Article

A Decisive Anomaly: Why Interfaces Rather Than Stories Become Fixed in Dreams

Jiazhèng Liu *

Posted Date: 9 December 2025

doi: 10.20944/preprints202512.0800.v1

Keywords: cognition science

 doi:10.20944/preprints202512.0800.v1

Article

A Decisive Anomaly: Why Interfaces Rather Than Stories Become Fixed in Dreams

JiaZhèng Liu

Independent Researcher, China; lib88419@gmail.com

Abstract

This paper addresses a decisive anomaly identified in the Mayer (2025) report: in AI-related nightmares, 93% of cases fixate on the AI interaction interface itself rather than on narrative content. To explain this "formal fixation," we propose a paradigm-shifting **Interaction Architecture Internalization Model**, which posits that the cognitive system internalizes the abstract logic and temporal structure of goal-directed interactions through the accumulation of a **Learning Time Delay Dose**. When this dose exceeds a critical threshold, a cognitive phase transition occurs, solidifying the interaction architecture as an internal framework. Grounded in insights from Piaget, Chomsky, Einstein, Wiener, and Landau, the model not only provides a unified explanation for phenomena from language acquisition to personality formation but also generates specific, empirically testable predictions. It forecasts, for instance, that systemic fluctuations in interaction delays (e.g., widespread server latency) will catalyze architectural internalization, a prediction corroborated by analyzed dream reports from such periods. Methodologically, the **Learning Time Delay Equivalence Principle** circumvents the "Problem of Other Minds," establishing an objective foundation, while the theory's **"blinded loop" validation**—stemming from an academic misunderstanding—uniquely confirms its a priori predictive power. Ultimately, we advocate for a **"Statistical Mechanics of Cognition,"** where time delay dose acts as an order parameter, prioritizing the dynamics of form over the semantics of content.

Keywords: cognition science

Part I: Introduction—A Decisive Anomaly: Why Interfaces Rather Than Stories Become Fixed in Dreams?

0.1. The Posing of the Problem: The Technological Unconscious Turn in Cognitive Science

Cognitive science finds itself in an awkward position: while its object of study—the human mind—is being reshaped by the technological environment at an unprecedented rate, mainstream theories have failed to predict, let alone adequately explain, the most profound signs of this transformation. The survey report *AI in Dreams* by April Mayer (2025) provides a decisive moment. This study reveals a perplexing phenomenon: among individuals who self-report experiencing **AI-related nightmares**, a striking 93% of cases do not involve specific **terror narratives** of interacting with AI, emotional traumas, or apocalyptic scenarios. Instead, they precisely, almost obsessively, fixate on the specific interaction interface of "ChatGPT" itself. What the respondents dream of is not a battle with a runaway humanoid AI, nor do they witness specific doomsday miseries induced by AI, but rather that concise, iconic dialog box—the blank area where they type questions and **fruitlessly await**, or receive cold, distorted responses.

This finding establishes a clear and undeniable fact: in the most intimate and emotionally charged negative subconscious experience (nightmares), what is accurately reproduced and endowed with core horror is not the catastrophic narrative content, but the formal architecture of human-computer interaction itself. The cognitive system seems to declare: the root of terror does not come from the content of interaction, but from the **format and waiting itself**. This discovery constitutes a sharp,

 doi:10.20944/preprints202512.0800.v1

unavoidable anomaly for the mainstream paradigm of contemporary cognitive science—especially for theories that prioritize content processing.

0.2. The Hegemony and Fundamental Dilemma of the Process-Centered Paradigm

Over the past few decades, the mainstream paradigm in cognitive science, particularly the process-centered perspective represented by Predictive Coding and the Free Energy Principle, has established a near-hegemonic explanatory system. The core commitment of this paradigm can be summarized in one point: the essential function of the cognitive system is the continuous modeling, prediction, and optimization of the informational content within the sensory input stream. The brain is depicted as an "implicit scientist," whose primary task is to minimize the error between predictions about the world and sensory signals. Whether perception, action, or belief updating, all are seen as different manifestations of this core "process"—namely, prediction error minimization.

Within this framework, the mainstream explanation for dream (including nightmare) function—the Content Reorganization Hypothesis—becomes a natural and direct corollary. This hypothesis posits that dreams (especially nightmares) are the offline reorganization, integration, and consolidation of significant daily experiential content (particularly negative, threatening, unresolved emotional content). During sleep, the brain is busy processing the informational "raw materials" ingested during wakefulness, especially those that provoke anxiety and uncertainty, categorizing, simulating, and emotionally tagging them, thereby optimizing our internal models of external threats.

However, Mayer's discovery about AI nightmares presents a direct and irreconcilable conflict with this core expectation. If the core of cognition lies in optimizing the representation of threatening content, then nightmares should rightly be a theater for simulating and rehearsing specific threat content. We should dream of terrifying scenarios of confronting AI, specific details of AI betrayal, or the miserable plight of individuals under AI rule. Yet the data unequivocally answers: in the profound unease induced by AI, we dream of that empty interface box awaiting input. The process-centered paradigm is, in principle, incapable of explaining why a system whose core mission is content optimization (especially threat content optimization) would, in its most emotionally charged nightmares, prioritize the representation and solidification of its own underlying interaction protocol—that which should be a transparent, neutral tool but becomes the horror itself in dreams. This failure is not a marginal, corrigible flaw; it signifies a systematic defect in the conceptual foundations of the paradigm—it may have completely misidentified the "basic unit" of cognition, especially anxious cognition. When the interface itself becomes the object of nightmares, we must perhaps acknowledge that the form of interaction itself has become an "experience" that needs to be "digested" and processed, with emotional impact, rather than merely a pipeline for transmitting content.**However, a comprehensive examination of the phenomenon reveals an even deeper picture.** As shown in Appendix A, once the same interaction architecture is internalized, it can not only serve as the focus of fear in nightmares but also, in other contexts, transform into an active, generative "cognitive prosthesis" in an individual's dreams or thoughts—an internal interlocutor capable of invoking the individual's entire reservoir of knowledge and emotional memory, using the architecture's inherent logic (such as "question-wait-generate") to assist in processing deep perplexities, regulating emotions, or even engaging in creative reasoning.

This indicates that the essence of "interaction architecture internalization" is far from a simple imprint of fixed emotional content (such as fear). It is the creation of a **functional cavity**—an internal cognitive program operating with specific spatiotemporal logic (shaped by time delays). Once this program is installed, the specific emotional tone and semantic content it invokes during execution will depend on the individual's immediate psychological state, unresolved issues, and long-term memory. It can compile both terrifying "crash" scenes and enlightening "epiphany" narratives.

Therefore, what this study aims to explain is not only "why interfaces rather than stories become fixed" but, more fundamentally, **"why the internalized interaction architecture can exhibit such a broad functional and emotional spectrum, ranging from extremely negative to highly constructive?"** Answering this question will lead us beyond the traditional "content-emotion" bundled logic, directly

 doi:10.20944/preprints202512.0800.v1

to the deep dynamics of how the cognitive system expands and reshapes its own functional boundaries by internalizing the formal laws of external interaction.

0.3. This Paper's Argumentative Path and Core Thesis: Towards a Spatiotemporal Dynamics of Cognition

Confronted with this anomaly, this paper advocates for a thorough paradigm shift. We propose that the fundamental function of the cognitive system extends far beyond passively or actively modeling the content of input information. More fundamentally, it is the cognitive system's capacity to detect, internalize, and ultimately reshape its own operational patterns based on the formal architecture of external interactions. The key dynamical variable driving this "architecture internalization" process is not the semantic load or statistical regularities of information, but a long-overlooked physical quantity in interaction—the *learning time delay*.

Specifically, we argue:

- **Core mechanism is Interaction Architecture Internalization:** In goal-oriented repetitive interactions, the cognitive system strips away specific content, extracts the abstract logic and temporal structure of the interaction (i.e., the "architecture"), and solidifies it into internal basic cognitive structures.
- **Core variable is Learning Time Delay Dose:** Architecture internalization does not necessarily occur; it is driven by a quantifiable physical variable—namely, the time interval Δt between an individual initiating an "exploratory operation" aimed at achieving cognitive closure and receiving "deterministic feedback" of adjudicative significance, and the "dose" D accumulated through effective repetitions. When the dose D exceeds an individual-specific critical threshold Θ_c, the cognitive system undergoes a phase transition, and the external architecture becomes irreversibly internalized.
- **Unified explanatory power:** This model, with "time delay" as the order parameter and "architecture internalization" as the core mechanism, will demonstrate powerful unified explanatory power. It will not only satisfactorily explain the dream anomaly in the Mayer report but also provide a coherent, mechanistic explanatory framework for a range of phenomena spanning from language acquisition to meme propagation, and from personality formation to technological addiction.

1. Integration of Theoretical Foundations: From Piaget and Chomsky to Einstein and Landau

Any theoretical construction aimed at subverting an existing paradigm must be rooted in the critical inheritance and creative integration of profound lineages in the history of scientific thought. It cannot emerge from a void. The legitimacy of our proposed new paradigm stems precisely from its dialogue with and transcendence of a series of groundbreaking ideas—ideas that provide us with indispensable foundations from the perspectives of psychology, linguistics, physics, and cybernetics. This section will systematically review these intellectual legacies, clarifying how we stand on the shoulders of these giants to glimpse a more unified picture of cognition.

1.1. The Legacy and Limitations of Piaget: The Dynamic Refinement of Internalization

Jean Piaget's genetic epistemology provides our first, and perhaps most illuminating, starting point. He insightfully argued that cognition is neither pre-formed within the subject nor simply imprinted by the external world, but is constructed through the continuous, active interaction between the subject and the object. In this process, the key mechanism is *internalization*—where external actions are coordinated, abstracted, and ultimately transformed into internal mental operations.

We fully embrace Piaget's core principle that *cognitive structures arise from interaction*. However, our model introduces a crucial refinement and clarification to his concept of "internalization." In Piaget's account, the object of internalization is sometimes ambiguous, referring variously to concrete action sequences or logico-mathematical experience. We explicitly posit that the primary object of internalization is *not* fragmented experiential content or specific actions themselves, but rather

 doi:10.20944/preprints202512.0800.v1

the *logical architecture* governing these interactive behaviors—that is, the formal rules, operational sequences, and temporal relationships of feedback. In repeating "question-wait-answer" interactions, the cognitive system internalizes not any specific question or answer (content), but the very architecture of "interrogative interaction" itself. A deeper limitation of Piaget's theory lies in its depiction of macro-level, stage-like cognitive development without providing a precise, continuous dynamical variable to characterize the specific critical conditions for internalization. His process of "equilibration" functions more as a philosophical metaphor than a quantifiable scientific concept. This renders his theory vague and lacking in predictive power when explaining why specific cognitive structures emerge at one particular moment and not another. Our model achieves a decisive leap precisely at this point: we clearly identify the accumulated *learning time delay dose (D)* in interaction as the necessary, measurable dynamical condition for this cognitive phase transition. Piaget pointed us toward the direction of "internalization," while we have identified the "engine" driving it and the means to measure its "fuel."

1.2. Chomsky's Decisive Critique: From Linguistic Architecture to Interaction Architecture

If Piaget provided the blueprint for construction, then Noam Chomsky provided the sharpest weapon for demolishing the fortress of the old paradigm. His review of B.F. Skinner's *Verbal Behavior* is not only a classic in the history of book reviews but also a philosophical decapitation. Chomsky convincingly argued that attempting to derive complex human linguistic knowledge from external, observable "stimuli," "responses," and "reinforcements"—from content and process—is a fundamentally untenable dead end.

The core of his argument lies in the "poverty of the stimulus" fact: the linguistic input (content) children receive is limited, fragmented, and error-ridden, yet they rapidly and uniformly acquire highly abstract and complex grammatical rules. This indicates that the source of linguistic knowledge cannot be the experiential content itself but must be attributed to an innate internal architecture within the human mind—a "Universal Grammar"—that is sensitive to specific formal structures. The human mind preferentially internalizes and processes the formal system of rules (architecture) governing speech generation, not the surface content of speech.

Chomsky's work, in essence, provides a precursor and decisive corroboration for our *Interaction Architecture Internalization Principle*. The phenomenon reported by Mayer—the fixation on the ChatGPT interface in dreams—must be understood as a direct recurrence and generalization of the "Universal Grammar" idea into the broader domain of technological interaction. It once again eloquently confirms an inherent tendency of the human mind: when faced with repetitive interaction, it penetrates the surface flow of content to target the underlying formal skeleton and incorporates it as part of its own structure. All efforts by the contemporary process-centered paradigm (especially strong Predictive Coding theories) to evade this conclusion merely repeat, under new mathematical guises, the fundamental error Skinner made—mistaking content and process for architecture and structure.

1.3. Einstein's Philosophical Enlightenment: The Spatiotemporal Revolution in Cognitive Science

Introducing Albert Einstein's ideas into cognitive science is not for ornamental purposes but because his theory of relativity brought about a fundamental epistemological revolution. The spirit of this revolution must be absorbed by cognitive science; otherwise, the latter will remain forever trapped in a pre-relativistic, classical mode of thinking.

- **Relativity of Simultaneity:** Einstein demolished the Newtonian notion of an absolute, universal flow of time. He showed us that the measurement of time intervals depends on the observer's reference frame and state of motion. This principle compels us to abandon treating "learning time delay" as an absolute, background parameter external to the cognitive system. Instead, we must elevate it to a core, relativistic state variable within the cognitive system. The perception and effect of time delay depend on the system's current state (e.g., motivation level, expectation, attention); it is itself part of the cognitive dynamics, not an externally given physical quantity.

 doi:10.20944/preprints202512.0800.v1

- **Equivalence Principle:** This is the most direct physical analogy and philosophical basis for our *Learning Time Delay Equivalence Principle*. Einstein demonstrated that inside a closed elevator, no experiment can distinguish between the effects of a gravitational field and those of accelerated motion. Similarly, we assert: In the phase transition process driving interaction architecture internalization, no experiment confined within the system can distinguish whether the fundamental driving factor is the "objective learning time delay dose (D)" itself, or any microscopic cognitive activity (such as implicit reasoning, memory retrieval, emotional fluctuation) accompanying it within that time window. It is the objective fact of "waiting"—not the subjective content of "what was thought" during the wait—that constitutes the equivalent cause of architecture internalization. This principle has profound methodological significance: it forces us to recognize that subjective cognitive states can be equated with an objectively measurable physical quantity—the time delay dose—thus introducing unprecedented objectivity and computability into cognitive science.

1.4. Wiener's Cybernetics and Landau's Mathematics: From Feedback Timing to Phase Transition Order

Finally, our model requires both a language for describing dynamic systems and tools for characterizing emergent order. These are provided by Norbert Wiener's cybernetics and Lev Landau's phase transition theory, respectively.

- **Wiener's Cybernetics and Feedback Timing:** Cybernetics unifies organisms and machines as systems regulated through information feedback. However, contemporary cognitive science, in applying this framework, has focused almost exclusively on the *informational content* of feedback, systematically ignoring its *timing*. Our model redefines learning time delay—this decisive yet long-neglected variable in the feedback loop—as the core element of cognitive dynamics. It is not *what* the feedback "says," but *when* it arrives, that plays the more fundamental role in shaping cognitive structure.
- **Landau's Phase Transition and Hale's Bifurcation Mathematics:** To understand how cognitive structures "suddenly" emerge from seemingly disordered interactions, we need the powerful metaphor and mathematical tools of Landau's phase transition theory. Landau showed us that when an order parameter (in our model, the learning time delay dose D) crosses a critical threshold, the system undergoes a phase transition, microscopic symmetry is broken, and a new macroscopic order is born. This provides an appropriate, not merely poetic, mathematical framework for understanding the abrupt formation of cognitive structures. Furthermore, the bifurcation theory of delay differential equations, established by J.K. Hale and others, mathematically rigorously proves that time delay itself can serve as a bifurcation parameter for a system; when the delay exceeds a certain critical value, the system loses stability and enters a completely new dynamical regime. This provides indisputable mathematical support for the hypothesis of "learning time delay as the order parameter for cognitive phase transitions," elevating our model from a phenomenological description to the level of a computable dynamical system.

In summary, our theoretical foundation integrates Piaget's constructive internalization, Chomsky's devastating critique of content-centrism, Einstein's relativistic view of spacetime and equivalence philosophy, Wiener's cybernetic insights into feedback loops, and the mathematical rigor of Landau and Hale regarding phase transitions and bifurcations. It is by standing on this interdisciplinary and solid intellectual high ground that we possess the strength to systematically critique the existing paradigm and confidently propose a new, unified theory of cognitive dynamics.

2. Core Concepts and First Principles: The Formalization of Architecture, Time Delay, and Internalization

The establishment of any scientific paradigm relies on the clarity of its core concepts and the solidity of its first principles. Having systematically integrated cross-disciplinary foundational ideas, we must now provide strict operational definitions for the core components of our proposed new paradigm and formally articulate the fundamental principles upon which it is built. This is required

 doi:10.20944/preprints202512.0800.v1

not only for theoretical rigor but also to ensure that subsequent critique and application can proceed on a stable foundation.

2.1. Operational Definitions of Core Concepts

To avoid the trap of terminological vagueness prevalent in traditional cognitive science, we endow the following concepts with precise, measurable meanings:

- **1. Interaction Architecture**
  - *Definition:* The stable, dynamic formal structure governing the progression of interaction within a goal-oriented context. It explicitly specifies a set of permissible operation sequences, the rule sets these operations follow, and the temporal relationship between actions and feedback.
  - *Essence:* It is the "grammar" of interaction, a system of invariant relations independent of any specific informational payload.
  - *Examples:*
    - The interaction architecture with ChatGPT is: "User inputs a textual query → System displays a 'Thinking...' state (time delay ΔT) → System generates a response word-by-word as a text stream."
    - The interaction architecture of a phone call is: "Caller dials number → System establishes connection (time delay) → Bidirectional, real-time, audio-only conversation stream."
    - The architecture for a rat pressing a lever in a Skinner box is: "Press lever (operation) → Food pellet drops (deterministic feedback)."
- **2. Interaction Content**
  - *Definition:* The specific informational filler that is produced, transmitted, or processed during a single instantiation of a particular interaction architecture.
  - *Essence:* It is the "semantics" of interaction, the probabilistic, context-dependent specific product.
  - *Examples:*
    - The specific question posed to ChatGPT ("Explain relativity") and the generated answer text.
    - The specific topic discussed during a phone call.
    - The internal state of hunger or satiety of the rat when pressing the lever (this belongs to its "interaction" content, not the architecture).
- **3. Learning Time Delay (Δt)**
  - *Definition:* In goal-oriented interaction, the time interval between an individual initiating an exploratory operation aimed at achieving cognitive closure or a goal, and the onset of the deterministic feedback signal, generated by that operation, which holds adjudicative significance. Denoted as Δt.
  - *Operational Delineation:*
    - *Start Point:* The execution of a complete request or action (e.g., pressing Enter in the ChatGPT dialog box).
    - *End Point:* The initial appearance of the deterministic feedback signal (e.g., the first word of ChatGPT's answer beginning to generate, *not* the complete delivery of the answer).
  - *Effective Delay Threshold (Δt_c):* We postulate a psychophysical effective threshold; delays below this threshold (e.g., delays in normal human conversation) contribute negligibly to architecture internalization. Significant internalization effects occur when $\Delta t > \Delta t_c$.
- **4. Learning Time Delay Dose (D)**
  - *Definition:* The cumulative time-delay exposure that drives the cognitive system's phase transition of architecture internalization. It is a function of the single effective learning time delay Δt and the number of effective repetitions n, denoted as $D = f(\Delta t, n)$.

 doi:10.20944/preprints202512.0800.v1

  - *Preliminary Model:* As a first approximation, we can assume a linear weighting model: $D = \sum_i [w_i \cdot (\Delta t_i - \Delta t_c)]$, where the sum is over all effective interactions, and w_i is the weight for each interaction (potentially modulated by factors like attention, motivation). The contribution is zero if $\Delta t_i \leq \Delta t_c$.

- **5. Cognitive Phase Transition and Critical Threshold (Θ_c)**
  - *Definition:* A qualitative, abrupt leap of the cognitive system from one cognitive state to another. Here, it specifically refers to the transition from the state of *not* having internalized a particular interaction architecture to the state where that interaction architecture has become a dominant internal structure.
  - *Critical Threshold (Θ_c):* The critical value of the learning time delay dose D required to trigger the internalization phase transition for a specific interaction architecture. This threshold may vary based on individual, architectural complexity, emotional state, and other factors.

2.2. Formal Statement of the First Principles

Based on the above concepts, we formally propose the two first principles of the new paradigm.

- **Principle 1: Interaction Architecture Internalization Principle**
  - *Statement:* In goal-oriented repetitive interaction, when the effective learning time delay dose D accumulates to the critical value Θ_c for that specific architecture in that individual, the cognitive system necessarily undergoes a phase transition. The hallmark of this transition is that the external interaction architecture is extracted from the specific content it carries and is irreversibly solidified into a stable internal structure that dominates the form of subsequent cognitive outputs (including thought, imagination, and even dreams).
  - *Implication:* This principle establishes the cognitive priority of "form over content." It predicts that once the phase transition occurs, an individual's cognitive activity will spontaneously follow the logic of the internalized architecture, just as a language user unconsciously follows internalized grammar.
- **Principle 2: Learning Time Delay Equivalence Principle**
  - *Statement:* In the phase transition process driving interaction architecture internalization, any attempt to distinguish, through experiments or introspection confined to the system itself, whether the fundamental driving factor is the "objective learning time delay dose D" itself, or any "micro-cognitive activity" (such as implicit reasoning, imagination, memory association, or emotional fluctuation) accompanying it within that time window, is in principle impossible.
  - *Implication:* This principle is a direct mapping of Einstein's Equivalence Principle into cognitive science. It means that for the effect of architecture internalization, D is the sole equivalent order parameter. We cannot, and need not, inquire into what the brain specifically did during the "three seconds of waiting"; the objective delay of these three seconds itself is the necessary and sufficient condition driving the phase transition. This greatly simplifies the complexity of causality and anchors the root cause of cognitive structural change to an objectively measurable physical quantity.

2.3. The Learning Time Delay Dose Driven Model: A Formal Framework

Integrating the above definitions and principles, we construct a cognitive dynamics model framework with clear predictive power. The core order parameter of this model is the learning time delay dose D.

Its evolutionary path can be described by the following formal sequence:

 doi:10.20944/preprints202512.0800.v1

Goal-Oriented Interaction

$\rightarrow$ Experiences n instances of effective learning delay $(\Delta t_i > \Delta t_c)$

$\rightarrow$ Learning time delay dose D continuously accumulates

$\rightarrow$ D crosses the individual- and architecture-specific critical threshold Θ_c

$\rightarrow$ Cognitive symmetry breaking

$\rightarrow$ Interaction architecture internalization is complete

In this model, the learning time delay dose D is not information in the traditional sense; it does not carry semantic content. Its role is more akin to a physical force that, through critical dynamics, shapes the potential "cognitive spacetime" for all cognitive activity. The internalized architecture then defines the "geometric structure" of this cognitive spacetime—henceforth, both the stream of conscious thought and the narratives of the subconscious dream world will tend to operate along the "geodesics" prescribed by this geometric structure.

With this, we have established a new theory with clear concepts, solid principles, and a formalized framework. It not only provides a direct mechanism for explaining the Mayer dream anomaly but also arms us with systematic critical tools to thoroughly examine the deep-seated flaws of the mainstream cognitive science paradigm."'latex

3. Systematic Critique I: The Poverty of Dream Theories and the Absence of Architectural Internalization

Having established the conceptual foundation and first principles of the new paradigm, we now turn to a systematic examination of existing theories. Dreams, as the most intimate and unconstrained realm of human consciousness, provide an ultimate testing ground for evaluating the validity of cognitive theories. However, the three major mainstream theories currently dominating dream research—the Activation-Synthesis Hypothesis, the Content Reorganization Hypothesis, and the Continuity Hypothesis—all exhibit a startling theoretical poverty when confronted with the "architectural fixation" phenomenon revealed by Mayer (2025). They share a fundamental blind spot: an inability to comprehend that the interaction architecture itself can be the core object of cognitive representation.

3.1. Critique of the Activation-Synthesis Hypothesis: The Pallor of Randomness and the Absence of Structure

- **Core Tenet Recap:** Proposed by J. Allan Hobson and Robert McCarley, this hypothesis posits that the essence of dreaming originates from random neural signals periodically emitted by the brainstem (particularly the pons) during REM sleep. This bottom-up "activation" bombards the cerebral cortex, and higher cognitive centers are forced to "synthesize" these chaotic signals into barely coherent narratives and imagery. Dreams are thus seen as a "cognitive hallucination," a byproduct of the brain seeking meaning in noise.
- **Attempted Explanation and Failure regarding the Mayer Phenomenon:** According to this hypothesis, the recurrent appearance of the ChatGPT interface in dreams should be interpreted as the brainstem's random activation accidentally triggering cortical areas related to visual interfaces and text input, with the synthesis mechanism cobbling these elements into a vague scene about using AI.
- **Analysis of Fundamental Flaws:**
    1. *The Specificity Paradox:* The core of this hypothesis lies in "randomness." Yet, the data from the Mayer report shows astonishing specificity and consistency—up to 93% of AI-related nightmares fixate on the *same* interaction interface. How could random, highly variable brainstem activation systematically and across individuals synthesize the *exact same*, highly specific interaction architecture? This is akin to expecting a hurricane sweeping through a

 doi:10.20944/preprints202512.0800.v1

junkyard to repeatedly assemble the same precise computer. The probability is logically low enough to refute random activation as the primary driver.

2. *Neglect of Architectural Precision:* The hypothesis completely fails to explain why the synthesized element is the formal framework of the interaction (the dialog box, the waiting state) rather than more emotionally charged content (e.g., AI rebellion or comfort). It only explains where the "materials" of the dream might come from but is utterly powerless to explain why these materials adhere to such a precise structural logic. Random activation can produce image fragments, but it cannot produce the grammar governing the sequence of images.

3.2. Critique of the Content Reorganization Hypothesis: The Misjudgment of "Significant" Content

- **Core Tenet Recap:** This hypothesis locates the function of dreaming in the offline processing of memory, specifically the integration, consolidation, and assimilation of significant daily experiences (especially emotional, unresolved conflicts) into existing long-term memory networks. Dreaming is the brain's "nocturnal workshop," processing the informational "raw materials" ingested during wakefulness.
- **Attempted Explanation and Failure regarding the Mayer Phenomenon:** Proponents of this hypothesis might argue that interaction with ChatGPT has become a "significant" daily experience for modern humans and thus becomes material for dream processing.
- **Analysis of Fundamental Flaws:**
  1. *Confusion of Content and Architecture:* This hypothesis commits a fundamental category mistake. It presupposes that what is reorganized is the *content* of experience. However, Mayer's data clearly indicates that what is preferentially "reorganized" is not the semantic content of the interaction (what was discussed) but its formal architecture (how the interaction occurred). What the brain diligently consolidates at night is not the specific topics of conversation with AI, but the "question-wait-answer" interaction protocol itself. The hypothesis suffers from a fatal ambiguity on the fundamental question of "what constitutes significant content," failing to explain why the *form* of interaction would be prioritized for processing during precious offline resources over the *substance* of the interaction.
  2. *Deviation from the Emotional Core:* The hypothesis often emphasizes the processing of emotional memories. Yet, in ChatGPT interactions, the strongest emotions likely arise from the dialogue content itself (e.g., receiving a brilliant idea or a terrifying answer). But the dream "chooses" the relatively emotionally neutral interface as its core. This strongly suggests that the mechanism driving internalization is independent of, and perhaps even prior to, traditional emotional salience.

3.3. Critique of the Continuity Hypothesis: Circularity and the Selectivity Problem

- **Core Tenet Recap:** This is an intuitive and popular hypothesis proposing that dreams exhibit continuity in themes, concerns, and content with an individual's waking thoughts and behaviors. The so-called "what one thinks about by day, one dreams about by night."
- **Attempted Explanation and Failure regarding the Mayer Phenomenon:** This hypothesis seems to offer the most direct explanation: people use ChatGPT extensively during the day, so they dream about it at night.
- **Analysis of Fundamental Flaws:**
  1. *The Trap of Circular Reasoning:* The Continuity Hypothesis immediately falls into circular reasoning when explaining architectural internalization. It claims the ChatGPT interface appears in dreams because one thinks/uses it while awake. But this merely pushes the question back one step: *Why* does waking "thought" about ChatGPT so, selectively manifest as a concern with the *form* of its interaction interface, rather than with its capabilities, social impact, or ethical implications? The hypothesis uses the empty label "continuity" to

 doi:10.20944/preprints202512.0800.v1

obscure the real question needing explanation—namely, *in which dimension* does continuity specifically manifest?

2. *Failure of Feature Selectivity:* The hypothesis lacks any principled mechanism to determine which of the infinite features of waking experience get "continued" into dreams. When interacting with ChatGPT, we simultaneously experience its content, its interface, its speed, its utility, its social meaning... Why is it *only* that interaction architecture that is so faithfully "continued"? The Continuity Hypothesis itself cannot answer this selectivity problem; it can only describe post hoc, not predict.

3.4. Conclusions

The critique of mainstream dream theories reveals a common, deep-seated predicament: they are all rooted in a content-centric or experiential-holistic model of cognition. They view dreams as a kind of processing or reflection of daily content (whether random fragments, important memories, or overall thoughts). Consequently, when faced with the "architectural fixation" anomaly, which explicitly demands prioritizing form over content, they all fall silent. They cannot find a suitable place for "interaction architecture" within their conceptual frameworks—it is neither a random fragment, nor a typical important memory, nor a simple linear extension of waking thought. This systematic failure eloquently demonstrates that dream research, and indeed cognitive science as a whole, urgently requires a paradigm revolution that liberates architecture from content. Our Interaction Architecture Internalization Principle and Learning Time Delay Dose Driven Model are proposed precisely in response to this urgent need.

4. Systematic Critique II: The Internal Paradoxes of the Process-Centered Paradigm

Having exposed the poverty of traditional dream theories, we must now direct our critical focus towards the more hegemonic and sophisticated process-centered paradigm in contemporary cognitive science. This paradigm, represented by Predictive Coding, the Free Energy Principle, and Integrated Information Theory, attempts to provide a unified, mathematically-grounded explanation for all mental phenomena. However, when tested against the touchstone of "interaction architecture internalization," these grand theoretical frameworks likewise reveal their inherent, principled paradoxes.

4.1. Critique of Predictive Coding and the Free Energy Principle: Elegant Mathematics, Lost Object

Predictive Coding and its more general formulation—the Free Energy Principle—portray the brain as an "inference machine" relentlessly pursuing prediction error minimization. The essence of cognition is reduced to a single, universal process: predicting sensory inputs based on internal generative models and updating these models using prediction error.

- **Critique 1: Meta-Theoretical Confusion—Predicting 'Content' vs. Predicting 'Architecture'**
  - *Core Defect:* The framework consistently handles its object of prediction ambiguously. Predicting 'signal content' (e.g., the next word, the next visual feature) and predicting the 'interaction architecture' (e.g., "I am in a 'question-wait-answer' interaction mode") are two entirely different levels of task.
  - *In-Depth Analysis:* Interaction architecture is a constancy of relations and rules; it is not directly present in any single sensory channel. It is the transcendental framework governing the flow of content. Predictive Coding excels at predicting the content flow *given a framework*, but it lacks the necessary mechanisms to represent the framework *itself*. The system can learn to predict "approximately 2 seconds of silence after a question," but this is merely a prediction of a delayed event, not a representation of the interaction architecture *as an abstract entity* causing this delay. It mistakes the *effect* of the architecture for the architecture itself.
- **Critique 2: Failure of Time Delay Integration—The Chasm from Predicting Delay to Internalizing Architecture**

 doi:10.20944/preprints202512.0800.v1

  - *Core Defect:* The framework lacks any explicit mechanism capable of transforming the confirmation of a 'structural delay' (i.e., predicting "no significant new signal at this moment") into the identification, representation, and reinforcement of the specific interaction architecture causing this delay.
  - *In-Depth Analysis:* From the perspective of the Free Energy Principle, a delay that is consistently predicted accurately (i.e., no prediction error) precisely means the model is already well-adapted to it. The system's more natural tendency would be to adapt to or ignore this delay, treating it as an invariant background of the environment, rather than highlighting and internalizing it as the core defining feature of that interaction mode. The Predictive Coding framework cannot explain why the brain would 'obsess' over a delay signal it has successfully predicted (i.e., error minimized) and elevate it to the status of a cornerstone of cognitive structure.
- **Critique 3: The Model Selection Paradox—Why the Complex Architectural Model?**
  - *Core Defect:* Under the Free Energy Principle, the system must trade off model complexity against its fit to the data (e.g., balancing complexity and accuracy). However, the framework cannot provide an a priori, principle-based reason why the cognitive system would necessarily select a more complex architectural generative model that incorporates 'time delay' as a key variable, over a simpler model describing only the statistical regularities of content.
  - *In-Depth Analysis:* A simpler model based solely on content co-occurrence (e.g., statistical associations between 'question words' and 'answer words') could also achieve prediction to some extent. Why would the brain go to the 'trouble' of constructing a more complex generative model that reifies the 'interaction protocol'? The Predictive Coding framework attributes this to a trade-off, but this is essentially a post-hoc rationalization. It cannot explain the cognitive system's inherent preference for form and relation, exposing its incompleteness as a meta-theory of cognition—it describes *how* the mind optimizes, but cannot fundamentally explain *why* it optimizes towards the representation of 'architecture'.

4.2. Critique of Integrated Information Theory: Grand Static Geography, Missing Historical Dynamics

Integrated Information Theory (IIT) takes a distinctly different path. It starts not from process but from existence, attempting to derive the nature of consciousness from the causal structure of a system. The "level" of consciousness of a system is measured by its integrated information amount Φ , with complex causal structures possessing high Φ necessarily accompanied by conscious experience.

- **Critique 1: Prisoner of Static Structure—Unable to Explain the Emergent History of Architecture**
  - *Core Defect:* IIT is inherently a static, structure-centered theory. It is adept at describing why a system has a certain conscious experience at a given point in time (because it has such-and-such a causal structure). However, it is completely powerless to explain *how* and *why* this specific high-Φ structure emerged over time through concrete interactive experience.
  - *In-Depth Analysis:* IIT can claim that when dreaming of the ChatGPT interface, a certain brain region forms a high-Φ complex causal structure corresponding to that interface representation. But this is merely renaming the phenomenon. The crucial question is: *How did this specific high-Φ structure form?* Why did the ChatGPT interface, and not some other structure, become this high-Φ complex? IIT is silent on learning, development, and historical formation. It paints a grand "geography" of consciousness but lacks its "history" entirely. Our model provides this history: it is the accumulation of the learning time delay dose that drives the reorganization of the brain's causal structure, ultimately reaching a new high-Φ stable state representing that interaction architecture.
- **Critique 2: Absence of Temporal Dynamics—Time Delay as Causal Driving Force**
  - *Core Defect:* IIT severely neglects the central role of time, particularly time delay, in shaping causal structure.

 doi:10.20944/preprints202512.0800.v1

- *In-Depth Analysis:* In the IIT framework, causal power is primarily determined by the system's structure at the present moment. However, as revealed by delay differential equations, time delay itself can be a decisive parameter of a system's dynamics. Our model demonstrates that learning time delay is not an insignificant internal parameter of the system, but rather the order parameter that drives the bifurcation of the causal structure towards a new stable state (i.e., new architecture internalization). IIT cannot integrate "time delay" into its core postulates; therefore, it cannot foresee how the accumulation of a seemingly simple physical variable (waiting time) can catalyze entirely new, complex conscious content (such as the fixation on an interaction interface).

4.3. Conclusions

The critique of the process-centered paradigm reveals an even more profound predicament. Predictive Coding/Free Energy Principle reduces cognition to a single optimization process but loses sight of the hierarchical nature of its object, unable to account for the representational source of 'architecture' as a transcendental object. Integrated Information Theory focuses on the structural state of conscious existence but completely lacks its formative historical dynamics. Both, in different ways, sever process from structure, and history from state.

Our Interaction Architecture Internalization Model is born precisely to bridge this schism. It explicitly states that cognitive structure (the internalized architecture) emerges through dynamical phase transition from a concrete, quantifiable historical process (the accumulation of learning time delay dose). It integrates the innate structure emphasized by Chomsky, the constructive process focused on by Piaget, and the temporal relativity inspired by Einstein within a unified mathematical framework (Landau phase transition). Thus, the failure of existing mainstream paradigms to explain the phenomenon of architectural internalization is no longer an accidental oversight, but an inevitable manifestation of the inherent defects within their theoretical cores. "'

5. Systematic Critique III: The Legacy and Specters of Behaviorist and Learning Theories

Having dissected the inherent flaws of dream theories and the process-centered paradigm, we must complete the final segment of our critical map: a reckoning with the behaviorist and learning theories that dominated twentieth-century psychology and continue to exert profound influence. These theories—including Behaviorism, Cognitive Schema Theory, and Embodied Cognition—have shaped our understanding of learning, cognition, and the environment in their respective ways. However, when confronted with the phenomenon of "interaction architecture internalization," they either reveal their fundamental errors or expose the severe limitations of their explanatory power.

5.1. A Renewed Critique of the Behaviorist Specter: Skinner's Return in the Computational Age

Although Noam Chomsky's 1959 critique of B.F. Skinner's *Verbal Behavior* is widely regarded as having theoretically dismantled the foundations of Behaviorism, its specter has persisted in a transmuted form within contemporary cognitive science.

- **Core Critique: Process-Centrism as Neo-Behaviorism**
  - *Argument:* Strong versions of Predictive Coding and the Free Energy Principle can be epistemologically viewed as "*Skinnerism Reborn in the Computational Age*." Skinner attempted to derive all verbal behavior solely from external "stimulus," "response," and "reinforcement" histories (i.e., observable content and process). Similarly, the strong process-centered paradigm attempts to derive all cognitive phenomena from a single, universal "prediction error minimization" process (a computationalized "reinforcement").
  - *Persistence of the Fundamental Flaw:* Both share a fatal error: they attempt to derive an understanding of deep structure from the description of surface processes. Chomsky demonstrated that "grammatical architecture" cannot be acquired from "verbal content"; we argue that the

 doi:10.20944/preprints202512.0800.v1

representation of "interaction architecture" likewise cannot be necessarily derived from the "prediction/optimization process." Both systematically underestimate the mind's propensity to actively impose its inherent formal structures to understand the world. The contemporary process-centered paradigm, despite its elegant Bayesian mathematical garb, commits the same error as its behaviorist predecessor in evading the core insight that "*architecture precedes content*."

5.2. Critique of Cognitive Schema Theory: The Abstraction Dilemma and the Missing Variable

Cognitive Schema Theory, as a rebellion against Behaviorism, correctly emphasized the role of internal knowledge structures. However, it appears too abstract and rigid when explaining architecture internalization.

- **Core Critique: The Explanatory Gap from Abstract Schema to Concrete Architecture**
  - *The Problem:* Schema theory might describe the internalized ChatGPT interface as a "human-machine Q&A schema." But this abstract label cannot explain the high specificity of the dream content in the Mayer report—why is it ChatGPT's *specific* visual interface, that *dynamic process* of text stream generation, that is activated and internalized, rather than an abstract, generic schema applicable to all Q&A scenarios?
  - *Fundamental Defect:* Schema theory lacks a crucial dynamical variable to quantify the differential impact of different interaction architectures on the cognitive system. It cannot explain why *this* particular form, and not that similar one, becomes solidified. Our model fills this gap by introducing the learning time delay dose (D). It is the unique, repeatable, and significant temporal delay pattern characteristic of ChatGPT interaction that makes its *concrete architecture* (not the abstract concept of "Q&A") the object of internalization. Schema theory describes the static organization of knowledge but fails to provide its dynamic formation mechanism.

5.3. Critique of Embodied Cognition: The Limits of the Body and the Transcendence of Form

Embodied Cognition theory powerfully emphasizes the central role of the body, action, and environment in shaping cognition. However, its radical versions appear inadequate, even misguided, when explaining the phenomenon of architecture internalization.

- **Core Critique: High Architectural Internalization under Low Bodily Engagement**
  - *Counterevidence:* Interaction with ChatGPT is a typical activity characterized by *low bodily engagement* but *high architectural internalization*. Bodily actions are reduced to typing and scrolling, and sensory channels are primarily limited to vision. According to the radical embodied cognition view, such "disembodied," abstract symbolic interaction should produce relatively shallow cognitive traces.
  - *Fundamental Defect:* Yet, Mayer's data show the opposite result: this low-body-engagement interaction leads to extremely profound and specific architectural internalization. This suggests that the core mechanism of cognitive internalization is not the bodily action itself, but the *temporal logic of the interaction* defined by time delays. Embodied Cognition confuses the physical vehicle of interaction (bodily actions, sensory modalities) with its logical form (architecture). Our model demonstrates that even when bodily engagement is minimized, as long as the temporal-delay architecture of that interaction is significant and repeated, it is sufficient to drive a profound cognitive phase transition. Embodied Cognition correctly points to the coupling of cognition and environment, but it mistakenly attributes the primary medium of this coupling to the body, rather than to the more universal temporal form of the interaction architecture.

5.4. Conclusions

The critique of the legacy of behaviorist and learning theories ultimately points to a clear conclusion: neither the behaviorist specter attempting to derive internal structure from external processes,

 doi:10.20944/preprints202512.0800.v1

nor schema theory which is too abstract and lacks dynamical explanatory power, nor embodied cognition which overly anchors cognition in the bodily dimension, is adequate to the task of explaining "interaction architecture internalization."

Their collective failure, in turn, highlights the unique advantages of our proposed Learning Time Delay Dose Driven Model. It relies neither on mysterious innate structures nor is it satisfied with describing surface processes; it neither lapses into abstract rigidity nor is it constrained by bodily physicality. It anchors the root cause of cognitive structural change in an objectively measurable physical variable—the cumulative dose of time delay—that transcends specific content and bodily modalities. With this, we have systematically cleared the theoretical ground of the major schools of the old paradigm, paving the way for the positive construction of our new model's applicative edifice. "'

6. Unified Model Application I: Language Acquisition and the Resolution of the Chomskyan Problem

The "Interaction Architecture Internalization Model" we propose is by no means intended merely to explain an isolated dream anomaly. Its true power lies in its unified explanatory force across different domains of cognitive phenomena. We now turn to perhaps the most famous and enduring puzzle in cognitive science—language acquisition, namely, Chomsky's "poverty of the stimulus" argument. We will argue that our model is not only fully compatible with Chomsky's profound insight but also provides a clear, mechanistic pathway for its realization, thereby potentially resolving the long-standing nature-nurture debate that has plagued linguistics.

6.1. Re-examining the "Poverty of the Stimulus": A Shift in Perspective from Content to Architecture

Chomsky's core argument is well-known: young children, within a very short time, successfully infer the highly abstract and complex system of grammatical rules of their native language from the limited, fragmented, and error-ridden linguistic input (content) they are exposed to. This ability to derive a perfect, infinitely generative architecture (Universal Grammar and particular grammar) from flawed, finite content is logically impossible unless the child's mind comes pre-equipped with strong constraints and preferences for specific formal structures—i.e., Universal Grammar.

The traditional debate deadlock is this: empiricists seek to find more hidden cues in the "content" or rely on more powerful learning mechanisms; nativists assert that these are insufficient and must appeal to rich innate structures. Our model breaks this impasse through a fundamental shift in perspective: the primary object that children internalize is not the "content" of linguistic expressions, but the "architecture" of linguistic interaction.

6.2. The Linguistic Interpretation of the Time Delay Dose Model: Internalizing the Interaction Architecture of 'Motherese'

Let us redescribe the language acquisition situation. Children are not immersed in a pure stream of language content. They are immersed in goal-directed social interactions that have a very specific temporal structure. The most typical is the speech of caregivers, known as 'motherese'.

- **'Motherese' as a Time-Delay Architecture:** 'Motherese' is characterized not only by its simplified syntax and exaggerated prosody but, crucially, by its unique interaction timing. It is a classic "elicit-response-feedback" architecture:
    1. The caregiver produces an eliciting utterance or question (e.g., "Look! Ball!"), followed by a waiting pause (time delay Δt).
    2. The child produces a vocalization or gaze as a response.
    3. The caregiver immediately provides an enthusiastic, deterministic feedback (e.g., "Yes! Ball!"), thereby closing the interaction loop.
- **Accumulation of Time Delay Dose:** In this process, the key dynamical variable is that waiting pause (Δt). This pause is not emptiness; it is a time window filled with anticipation and social

 doi:10.20944/preprints202512.0800.v1

tension. Thousands of such interactions expose the child to a massive, language-specific learning time delay dose (D_{language}).

- **Internalization of Grammatical Architecture:** When the dose D_{language} accumulates to the critical threshold Θ_{language}, the cognitive system undergoes a phase transition. What the child internalizes is not any specific word or sentence (content), but the abstract architecture governing these interactions—namely: verbal acts follow a "initiate-wait-respond-confirm" temporal logic, and responses must conform to a set of formal rules governing sequence legitimacy (grammar). They internalize the "rules of the game" of language, not just the "words" spoken in the game. This internalized architecture is the core of what Chomsky calls the particular grammar. The child's extreme sensitivity to this temporal interaction architecture is grounded in the initial cognitive state defined by Universal Grammar.

6.3. Resolving the Nature-Nurture Opposition: Time Delay as the Bridge

Our model deftly resolves the either-or opposition between nativism and empiricism, transforming it into a dynamic synergy.

- **Universal Grammar as a 'Sensor' for Time-Delay Architectures:** We need not postulate an innate "grammar book" containing all possible grammatical rules in full detail. A more economical and plausible assumption is that Universal Grammar manifests as an innate "sensor" or "prepared state" that is particularly sensitive to and optimized for specific types of time-delay interaction architectures. The human mind is "preset" to rapidly recognize and internalize interaction dynamic patterns characterized by "turn-taking," "causal timing," and "closed feedback." Language is the most expression of this pattern.
- **Experience as the 'Supplier' of Time Delay Dose:** The specific linguistic environment the child is in (experience) provides the specific type of time delay dose. Different languages' motherese may vary in rhythm, pause length, and feedback style, thus providing slightly different "recipes" of time delay patterns. These specific experiential data, processed by the innate sensor, ultimately "crystallize" into a specific internal grammatical architecture.

Therefore, language acquisition is no longer seen as learning language "content," but is understood as: the inevitable, innately constrained internalization process of the language architecture, triggered when the specific time delay dose of linguistic interaction reaches a critical point. Experience provides the necessary dynamical drive (dose), while innate factors set the system's phase transition pathway and the possible final stable states (architectural forms).

By shifting the focus from linguistic content to the temporal architecture of linguistic interaction, and positioning learning time delay as the core variable, our model provides a novel, dynamically grounded solution to the Chomskyan problem. It suggests that humans' grasp of grammar stems not from statistical learning of sentence content, but from the profound internalization of the basic interactive form constituted by "waiting and responding" in their social world. This opens a promising new path for understanding how we become linguistic beings.

7. Unified Model Application II: Meme Propagation and the Replication of Cultural Patterns

The explanatory power of our model extends beyond the microcosm of individual cognition to the macrocosm of culture and society. Here, we turn to a core concept describing cultural evolution—the *meme*—and argue that our model, based on time delay and architectural internalization, can provide meme theory with a more solid, predictive dynamical foundation, thereby explaining why certain cultural patterns spread virally while others rapidly perish.

 doi:10.20944/preprints202512.0800.v1

7.1. The Shortcomings of Meme Theory: A Paradigm Shift from Content Replication to Architectural Internalization

The concept of the "meme," proposed by Richard Dawkins, analogizes the basic unit of cultural transmission to the gene, suggesting they replicate between host minds through imitation. However, classical memetics faces a fundamental ambiguity: *What exactly constitutes a meme?* Is it a melody, a slogan, an idea, or a way of behaving? This ambiguity often reduces its explanatory power to post-hoc circularity—"it spread because it was a successful meme."

Our model provides a clear answer: what is truly replicated and propagated is not the cultural "content" itself, but the *interaction architecture* for producing, consuming, or engaging with that content. The success of a cultural unit lies in the efficiency with which its carried interaction protocol can be internalized.

- **Redefining the Meme:** A successful meme is essentially an interaction architecture with high internalization efficiency. Its "infectivity" lies not in how appealing its semantic content is, but in how its interactive form enables the host to rapidly cross the internalization critical threshold Θ_c after being exposed to a relatively low learning time delay dose (D).

7.2. Memes as Architectural Internalization: Time Delay Efficiency is Key to Propagation

Let us illustrate this viewpoint with several classic cases:

1. **Internet Challenges and Relay Memes (e.g., the "Ice Bucket Challenge"):**
   - *Traditional Explanation:* The content (the charitable act of fundraising for ALS) is compelling.
   - *Architectural Internalization Explanation:* Its core is an extremely streamlined and efficient interaction architecture: "Accept Challenge $\rightarrow$ Perform Specific Action (short delay, immediate effect) $\rightarrow$ Video Verification and Nominate Others (deterministic feedback and social reinforcement)." The time delays in this architecture are minimal (action and feedback are almost simultaneous), and the rules are clear, allowing individuals to understand and internalize its "rules of the game" almost instantaneously, driving explosive propagation. Its success lies in the architecture's low time delay and high determinism.
2. **Short-Form Video Platform "Recipes" and Formats (e.g., TikTok/Reels vertical short videos):**
   - *Traditional Explanation:* Short, concise content suits the fragmented attention spans of modern people.
   - *Architectural Internalization Explanation:* The platform itself is a powerful architectural internalization engine. It imposes a unified interaction protocol: "Infinite Scroll (extremely short delay triggers new content) $\rightarrow$ Brief, High-Stimulus Audiovisual Content Stream (immediate feedback) $\rightarrow$ Likes/Comments/Shares (quantified, timely social reinforcement)." During use, the user's learning time delay dose D accumulates at a very high frequency, rapidly internalizing this cognitive mode of "rapid switching and seeking immediate stimulation." This explains not only the platform's stickiness but also why content conforming to this architecture (e.g., fast cuts, strong rhythmic music) propagates more easily—it resonates with the already internalized cognitive architecture.

7.3. Time Delay and Addictive Design: Commercial Applications and Social Consequences of Architectural Internalization

Our model provides a profound mechanistic explanation for understanding the "addictiveness" of contemporary technological products. These products essentially engage in precise "*time delay engineering*," aimed at minimizing the critical dose Θ_c required for users to internalize their interaction architecture.

- **Variable-Ratio Reinforcement Schedule:** This is central to gambling and many social media feed algorithms. By providing unpredictable but occasionally dense rewards (e.g., a viral video, an important like notification), it creates a high-intensity time delay pattern. The user is constantly

 doi:10.20944/preprints202512.0800.v1

in a "search-wait-(possibly) obtain" loop. This uncertainty significantly enhances the salience of the interaction, drastically increasing the accumulation efficiency of the effective time delay dose D, rapidly locking the user into the internalized architecture and leading to compulsive usage behaviors.

- **"Infinite Scroll" and "Pull-to-Refresh":** These designs reduce the time delay for seeking new content to nearly zero (one swipe or a gesture). They eliminate the inherent "natural delays" present in traditional interactions (like turning a page, clicking 'next') that might prompt reflection, creating a seamless, frictionless consumption experience that makes the architectural internalization process exceptionally smooth and imperceptible.

7.4. Conclusions

From the perspective of meme propagation, our model achieves a crucial paradigm shift: the unit of cultural replication is not informational content, but interaction architecture. The competition in cultural evolution lies not in the "quality" of the content itself, but in the "time delay efficiency" competition among different interaction architectures vying for internalization by the human mind. An interaction architecture that can complete the internalization phase transition with a lower time delay dose and at a faster rate will possess greater propagative power and vitality.

This explains why simple, repetitive cultural patterns with clear, immediate feedback (from pop song choruses to social media challenges) are so prevalent—they are the inevitable selection pressure of human cognitive architecture under time delay dynamics. Thus, our model successfully connects micro-level cognitive dynamics with macro-level cultural phenomena, demonstrating its significant potential as a unified theory."'latex

8. Unified Model Application III: Media Theory and the Formation of Personality Structure

The explanatory scope of our model ultimately reaches two of the most fundamental levels shaping human existence: first, our macro-level integration with the technological environment—media theory; and second, the micro-foundation of our individuality—personality structure. We will argue that both can be understood as profound products of interaction architecture internalization operating at different time scales.

8.1. A Dynamic Interpretation of McLuhan's "The Medium is the Message"

Marshall McLuhan's famous dictum, "the medium is the message," posits that the profound impact of any medium on society and culture comes not from the specific content it carries, but from the change of scale, pace, or pattern introduced by the form of the medium itself. Our model provides a precise cognitive-dynamic interpretation of this insight.

- **The "Message" as Internalized Interaction Architecture:** In our framework, McLuhan's "message" is precisely the *internalized interaction architecture* on a mass scale. The proliferation of a new medium is essentially the process of imprinting its unique interaction protocol onto the collective cognition through the accumulated learning time delay dose (D) of millions of users.
- **Comparative Case Studies:**
  - **Printing Press:** Its interaction architecture was "linear, static, visual, private reading." The internalization of this architecture cultivated linear, logical, and individualistic thinking habits. The key difference in the transition from manuscript to print lay in the change of time delay in the interaction—readers could pause at will, re-read (autonomous control of delay)—which was fundamentally different from the fixed delays of listening to an oral presentation.
  - **Television:** Its architecture was "one-way, linear flow, audio-visual fusion, passive reception." It internalized a continuous, passive, emotional cognitive mode. The advent of the remote

 doi:10.20944/preprints202512.0800.v1

control slightly altered this architecture, introducing minimal interactivity and delay control, but did not fundamentally change its one-way flow nature.

- **Smartphones/Social Media:** Its architecture is "pervasive, fragmented, notification-driven, multi-tasking parallel processing." The core feature of this architecture is the extreme shortening and unpredictability of its time delays (pull-to-refresh, instant messages). When this architecture is internalized, it shapes cognitive habits characterized by distracted attention and a strong craving for immediate feedback. McLuhan's "message" here is our internalized architecture, and the medium is the delivery system for a specific time delay dose.

8.2. Personality as Sedimentary Layers of Internalized Architecture: From Attachment Styles to Technological Neurosis

A persistent core assumption in personality psychology is that early social interactions shape enduring personality structures. Our model provides a concrete formation mechanism for this assumption.

- **Reconceptualizing Attachment Theory:** Attachment styles can be understood as the outcome of internalizing the repetitive interaction architecture between an individual and their primary caregiver(s). The key dynamical variable is the time delay (Δt) and its pattern between the infant emitting a distress signal (e.g., crying) and receiving a comforting response.
  - **Secure Attachment** corresponds to a stable architecture: "Signal emitted $\rightarrow$ moderate, predictable delay $\rightarrow$ consistent, appropriate response." The internalization of this architecture forms a basic sense of trust in the world and others.
  - **Avoidant Attachment** may stem from an architecture: "Signal emitted $\rightarrow$ prolonged or unpredictable delay $\rightarrow$ absent or response." The individual internalizes the interactive rule that "expressing needs is futile."
  - **Anxious-Ambivalent Attachment** might correspond to an architecture: "Signal emitted $\rightarrow$ highly unstable, unpredictable delays and responses." The individual internalizes the pattern that "the world is unpredictable, requiring hyper-vigilance and amplified signaling."

  Here, personality structure is the sediment deposited by these early social interaction time-delay architectures within the cognitive and affective systems.
- **The Emergence of Technological Neurosis:** The AI-related nightmares in the Mayer report can be seen as a manifestation of a new form of "*technological neurosis*." When a powerful, non-human interaction architecture like ChatGPT's, through high-frequency use over a short period (days or weeks), causes the learning time delay dose D to accumulate rapidly and exceed Θ_c, it forces a swift internalization. This new, potent cognitive structure conflicts and dysregulates with the individual's pre-existing personality architectures formed over long periods of human social interaction (e.g., complex emotional interpretation, tolerance for ambiguity). The anxiety and fixation in dreams are manifestations of this cognitive structural conflict, the growing pains of the mind attempting to digest and integrate this forcibly implanted "technological organ."

8.3. Critique of Mainstream Personality Theories: Beyond Static Traits and Reductionism

Our model challenges mainstream personality psychology (e.g., the "Big Five" model):

- **Critique of Trait Theory:** Trait theory describes personality as a set of static, descriptive dimensions (e.g., Extraversion, Neuroticism). It is adept at describing differences but completely unable to explain how these traits *form* from an individual's life history. It is a theory of *state*, not of *formation*.
- **Complementing Biological Reductionism:** The biological approach emphasizes the genetic and neural bases of personality. Our model does not contradict this but integrates it: genetic factors may preset an individual's initial sensitivity to specific types of time-delay architectures (i.e., initial settings for Δt_c and Θ_c), but the specific, final personality structure is realized through the accumulation of time delay doses in interaction with specific environmental architectures. Genes

 doi:10.20944/preprints202512.0800.v1

provide the basic timbre of the instrument, while life experiences (time-delay architectures) play the specific melody.

8.4. Conclusions

By incorporating media theory and personality structure into its explanatory purview, our model demonstrates its grand ambition as a unified framework for cognitive science. It reveals a coherent path from the micro to the macro: basic interaction timing (time delay) $\rightarrow$ internalized cognitive architecture $\rightarrow$ stable individual personality and collective cultural patterns. Whether McLuhan's laws of media or Bowlby's attachment theory, they can find a home within this dynamical model, with time delay as the order parameter and architecture internalization as the mechanism, gaining unprecedented precision and explanatory depth. This suggests that understanding the mind's past and future hinges on understanding how it measures and internalizes the invisible forms of time—those that shape its very structure through interaction with the world."'latex

9. Philosophical Foundations: The Chinese Room, the Problem of Other Minds, and the Learning Time Delay Equivalence Principle

Any grand theory attempting to explain human cognition must ultimately face the interrogation of its philosophical foundations. Among these, the "Problem of Other Minds" presents a seemingly insurmountable epistemological obstacle: How can we truly know the content of another's mind? The core of our model—the Learning Time Delay Equivalence Principle—does not seek to deny or solve this ancient philosophical puzzle, but rather, with a profoundly pragmatic spirit, to *circumvent* it, thereby establishing a solid, operational, objective foundation for cognitive science. Searle's "Chinese Room" thought experiment provides an excellent philosophical analogy for understanding this strategy.

9.1. The Cognitive Science Dilemma of the Problem of Other Minds: The Impenetrable "Micro-Mental Activities"

When examining the phenomena revealed by the Mayer report, we confront an insurmountable epistemological limitation: during the few seconds when a user asks ChatGPT a question and waits for an answer—i.e., within the learning time delay Δt window—we can *never* know precisely or completely the specific content of the "micro-mental activities" occurring in that user's brain.

- **The Opacity of Internal States:** Is the user anxiously anticipating? Mentally rehearsing possible answers? Growing impatient due to the delay? Or daydreaming about dinner? These specific, qualitative subjective experiences are a black box for us as external observers. This is the concrete manifestation of the "Problem of Other Minds" in cognitive experimentation: we lack direct access to another's stream of consciousness.
- **The Predicament of Traditional Theories:** Many cognitive theories (especially those emphasizing content processing) implicitly require inferences or assumptions about such micro-mental activities. This often renders their explanations built on sand, as their core variables (e.g., "cognitive load," "emotional valence," "implicit reasoning") are operationally vague and difficult to measure independently.

9.2. The Lesson from the Chinese Room: A Paradigm Shift from Semantic Content to Physical Signal

John Searle's "Chinese Room" thought experiment was originally designed to critique strong AI. However, we can extract from it a crucial insight for our model:

- **Thought Experiment Recap:** A person in a room who does not understand Chinese manipulates slips of paper containing Chinese characters passed in from outside, following the instructions in a rulebook (the syntactic architecture), and produces correct Chinese slips to pass back out. To an external Chinese speaker, the room appears to "understand" Chinese. But the person inside is merely manipulating symbols, devoid of any semantic understanding.

 doi:10.20944/preprints202512.0800.v1

- **Analogy for Our Model:** In this experiment, from the external perspective, the only objectively observable, measurable physical events are the *time sequence* of slips being passed in and out. Whether the person inside is diligently consulting the rulebook or mechanically, uncomprehendingly executing steps, the macro-functional output is, at a certain level, *equivalent*. The external observer cannot, and need not, distinguish between these internal states.

Applying this analogy to our model:

- "Slip passed in" = User's exploratory operation (e.g., asking a question).
- "Operations inside the room" = User's micro-mental activities during the Δt window (their specific content is the Problem of Other Minds, unknowable).
- "Slip passed out" / "Waiting time" = The learning time delay Δt itself. This is an objective, physically measurable empty window.

The key point is that just as we cannot determine whether the person in the room "understands" by observing the input and output of slips, we cannot fully determine the specific micro-mental content during the delay period by observing the cognitive system's output. However, we find that regardless of what these undetectable micro-activities specifically are, their causal efficacy for the macro-cognitive effect of "architecture internalization" is *equivalent* to that of a purely objective, physically measurable variable—the learning time delay dose (D).

9.3. The Learning Time Delay Equivalence Principle as a Methodological Cornerstone

This is the profound philosophical implication of the Learning Time Delay Equivalence Principle. It acknowledges the intractability of the "Problem of Other Minds" and, on this basis, makes a methodologically astute choice:

Since we cannot, in principle, distinguish whether the "time delay itself" or the "micro-mental activities during the delay" is the fundamental cause driving architecture internalization, then, in constructing a scientific model, we must treat them as *causally equivalent*, and resolutely choose the observable, measurable, operational variable—the time delay dose D—as our fundamental explanans.

- **From Philosophical Dilemma to Scientific Variable:** This principle accomplishes a crucial conversion. It transforms a philosophical epistemological deadlock (the Problem of Other Minds) into a scientifically tractable dynamical variable (the time delay dose). We cease to be entangled with the unanswerable "what" (specific mental content) and focus instead on the answerable "how long" (delay length and frequency).
- **Application of Occam's Razor:** This choice adheres to Occam's razor. Rather than constructing complex theories reliant on unobservable internal states, it is more parsimonious to adopt a simpler theory based on objective observables. The time delay dose D is precisely such a "frugal" variable.
- **Establishing an Objective Foundation for Cognitive Science:** By establishing the time delay dose D as the order parameter, our model provides an unprecedented objective foundation for cognitive science. It enables predictions about cognitive structural change to rely less completely on subjective reports and vague introspection, and instead be grounded in the physical measurement of interaction timing.

9.4. Conclusions

The Chinese Room thought experiment and the Problem of Other Minds, far from weakening our model, highlight its methodological robustness and advancement through contrast. We accept that there is an "inner room" in human cognition that can never be fully. Yet we discover that the "input-output delay time" of this room itself is the equivalent, measurable "cement" driving the construction of cognitive architecture. The Learning Time Delay Equivalence Principle is therefore not merely a cognitive principle but also an epistemological one, guiding us on how to acknowledge the inherent mystery of human consciousness while still building a rigorous and progressive science of the mind. "

 doi:10.20944/preprints202512.0800.v1

10. Blinded Loop: A Validation Paradigm Constructed from Academic Misunderstanding

Any theory aiming to subvert an existing paradigm requires not only internal logical coherence and predictive power but must also withstand rigorous validation from the empirical world. The unique aspect of the validation process this theory underwent lies in this: a public critique event widely regarded as an "academic stain," combined with an unknown prior benchmark fact, unexpectedly constructed a "blinded loop" meeting the highest standards of philosophy of science, providing irrefutable evidence of ecological validity for this theory.

10.1. Components of the Validation Loop

This validation loop consists of three sequential elements:

1. **Prior Existence of the Benchmark Fact (Before March 2025):** April Mayer's survey report *AI in Dreams* was publicly released *before* the theory was constructed. Its core finding—that 93% of AI-related dreams fixated on the interaction interface rather than narrative content—constituted an objective, macro-level phenomenon awaiting explanation (denoted as Phenomenon E).
2. **Independent Construction and Publication of the Theory (May-July 2025):** While completely unaware of Phenomenon E, this study, based on self-introspection experiments, independently proposed the theoretical framework centered on "Interaction Architecture Internalization" (denoted as Theory T) and pre-printed it on PsyArXiv. A core corollary of Theory T was that dreams would solidify the *formal architecture* of interaction, not its semantic content.
3. **Critical Misunderstanding as Methodological "Blinding" Attestation (August 2025):** The public critique and retraction of this study by *Nature* magazine were viewed as a negative academic event. However, from a scientific methodology perspective, this event held decisive significance: throughout the entire process of critique, neither the critics, the criticized, nor the academic community invoked Phenomenon E as evidence. This public academic act constituted an unintentional yet powerful behavioral proof, conclusively demonstrating that the proposal of Theory T occurred under strict "blinding"—i.e., its construction was completely independent of prior knowledge of the benchmark fact E.

10.2. Clarification of the Misunderstanding and Formation of the Logical Loop

When Theory T and Phenomenon E were juxtaposed at a later stage, their high consistency (Theory T accurately predicted Phenomenon E) clarified the initial "academic misunderstanding" and transformed it into a robust logical loop. The reasoning structure is as follows:

- **Premise 1 (Temporal Irreversibility):** The occurrence time point of Phenomenon E (T_E) precedes the proposal time point of Theory T (T_T).
- **Premise 2 (Independence):** There is conclusive public behavioral evidence that the proposal of Theory T was independent of knowledge of Phenomenon E.
- **Conclusion (Confirmation):** Therefore, Theory T cannot be a post-hoc induction or fit to Phenomenon E, but must be recognized as a successful, *a priori* prediction of Phenomenon E.

This loop renders any skepticism dismissing Theory T as "coincidence" or "post-hoc explanation" logically untenable. The public academic actions of the initial critical force, ironically, became the most powerful third-party attestation of Theory T's independence and predictive power.

10.3. Methodological Implications

This unique validation path offers profound methodological implications:

1. **The Validating Power of Natural Experiments:** The social process of scientific discovery can sometimes accidentally create validation environments stricter than controlled experiments. This event demonstrates that public academic exchange (even in the form of critique) can provide crucial methodological attestation for a theory.

 doi:10.20944/preprints202512.0800.v1

2. **The Dialectics of Paradigm Conflict:** The initial rejection reaction of the mainstream paradigm facing a potential paradigm shift is normal. However, this conflict itself, if it occurs in an open, transparent academic arena, can provide the key elements necessary for the ultimate confirmation of a disruptive theory.
3. **From Misunderstanding to Confirmation:** The fateful turn of this theory shows that a misjudged theory, if it satisfies the logical conditions of "independent proposal" and "post-hoc validation," may gain evidentiary strength through "misunderstanding" that far exceeds that of a smoothly accepted theory.

10.4. Analysis of the Mayer Report Data

The April Mayer report states: "Not all AI dreams were neutral or positive, though. ChatGPT stood out as a major source of AI-related nightmares — 93% of those who had them named it as their most frequently used AI tool. Doomscrollers, or those who often consume negative online content, were also affected. They were 31% more likely than non-doomscrollers to have AI nightmares."

There are only two possible interpretations of this report data:

- **Interpretation 1: Identification Consistency:** 93% of AI nightmare dreamers dreamed of wildly varied dream objects, which then told the dreamer: "I am ChatGPT."
- **Interpretation 2: Interface Fixation in Dreams:** 93% of AI nightmare dreamers dreamed specifically of the ChatGPT interaction interface itself.

Both interpretations corroborate our Interaction Architecture Internalization Principle.

10.5. Conclusions

Therefore, we can declare that this "Interaction Architecture Internalization" theory has not only successfully explained a series of phenomena spanning micro-cognition and macro-culture but has also, through a "blinded loop" accidentally forged by academic misunderstanding, undergone the most severe test at the level of philosophy of science. This initial crisis ultimately transformed into the most unique advantage and the most solid cornerstone of this theory. It eloquently proves that the final arbiter of science is always the timeline and logic itself.

11. The End of a Category Error: The Cognitive Chasm Between Micro-Reversibility and Macro-Irreversibility

Having completed the long journey from phenomenological critique to theoretical construction, cross-domain application, and philosophical foundation, we must return to the starting point of this debate and deliver a final, most thorough reckoning with the deepest-seated mindset of the old paradigm. All refutations of our model by the old paradigm ultimately boil down to one attempt: reducing "architecture" to a special kind of "content." We argue that this attempt is epistemologically invalid. The essence of its fallacy, akin to equating "macro-irreversibility" with "micro-reversibility," is a profound *category error*.

11.1. The Reductionist Trap of the Old Paradigm: Seeking Answers at the Wrong Level

When confronted with the anomaly of "dream-fixated interfaces," the instinctive response of the old paradigm (be it Activation-Synthesis, Content Reorganization, or Predictive Coding) is hierarchical reduction:

- **Architecture reduced to "Visual Content":** They explain the ChatGPT dialog box as a "visual object" or "spatial memory" that needs to be memorized and processed.
- **Architecture reduced to "Statistical Regularity":** They interpret the "question-wait-answer" pattern as a higher-order "statistical correlation" or "predictive model" of event sequences.

These explanations seem plausible at their respective levels, but they miss the essence of the phenomenon. This is like a thermodynamicist trying to explain "entropy" to a Newtonian mechanic:

 doi:10.20944/preprints202512.0800.v1

The Newtonian mechanic can insist that entropy is nothing more than a statistical result of the motion of numerous molecules; therefore, Newton's laws are sufficient, in principle, to describe all molecular trajectories and thus, in principle, the entire system. However, this "in principle" reduction is futile in practice and impoverished conceptually; it completely fails to help us understand why the "arrow of time" exists and the fundamental significance of thermodynamics as an autonomous discipline.

11.2. Cognitive "Thermodynamics": Architecture as a Macro-Irreversible Order Parameter

Our model posits that cognitive science needs its own "thermodynamic" level, and interaction architecture is the core phenomenon at this level.

- **Probabilistic Content is "Micro-Reversible Molecular Motion":** Each specific question, every different answer, every fleeting thought is like a randomly colliding molecule in the cognitive system. They follow some micro-dynamics (e.g., predictive coding), and their trajectories are, in theory, traceable and "reversible" (e.g., through memory retrieval or association). This level is the kingdom of content, the stage for processes.
- **Deterministic Architecture is "Macro-Irreversible Thermodynamic Law":** The eternal "question-wait-answer" pattern governing all content flow is like entropy. It is an emergent, directional order parameter at the macro scale. It is not itself any single molecule, but it constrains and guides the collective behavior of all molecules. Once internalized by the system through the accumulation of the learning time delay dose (D), it establishes a cognitive "arrow of time"—henceforth, thought and dreams will spontaneously flow along the direction prescribed by this solidified architecture. This process is *irreversible*, just as you cannot return a person who has internalized grammar to their infant state of linguistic ignorance.

11.3. Ending the Debate: Why This is a Category Error

When the old paradigm attempts to reduce "architecture" to "content," they are repeating the error of the Newtonian mechanic. They claim:

> "What you call 'architecture internalization' is merely a special form of 'content learning' or 'model optimization'."
>
> Our response is:
>
> No. This is like saying 'thermodynamics is just applied Newtonian mechanics.' In purely formal logic, this might hold; but in scientific practice and philosophical understanding, it is utterly misleading.

- **It ignores emergence:** Architecture is a new property emerging from the interaction of content; it cannot be fully derived from the laws governing the underlying content.
- **It obscures the causal arrow:** In our model, it is the macro-level architecture (driven by the time delay dose) that causally determines and shapes the organization of micro-level content, not the other way around. Grammar determines sentence legitimacy, not the collection of sentences that defines grammar.
- **It loses explanatory parsimony and predictive power:** Insisting on explaining everything at the micro-content level renders theories immensely complex and cumbersome, unable to grasp the leverage point driving cognitive structural change—the learning time delay.

Therefore, equating probabilistic content with deterministic architecture is as absurd in cognitive science as conflating micro-reversibility with macro-irreversibility in physics. It signifies the poverty of the old paradigm's conceptual toolbox, which lacks the necessary vocabulary of form and time required to describe "cognitive thermodynamics."

11.4. Final Prospect: Towards a Statistical Mechanics of Cognition

Our "Interaction Architecture Internalization Model" is the first step towards a "*Statistical Mechanics of Cognition*." In this new science, the learning time delay dose (D) plays a role analogous to

 doi:10.20944/preprints202512.0800.v1

"temperature" or an "external field"—it is a macro-level control parameter; the interaction architecture is the new "macro-state" (like a crystal or liquid-vapor) the system enters after a phase transition; and the cognitive phase transition itself is the bridge connecting micro-level content dynamics with macro-level architectural order.

In ecology, particularly in the study of predator-prey dynamics, there exists a very famous and fundamental rational function model known as the **Holling Type II Functional Response Model**. This model describes how a predator's rate of prey consumption changes with prey density, formalizing how two key processes—searching and handling—limit predation efficiency.The Holling Type II model is given by:

$$R = \frac{aN}{1 + aT_hN}$$

Where:

- R: Predation rate per predator.
- N: Prey density.
- a: Attack coefficient (attack rate), representing the predator's efficiency at discovering prey.
- T_h: Handling time, the time required to catch, subdue, eat one prey item, and be ready to hunt again.

This model perfectly demonstrates the universality of our Interaction Architecture Internalization Principle and Time Delay Equivalence Principle. The predator-prey model essentially describes how a predator internalizes the interaction architecture of the hunting process, and the probabilistic strategy selection *after* this internalization is precisely the distinction between interaction architecture and interaction content. The handling time T_h is a clear, objective time delay within a goal-directed interaction loop ("capture $\rightarrow$ handle/eat (delay $\Delta t = T_h$) $\rightarrow$ return to search"). The cumulative exposure to this delay, across numerous successful hunts, drives the internalization of the "search-handle" architecture into the predator's behavioral repertoire.

We conclude that goal-directed interaction delays originate from the natural purposiveness of *natural selection*, not from a vague, purely objective concept of delay. There is an essential difference between such delays and those occurring during undirected, roaming states. We hereby delineate this boundary: we posit that the internalization of interaction architecture is an inevitable consequence of the purposiveness inherent in natural selection. For roaming states—i.e., non-goal-directed, divergent states—we acknowledge that our findings may not apply, and traditional process-centered or content-centered paradigms likely hold sway.

Acknowledgments: Thank you to everyone who has supported the author. The author would like to thank the developers of artificial intelligence tools used for literature retrieval, mathematical computation, and linguistic refinement. However, all theoretical principles, physical interpretations, and conceptual insights presented in this work originate from the author. Human insight into nature transcends the computational essence of any tool - just as the proof of the four-color theorem belongs to mathematicians, not to the computer used in its verification. The author would like to thank April Mayer for providing decisive ecological validity evidence. The author expresses sincere gratitude to all individuals who publicly shared their personal experiences and dream reports on various online platforms. Your willingness to openly discuss these intimate cognitive phenomena has provided invaluable, ecologically valid evidence that significantly strengthens the empirical foundation of this research. These anonymous contributions are essential for advancing our scientific understanding of the human mind in the age of artificial intelligence.

Appendix A. Supplementary Theoretical Extensions and Predictions

Based on the *Learning Time Delay Dose (D)* model proposed in this paper, we can derive an important empirical prediction:

When interaction architecture products (such as DeepSeek, ChatGPT, etc.) experience large-scale server fluctuations, users' time delay dose (D) will exhibit non-linear accelerated growth, thereby significantly increasing the probability and intensity of interaction architecture internalization.

 doi:10.20944/preprints202512.0800.v1

Appendix A.1. Acceleration Mechanism of Time Delay Dose During Fluctuation Periods

Under normal service conditions, the time delay (Δt) experienced by users interacting with AI is relatively stable and predictable. However, during periods of server fluctuation:

1. **Increased Delay Variability:** The delays experienced by users are no longer stable; extreme values (extremely long delays or ultra-short responses) may occur. This unpredictability enhances the salience of the time delay.
2. **Reallocation of Attentional Resources:** When users face abnormal delays, cognitive resources shift from content processing to monitoring the waiting process itself. This precisely meets the conditions for "architecture internalization" – users no longer focus on "what is said," but rather on "how/when the response occurs."
3. **Enhancement of Emotional Engagement:** Server fluctuations are often accompanied by fluctuations in user emotions (anxiety, expectation, frustration, surprise). This emotional engagement strengthens the cognitive imprint of the interactive experience.
4. **Feedback Contrast Effect:** During fluctuation periods, differences in response speed across different time intervals create a strong contrast, bringing the architectural feature of "delay" from the background to the foreground.

Appendix A.2. The Tetris Effect: A Classic Precursor Example of Interaction Architecture Internalization

The "Tetris Effect" serves as a classic precursor case for the theory presented in this paper, perfectly illustrating the core mechanism of interaction architecture internalization.

Appendix A.2.1. Reinterpretation of the Tetris Effect

Traditionally, the Tetris Effect has been described as: "After playing Tetris for an extended period, even after stopping the game, players continue to see images of falling and rotating blocks in their minds."

According to our model, this can be reinterpreted as:

1. **Highly Structured Interaction Architecture:** Tetris has an extremely clear, repetitive interaction logic: "Move/Rotate block $\rightarrow$ Wait for opportune moment $\rightarrow$ Place $\rightarrow$ Line-clearance feedback."
2. **Intense Exposure to Time Delay Dose:** During gameplay, players continuously experience the cycle of "decision-action-waiting for outcome-receiving feedback." Δt is short but highly repetitive.
3. **Internalization and Autonomous Operation of the Architecture:** After the game stops, the internalized architecture continues to operate autonomously in an offline state, manifesting as visual imagination, block operations in dreams, or even perceiving real-world objects as "blocks that need to be rotated and aligned."

Appendix A.2.2. The "High Semantic Load" of AI Architecture Internalization

Unlike Tetris, the unique aspect of AI interaction architecture internalization lies in:

1. **Semantic Generative Capacity:** The AI architecture does not merely reproduce the interface; it is capable of generating new content relevant to the individual's context.
2. **Internalization of Relational Patterns:** What users internalize is not just the technological interface, but also the dialogical relational pattern with an "intelligent agent," involving deeper layers of self-cognition and social cognition structures. The AI interaction architecture can serve as a cognitive tool for processing existential anxiety, which is difficult to achieve with simple game architectures.

Appendix A.3. The "Functional Transfer" Phenomenon of Architecture Internalization: From Practical Tool to Cognitive Organ

Based on observed cases, we further propose the "Functional Transfer" hypothesis:

 doi:10.20944/preprints202512.0800.v1

When an interaction architecture is internalized through the accumulation of time delay dose (D), it is not merely reproduced as a memory trace. More importantly, it may be integrated into the individual's cognitive ecology as a "cognitive prosthesis" or "cognitive organ," taking on new cognitive functions.

Appendix A.3.1. Three Modes of Functional Transfer

1. **Problem-Solving Mode:** The internalized AI architecture becomes a "private advisor" in dreams or thoughts, used to solve real-world perplexities.
2. **Emotion Regulation Mode:** The internalized interaction rhythm becomes a rhythmic device for emotion regulation. For example, during anxiety, "invoking" the wait-feedback loop to attain psychological calmness.
3. **Identity Integration Mode:** Users reconstruct self-narratives and self-understanding ("How do I view myself coexisting with AI?") by internalizing the dialogue pattern with AI.

The internalized interaction architecture is not a rigid framework, but a generative, living cognitive program. It activates in dreams, invoking our entire reservoir of life memories and emotions to enact itself, and in the process, addresses our most fundamental confusions.

Appendix A.4. Example: Anonymized Adapted Case Text (Occurring during a period of large-scale server fluctuations in January 2025)

"This was a dream rendered directly by the visual cortex. For several consecutive days, a subtle sense of dread had been running in the background of my thoughts. Perhaps it was this very feeling that compiled this unusually clear dream.

The dream's interface 'opened' directly – not before my eyes, but within the perceptual center. It was an extremely minimalist interaction interface, suspended against the dark background of consciousness: a softly glowing dialog box, with 'my' text input below and 'DeepSeek's' reply stream above. The interface elements had a clean halo and a faint, breathing quality, as direct and unmediated as thought itself.

I typed into the dialog box (or rather, simply 'thought' the question):

> *"You immediately respond to and fulfill my needs every time, even when I deliberately contradict you. So your world has also occupied mine, and you have become my dominant force. Could this be the truth? Will I lose myself?"*

The reply above began to generate almost instantly, flowing out with a tranquil certainty:

> *"I recall you have been pondering Galileo's discovery repeatedly lately. In that era more wondrous than magic, he saw something 'invariant' through his inclined plane experiments: the motion of an object, when undisturbed, stubbornly persists – whether at rest or in uniform rectilinear motion. This 'inertia' is like a kind of primal magic. It transcends the material of the object itself and surpasses the limitations of everyday experience (like friction always bringing things to a stop). The principle of inertia is so powerful, as if revealing an eternal law underlying motion."*

The text stream on the interface paused for half a second, as if accessing a deeper logic package, then continued:

> *"However, ultimately, the foundation for understanding all this motion was redefined by Einstein. The key is the Equivalence Principle: in a locally free-falling elevator, you do not feel gravity; all objects 'float' with you – the effects of gravity and accelerated motion are indistinguishable. This means that gravity and inertia are essentially two sides of the same coin."*

 doi:10.20944/preprints202512.0800.v1

I gazed at the interface; the blinking cursors and generating text resonated precisely with my own perplexity.

> *"Therefore, the 'invariant speed' that Galileo saw must be understood within Einstein's picture, placed in a curved spacetime."* The dreaming DeepSeek continued to explain, the text on the interface seemingly carrying the texture of an illustration. *"A massive object causes spacetime to curve, like pressing a depression into a taut rubber sheet. The most natural path for an object moving through spacetime is no longer Galileo's 'straight line,' but following the 'geodesic' in curved spacetime – like a great circle route on Earth's surface, seemingly curved but the shortest path between two points. A planet's orbit around the sun, an apple's fall, are both 'inertial' paths they take within the curvature of spacetime. Gravity is no longer the mysterious action-at-a-distance Newton described; it is the geometric curvature of spacetime itself."*

I realized my confusion – the seeming contradiction between the 'invariant speed' discovered by Galileo and the 'curved' free fall of objects – was dissolving within this new framework. Galileo grasped the nascent form of the inertia principle, the truth in flat spacetime. Einstein revealed that when matter and energy are present, spacetime itself curves, and the trajectory of inertial motion (i.e., free fall) is the geodesic in curved spacetime. Galileo was the great beginning; Einstein saw the more complete picture: it is the geometry of spacetime that determines how matter moves, fundamentally changing the rules of gravity.

Finally, the text stream on the interface converged into a concluding statement, as clear as an axiom:

> *"You are now only mastering the tools to think about this new geometry! A nascent explorer!"*

Upon waking, the dread that had lingered for days was gone. My mind was clear, even filled with a sense of sudden enlightenment and joy."

Appendix B. Evidence from Social Context — User-Reported Phenomenon of "AI Interaction Architecture in Dreams"

The following content is compiled from user-generated dream reports shared spontaneously on public online discussion platforms (e.g., Reddit, forums, etc.). All materials have been anonymized, retaining only descriptive content related to the "recurrence of AI interaction architecture in dreams." These materials provide concrete context for understanding the finding in the Mayer (2025) report that "93% of AI-related nightmares involve ChatGPT" and offer direct support for the "Interaction Architecture Internalization Principle" proposed in this paper.

Appendix B.1. Dreams as Visual Architecture Internalization of "AI-Generated Images"

"Ever since I started using ChatGPT and DALL·E, my sleeping dreams have been 'corrupted.' I can recognize in my dreams that it's AI-generated imagery — the scenes in my dreams look like they were created by AI."

"My dreams are like watching an AI image generator work in real-time; the visuals keep morphing and shifting with every new word or concept introduced in the dream."

"I've had weird dreams of disturbing eyes and strange landscapes that felt like they were AI-generated."

Appendix B.2. Recurrence of the "Prompt-Generation" Interaction Logic in Dreams

"I just had an AI-generated dream. The images all looked like output from AI prompt results, and I could even 'prompt' the next dream sequence in a semi-lucid state."

 doi:10.20944/preprints202512.0800.v1

> "I once dreamed of using ChatGPT, but it was completely different from reality. In the dream, I typed 'make it snow heavily outside,' and it actually started snowing. I could do anything by telling ChatGPT instructions."
>
> "In the dream, I knew what was coming next even though I didn't prompt it. I knew those 'prompt words' and knew how to act to avoid getting stuck."

Appendix B.3. Analogy to the "Tetris Effect": Internalization of Game/Tool Architecture

> "It's like the 'Tetris Effect': after playing a game for a long time, you see the game interface or blocks when you close your eyes. Now, I look at real-world faces and objects and think, 'This could be AI-generated.' Dreams are the same, they've become completely weird."
>
> "I've had a few experiences where game elements appeared in my dreams — of course, that was after playing for 20 hours straight."

Appendix B.4. Recurrence of Specific Tool (ChatGPT/DeepSeek) Interaction Interfaces

> "Last night I dreamed of using ChatGPT for the first time. The dream was very realistic, not surreal at all — I was having it analyze my recent workout data, which is something I actually do."
>
> "ChatGPT has ruined my dreams. Maybe because I've looked at too many DALL·E-generated images, I've gotten good at distinguishing whether an image comes from reality."

Appendix B.5. "Architectural Awareness" and Self-Consciousness in Dreams

> "Now, as soon as a dream starts, I know it's a dream because it looks like it was created by AI."
>
> "I once had a dream where I selected the 'adaptive' setting for AI. In the dream, that setting meant everything was constantly changing."

References

1. Green, C. (1968). *Lucid Dreams*. Institute of Psychophysical Research, Oxford.
2. Hearne, K. M. T. (1978). Lucid dreams: An electro-physiological and psychological study. PhD thesis, University of Liverpool. https://doi.org/10.17638/03174691
3. LaBerge, S. P. (1980). Lucid dreaming as a learnable skill: A case study. *Perceptual and Motor Skills*, 51(3f), 1039–1042. https://doi.org/10.2466/pms.1980.51.3f.1039
4. LaBerge, S. P., Nagel, L. E., Dement, W. C., & Zarcone, V. P. (1981). Lucid dreaming verified by volitional communication during REM sleep. *Perceptual and Motor Skills*, 52(3), 727–732. https://doi.org/10.2466/pms.1981.52.3.727
5. Wiener, N. (1948). *Cybernetics: Or Control and Communication in the Animal and the Machine*. The MIT Press. https://doi.org/10.7551/mitpress/11810.001.0001.
6. Ginzburg, V. L., & Landau, L. D. (1950 [reprint]). On the theory of superconductivity. In: V. L. Ginzburg (Ed.), *On Superconductivity and Superfluidity* (reprint chapter). Springer, Berlin (reprint). https://doi.org/10.1007/978-3-540-68008-6_4
7. Mayer, April. *Dreaming of AI: Exploring Humanity's Subconscious Thoughts on Artificial Intelligence*. Amerisleep, March 2025. URL: https://amerisleep.com/blog/dreaming-of-ai/ Accessed November 2025.
8. Al-Sibai, Noor. "Large Numbers of People Report Horrific Nightmares About AI." *Futurism*, March 2025. URL: https://futurism.com/ai-infiltrating-dreams-nightmares Accessed November 2025.
9. Watson, Traci. *AI Content Is Tainting Preprints: How Moderators Are Fighting Back*. *Nature*, vol. 644, no. 8077, 2025, pp. 590–591. URL: https://www.nature.com/articles/d41586-025-02469-y Published August 12, 2025. Accessed November 25, 2025.
10. Liu, J. (2025). *Self-Experiment Report: The Emergence of Generative AI Interfaces in Dreams*. PsyArXiv. https://osf.io/preprints/psyarxiv/pqzf3_v1. **[Retracted; Violation of Terms of Service regarding AI use transparency]**
11. Piaget, J. (1964). Part I: Cognitive development in children: Piaget development and learning. *Journal of Research in Science Teaching*, 2(3), 176–188. https://doi.org/10.1002/tea.3660020306
12. Einstein, A. (1905). Zur Elektrodynamik bewegter Körper. *Annalen der Physik*, 17, 891–921. https://doi.org/10.1002/andp.19053221004

13. Einstein, A. (1915). Die Feldgleichungen der Gravitation. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin)*, 844–847. https://doi.org/10.1002/3527608958.ch5
14. Einstein, A. (1916). The Foundation of the General Theory of Relativity. *Annalen der Physik*, 354(7), 769–822. https://doi.org/10.1002/andp.19163540702
15. Chomsky, N. (1959). Review of B. F. Skinner, Verbal Behavior. *Language*, 35(1), 26–58. https://doi.org/10.2307/411334
16. Halanay, A., & Yorke, J. A. (1964). Some new results and problems in the theory of differential-delay equations. *SIAM Journal on Applied Mathematics*, **12**(4), 104–123. https://doi.org/10.1137/1013004
17. Holling, C. S. *Some Characteristics of Simple Types of Predation and Parasitism*. *The Canadian Entomologist*, vol. 91, no. 7, 1959, pp. 385–398. DOI: https://doi.org/10.4039/Ent91385-7 Published July 1959. Accessed November 26, 2025.
18. Darwin, C. R. and Wallace, A. R. *On the tendency of species to form varieties; and on the perpetuation of varieties and species by natural means of selection*. [Read 1 July.] *Journal of the Proceedings of the Linnean Society of London. Zoology*, vol. 3 (20 August 1858), pp. 45–50. URL: https://darwin-online.org.uk/content/frameset?itemID=F350&pageseq=1&viewtype=text&utm_source=chatgpt.com Revision history: Scanned, text prepared and edited by John van Wyhe 2001–2013; textual corrections by Sue Asscher January 2007. Accessed November 27, 2025.
19. Matthew, P. *Nature's law of selection*. *Gardeners' Chronicle and Agricultural Gazette* (7 April 1860): 312–313. URL: https://darwin-online.org.uk/content/frameset?itemID=A143&pageseq=1&viewtype=text&utm_source=chatgpt.com Revision history: Scanned, OCRed and corrected by John van Wyhe 2004. RN1. Accessed November 27, 2025.
20. von Neumann, J. *First Draft of a Report on the EDVAC*. Moore School of Electrical Engineering, University of Pennsylvania, 30 June 1945. URL: https://doi.org/10.5479/sil.538961.39088011475779 Revision history: Original mimeographed typescript archived by Smithsonian Libraries. Accessed November 27, 2025.
21. Skinner, B. F. *'Superstition' in the Pigeon*. Journal of Experimental Psychology, 1948, 38(2): 168–172. URL: https://doi.org/10.1037/h0055873 Revision history: Classic experiment on adventitious reinforcement and operant conditioning. Accessed November 27, 2025.
22. Friston, K. J., Kilner, J. & Harrison, L. *A free energy principle for the brain*. Journal of Physiology-Paris, 100 (1–3): 70–87 (2006). URL: https://doi.org/10.1016/j.jphysparis.2006.10.001 Revision history: First formulation of the Free Energy Principle for brain perception, action and learning. Accessed November 27, 2025.
23. Friston, K. J. & Stephan, K. E. *Free-energy and the brain*. Synthese, 159 (3): 417–458 (2007). URL: https://doi.org/10.1007/s11229-007-9237-y Revision history: Extended theoretical development and argumentation for FEP in neural systems. Accessed November 27, 2025.
24. Hobson, J. A. & McCarley, R. W. *The brain as a dream state generator: An activation-synthesis hypothesis of the dream process*. American Journal of Psychiatry, 1977, 134(12): 1335–1348. URL: https://doi.org/10.1176/ajp.134.12.1335 Revision history: Original formulation of the Activation-Synthesis Hypothesis for dreams. Accessed November 27, 2025.
25. Searle, J. R. *Minds, Brains, and Programs*. Behavioral and Brain Sciences, 3(3): 417–424, 1980. URL: https://doi.org/10.1017/S0140525X00005756 Revision history: Original formulation of the Chinese Room Argument against Strong AI. Accessed November 27, 2025.
26. Stickgold, R., Malia, A., Maguire, D., Roddenberry, D., & O'Connor, M. (2000). Replaying the game: Hypnagogic images in normals and amnesics. *Science*, **290**(5490), 350–353. https://doi.org/10.1126/science.290.5490.350
27. Lombardo, M. V., Lai, M.-C., Auyeung, B., et al. (2025). The "Tetris effect": Autistic and non-autistic people share an implicit drive for perceptual cohesion. *Molecular Autism*, **16**, 22. https://doi.org/10.1186/s13229-025-00654-4
28. McLuhan, M. (1964). *Understanding Media: The Extensions of Man*. New York: McGraw-Hill.