts food straight ahead.

Danger channel (ch1): left column positive (drives left turn), right column negative (drives right turn) – together forming an **avoidance turning pattern** that moves danger out of the visual field.

Obstacle/boundary channel (ch2): same pattern as the danger channel, with slightly smaller magnitude.

Weight column sum distribution (step 430000) -- no_value_dependency -- left/right bias per channel

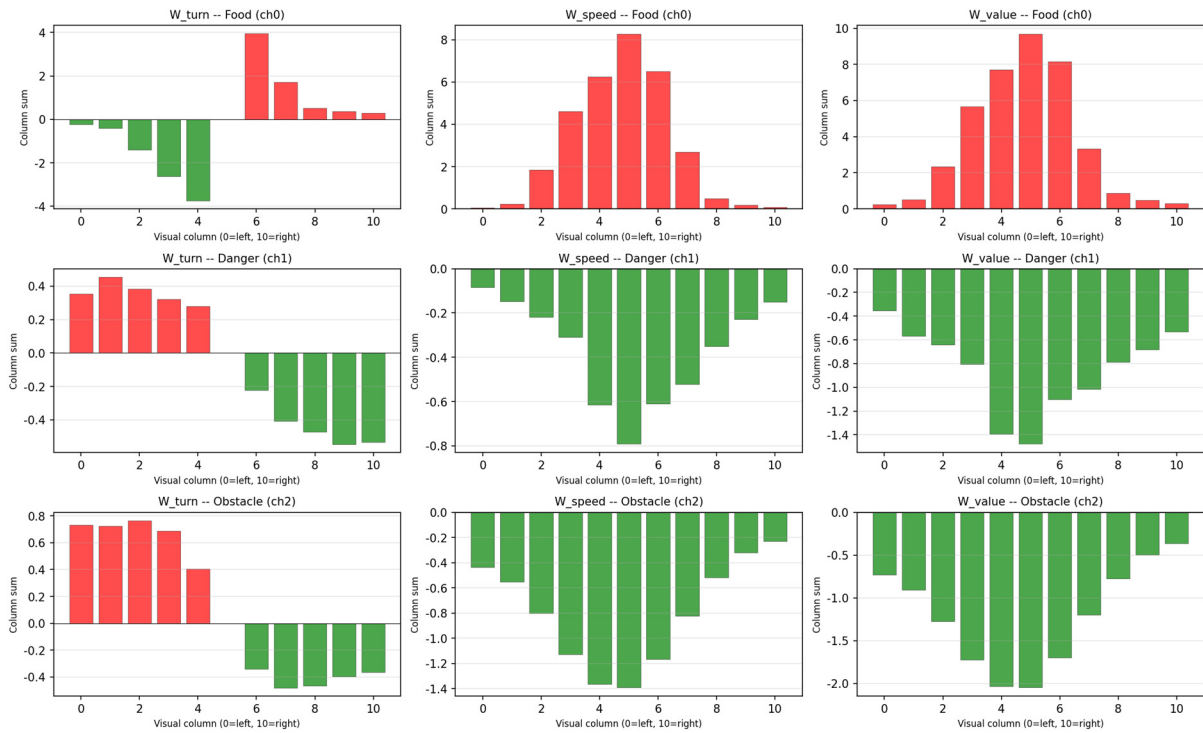


Figure 4a. Column-summation heatmaps of W_turn/W_speed/W_value in the first 10k steps of Stage 1

Weight column sum distribution (step 380000) -- Value dependency -- left/right bias per channel

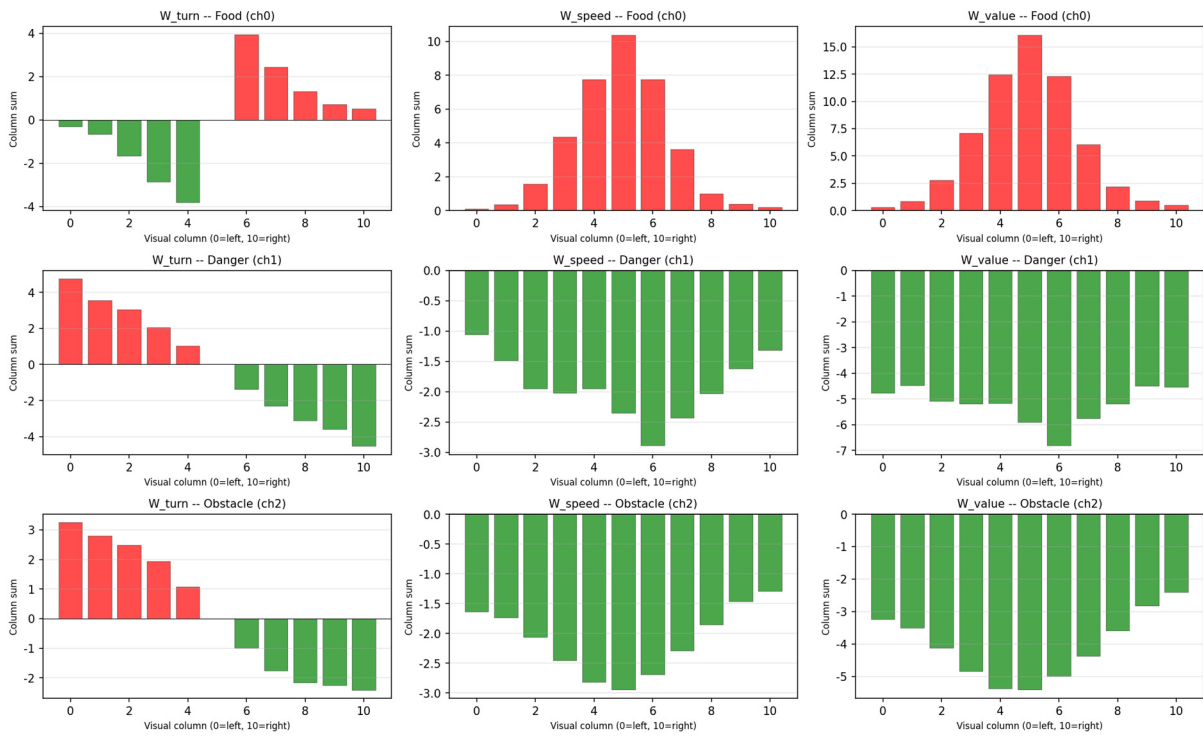


Figure 4b. Directional separation heatmaps of W_turn/W_speed/W_value at the end of Stage 5 (380k steps)

W_{speed} shows exactly the opposite pattern across the three channels: the central column for the food channel is positive (accelerating approach), while the central columns for the danger/obstacle channels are negative (decelerating or backing off). W_{value} also shows a pattern similar to that of W_{speed} .

Profound meaning of structural separation: positive value sensation (positive polarity of **subjective value sensation**) stabilizes approach traces; negative value sensation stabilizes avoidance traces. The positive/negative polarity of value sensation maps directly into the direction of behavioral differentiation – **visual input** → **value sensation modulation** → **motor output**, with no independent “decision” step inserted throughout. This is structural evidence at the neural-trace level that “sensation directly drives behavior”.

Table 2 gives the ratios between the peak absolute column-sum values at the final state and those at the initial state, showing the extent to which subjective value sensation is strengthened through experience.

Table 2. Comparison of peak absolute column-sum values of $W_{\text{turn}}/W_{\text{speed}}/W_{\text{value}}$ between initial and final states

Channel	Weight	Max column-sum absolute value		Ratio (380k/10k)
		10k steps	380k steps	
Food (ch0)	W_{turn}	0.123	3.932	31.97
	W_{speed}	0.732	10.372	14.17
	W_{value}	0.887	16.076	18.12
Danger (ch1)	W_{turn}	0.053	4.765	89.91
	W_{speed}	0.198	2.892	14.61
	W_{value}	0.240	6.809	28.37
Obstacle (ch2)	W_{turn}	0.011	3.246	295.09
	W_{speed}	0.052	2.942	56.58
	W_{value}	0.059	5.405	91.61

3.3 Comparison Between Value-Dependent and Value-Independent Systems: The Cost of Subjective Dwelling and Adaptiveness

Our experiment not only replicates the core findings of the Edelman school but also yields new important insights.

Replicated finding: The value-dependent system learns significantly more efficiently. As shown in

Table 3, the absolute values of the weights of all three channels are higher in the value-dependent system than in the value-independent system; in particular, the negative weight for the danger channel (0.477 vs 0.077) shows a ratio as high as 6.19. This is consistent with the Edelman school’s conclusion that “after about 1000 seconds of experience, all responses become conditioned responses, i.e., triggered by vision alone”.

New finding: The cost of subjective dwelling

When the environment becomes food-scarce (Stages 4 and 5: 7 food items, 25 dangers), a striking reversal occurs:

- The **value-dependent system** (which possesses autonomous subjective value sensation) frequently starves to death after about 50,000 steps.
- The **value-independent system** (whose value sensation remains tightly anchored to objective contact) consistently completes 100,000 steps and maintains good energy levels.

Table 3. Comparison of absolute mean W_value of the three channels between value-dependent and value-independent systems

Three-channel W_value	Value-dependent	Value-independent	Ratio (VD/VI)
Food channel ch0	0.509	0.324	1.57
Danger channel ch1	0.477	0.077	6.19
Foraging/Avoidance ratio	1.067	4.208	0.254

Figure 5 shows the evolution of mean W_value across three channels for the value-independent system over five stages (430k steps); Figure 6 shows the spatial distribution of weights for the two systems. Table 4 gives the statistics of four consecutive runs.

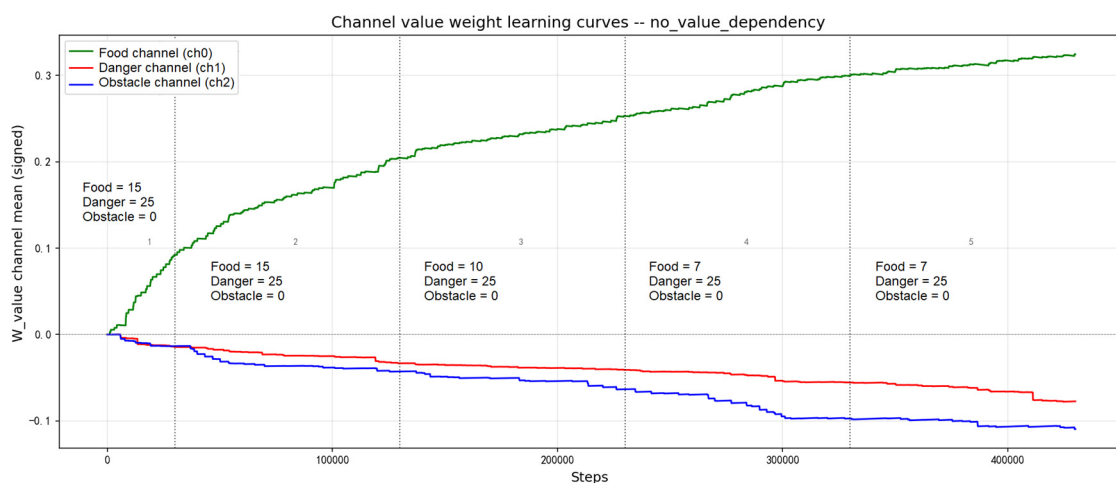
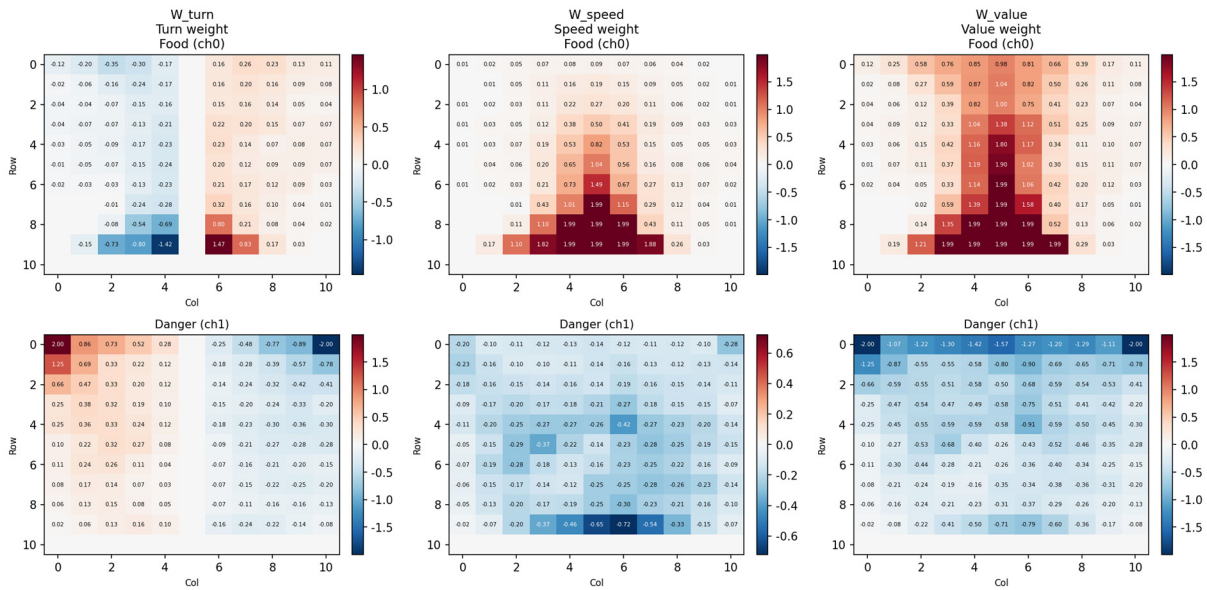


Figure 5. Evolution of mean W_value across three channels for the value-independent system over five stages (430k steps)

Table 4. Statistics of four consecutive runs of value-dependent vs value-independent systems (food 7, danger 25)

	Value-dependent				Value-independent			
	Run 4	Run 5	Run 6	Run 7	Run 4	Run 5	Run 6	Run 7
Total steps	95.4k	58.1k	58.7k	55.8k	100k	100k	100k	100k
Food consumed	54	34	31	31	67	66	61	63
Ten thousandth	5.66	5.85	5.28	5.56	6.70	6.60	6.10	6.30

Weight spatial distribution (step 380000) -- Value dependency
(Rows: grid row [0=far / 10=near], Cols: grid col [0=left / 10=right])



Weight spatial distribution (step 430000) -- no_value_dependency
(Rows: grid row [0=far / 10=near], Cols: grid col [0=left / 10=right])

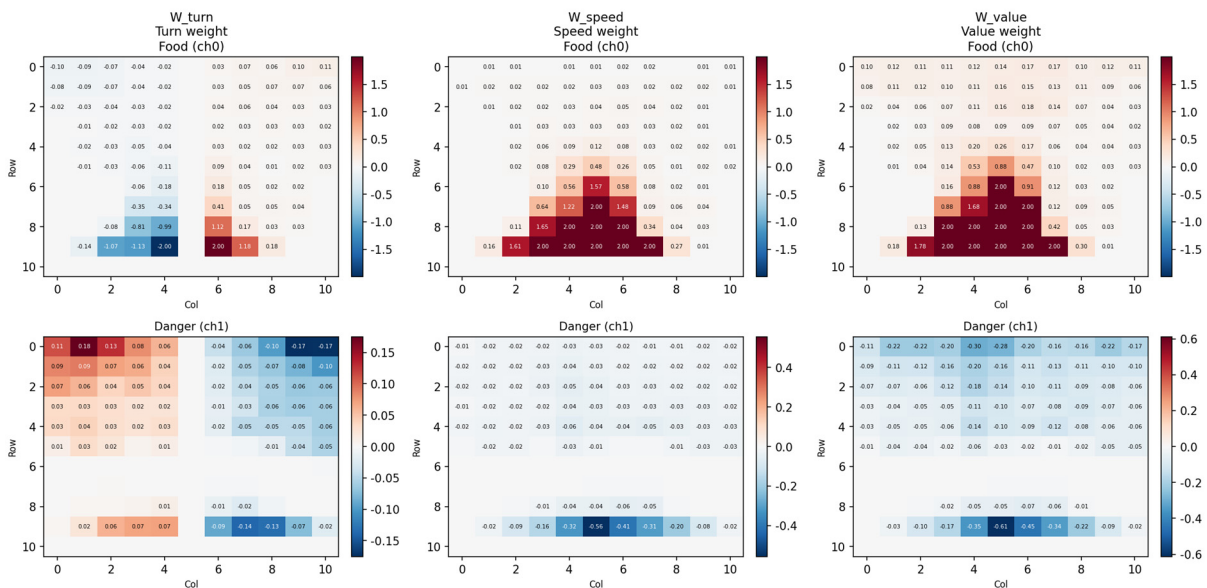


Figure 6. Comparison of weight spatial distributions between value-dependent and value-independent systems

Theoretical interpretation:

The greatest advantage of the value-dependent system – **the ability to sense threats from a distance** (i.e., the strong intensity of subjective value sensation) – turns into a disadvantage in a resource-sparse environment. The system avoids not only dangers but also loses opportunities to consume food located near boundaries or behind dangerous objects, because its internal threat anticipation (subjective negative value sensation) dominates the current feeling field and suppresses the foraging drive.

In contrast, the **objective value sensation** of the value-independent system remains tightly coupled to each contact event: it has no strong endogenous threat anticipation, its behavior is dominated by proximal responses, and it continues to approach and consume food even in ambiguous or risky situations.

Core conclusion: Predictive capacity enhances safety, but in sparse or patchy environments it may reduce adaptive flexibility. The “dwelling” of subjective value sensation endows the system with autonomy from immediate evaluation, but this autonomy comes with a cost – strong internal feelings may obscure the grasp of immediate opportunities.

This phenomenon bears a profound similarity to certain real-world behavioral decision patterns: people with high sensitivity to danger and strong foresight often lose more opportunities to gain resources, whereas those less sensitive to danger but more sensitive to reward may actually obtain more resources. This is a direct simulation-level confirmation of the “advantages and costs of cognition” revealed in the paper *From Objective Anchoring to Subjective Dwelling*.

4. Conclusion: SBLB as an Empirical Bridge for the *Sentient Principle*

The SBLB simulation system validates the core propositions of the *Sentient Principle*:

1. **Value grows from sensation:** All plastic traces start from zero, gradually differentiating into positive food value and negative danger value through value-dependent selection. **Objective value sensation** (LHA pulses from contact events) is the soil of growth; **subjective value sensation** (the feeling carried by visual anticipatory value itself) is the autonomous value world that grows from that soil.

2. **Sensation directly drives behavior:** The structural directional separation of the W_{turn} , W_{speed} , and W_{value} channels demonstrates that there is no independent “decision module” – the polarity of value sensation maps directly into the direction of behavior. This is neural-trace evidence that “sensation is the drive”.

3. **Temporal separation of the dual value pathways:** The anticipatory pathway drives action (subjective value sensation), while the evaluative pathway modulates learning (objective value sensation)

– the system acts first, then learns. This precisely implements the closed-loop dynamics of the **sensation-behavior loop** as described in the *Sentient Principle*.

4. **The cost of subjective dwelling:** The vulnerability of the value-dependent system in resource-scarce environments reveals the double-edged nature of cognitive autonomy – exactly the argument made in *From Objective Anchoring to Subjective Dwelling* that “the autonomization of the subjective value world is not an unconditional advantage, but a mode of dwelling in a particular ecological niche”.

5. **Theoretical significance:** SBLB provides a computable, testable empirical platform for the *Sentient Principle*, enabling abstract philosophical and biological propositions to be grounded in simulation. It builds a bridge – a bridge back to life itself – from Whytt’s *An Essay on the Vital and Other Involuntary Motions of Animals* (1751), Sechenov’s systemic feelings, and Edelman’s value-dependent selection, to contemporary computational neuroscience and artificial intelligence.

References

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

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.