The Information Integration Model (IIM): Consciousness, Self, and Anxiety Under Finite Causal Constraints

Yu Qinghong — Independent Researcher

Abstract

This paper develops a theoretical framework that attempts to explain why consciousness, the sense of self, and anxiety appear in the kind of universe we inhabit—one in which information does not arrive instantly but must travel through space and neural tissue at finite speed. What seems like an abstract physical limitation turns out to matter deeply for the structure of experience. If perception always lags behind the world, any organism that must act in real time needs a way to compensate for the delay. I argue that consciousness emerges as the system responsible for integrating predictions with delayed sensory evidence; that the self is the long-term record of these predictive efforts; and that anxiety arises when the demands placed on prediction and integration exceed the system's limits.

To explore this idea, I introduce two conceptual quantities: **Integration Efficiency (η_c)**, which reflects how effectively prediction errors are resolved within behaviorally meaningful time windows, and **Self-Entropy (H_{self})**, which describes the disorder or instability within the organism's self-model based on unresolved, self-relevant errors accumulated over time. The paper proposes a simplified mathematical structure, outlines several empirical predictions, and sketches behavioral, narrative, and artificial-system experiments that could test the model. While the proposal remains

speculative, it aims to offer a coherent starting point for linking subjective experience to universal physical constraints.

Contents

1. Introduction

The everyday fact that light, sound, and neural signals require time to reach us seems too trivial to matter. Yet this delay shapes everything an organism can know or do. When I stare at a bright light and close my eyes, the afterimage lingers in a way that feels slightly "out of sync" with the present. When I am anxious, time itself feels strained—too fast, too thin, or somehow too close to slipping away. These experiences are different, yet both reveal the same fact: the mind is not receiving the world in real time. It is always catching up.

For many years I found myself returning to this simple point. If the present moment is always a little out of reach, then the mind must constantly make up the difference. It cannot wait for perfect information before acting; it must guess, anticipate, and revise. This thought gradually became the starting point for the model I develop here.

Contemporary theories of mind often begin with higher-level concepts such as representation, inference, or integration. Predictive processing theories describe the brain as a machine that continually anticipates incoming signals and corrects its predictions when they are wrong. Research on time perception shows that human experience depends on the way information is grouped within tight temporal windows. Philosophers have argued that the self is not a mysterious inner entity but a model constructed by the brain. And theorists have attempted to ground consciousness in the structure of causal influence within the brain.

Each of these ideas contributes something important. Yet they typically treat temporal delay —the simple fact that information takes time to move—as an implementation detail rather than a first principle. Predictive processing often begins with the assumption that prediction is useful, but not with the deeper question of why prediction is indispensable for any organism that must move and react in a world governed by finite causal speed.

This paper attempts to place that physical constraint at the foundation of the discussion. If no signal arrives immediately, an organism that wants to survive must function in a world it can never perceive exactly as it is. It must predict the near future and then integrate incoming evidence that confirms or contradicts those predictions. Conscious experience, on this view, is not a passive display but the work of maintaining this integration in real time.

The self, in turn, becomes the accumulation of predictive successes and failures—a record that guides how future predictions are shaped. And anxiety becomes the experiential sign of a system falling behind: prediction demands rise, integration becomes strained, and unresolved errors accumulate into an unstable self-model.

> *The aim of this paper is modest. I do not claim to settle the nature of consciousness or to offer a fully polished computational theory. My intention is to sketch a framework that links subjective life to constraints that apply to any information-processing system in a causal universe. I hope this framework is simple enough to be tested, flexible enough to be extended, and clear enough to be criticized.*

2. Related Work

The model developed in this paper sits at the intersection of several approaches to mind and cognition. Rather than offering a broad survey, I highlight the specific lines of research that shape the framework presented here. The intention is not to place the model within any single tradition but to show how it grows out of questions that these traditions have not fully resolved.

2.1 Predictive Processing and the Demands of Anticipation

Predictive processing has become one of the most influential frameworks in cognitive science. It describes the mind as a system that continuously anticipates incoming signals and adjusts its expectations when those predictions fail. Much of this work presents prediction as a computationally efficient strategy. Yet the deeper problem—why prediction is indispensable in the first place—is usually taken for granted. In a universe where signals propagate with finite speed, prediction is not merely useful but necessary. This physical necessity serves as one of the foundations of the model developed here.

2.2 Temporal Integration and the Structure of the Present

Research on time perception shows that experience is not instantaneous but knitted together across short windows of time. The work of David Eagleman and others demonstrates that simultaneity, delay, and continuity are constructed through post hoc integration. These findings suggest that the brain must process events in discrete temporal packets rather than as a perfectly continuous stream. The IIM uses this idea to motivate the concept of **Integration Efficiency**, which concerns how much prediction error can be resolved within these temporally bounded windows.

2.3 The Self as a Constructed Model

Philosophers like Thomas Metzinger argue that the self is not a thing but a process: a model that the brain builds to stabilize action and perception across time. This idea aligns with the IIM's notion that the self is an accumulation of resolved and unresolved predictive events. What the present model adds is an explicit role for temporal delay: the self emerges as the long-term memory of how prediction has bridged the constant gap between the organism and the world.

2.4 Anxiety as Anticipatory Overload

The psychological literature on anxiety often describes it as a state in which the mind is pulled too far into the future while losing grip on the immediate present. Some authors note that anxious individuals expend cognitive resources imagining possibilities faster than they can be evaluated. These descriptions map naturally onto the IIM's view of anxiety as a condition in which predictive demand exceeds the system's ability to integrate incoming evidence in time, leading to unresolved errors that accumulate into an unstable self-model.

2.5 Integrated Information and Causal Organization

Integrated Information Theory proposes that consciousness corresponds to the amount of integrated causal structure in a system. While the IIM diverges from IIT in both motivation and method, the two share an interest in integration. The IIM does not attempt to quantify consciousness strictly through causal topology; instead, it focuses on how well a system integrates prediction with delayed sensation in real time. The temporal dimension here is central, not secondary.

2.6 The Gap in the Literature

Taken together, these approaches offer valuable tools but leave an important question unanswered: what unifies prediction, the self, and affective states like anxiety under a single principle? The IIM attempts to answer this by grounding all three in the unavoidable fact that information arrives late. This physical constraint forces organisms to predict, forces predictions to be integrated with delayed feedback, and forces the history of these operations to accumulate into a model of the self.

3. Theoretical Foundations

The starting point of this model is a simple physical observation: **information takes time to move**. Light travels at a finite speed, sound travels slower, and neural signals travel slower still. Every perceptual system therefore faces a delay between what happens in the world and when that event becomes available to be processed. This fact is so familiar that it is usually ignored, but if taken seriously, it has surprisingly deep consequences.

An organism that must move, avoid threats, and interact with other agents cannot afford to wait for perfect information. It must act in a world it cannot perceive exactly as it is, but only as it was a moment ago. It must fill the gap between the delayed world and the unfolding present. In this sense, prediction is not a convenient feature of biological cognition; it is a requirement forced by physics.

3.1 Finite Delay as the Central Constraint

Consider a simple example. If an object is thrown toward me, the light reflecting from it takes time to reach my eyes, and the neural signals that process that image take even longer to reach the motor systems that move my body. By the time I perceive the object's position, it has already moved. To catch it, my mind must estimate where the object will be, not where it appears to be.

This example generalizes. At every moment, the organism is acting on stale information. Even the sense of touch involves delays measured in tens or hundreds of milliseconds. In more abstract domains—social interactions, expectations about others' intentions—the delays are even more complicated, involving memory and inference.

> *From this perspective, one can think of **life as a continuous race between delay and prediction**. If delay wins, the organism falls behind and fails to coordinate with the world. If prediction wins, the organism keeps pace.*

3.2 Prediction as a Forced Strategy

Because of these delays, organisms have no choice but to predict the near future. Without prediction, perception would always be too late to be useful. Prediction bridges the gap between the sensed past and the actionable present.

This idea is not new. Theories of anticipatory behavior go back at least to Helmholtz. More recently, the predictive processing literature has described the brain as a system that continually generates hypotheses and evaluates them against incoming signals. But what is often overlooked is that these theories assume prediction as a functional preference rather than a physical necessity. The IIM emphasizes that **prediction is the organism’s only viable response to finite causal speed**.

3.3 Integration as the Second Necessity

Prediction alone is not sufficient. Predictions must be tested against the sensory evidence that finally arrives, despite being delayed. This demands a process of **integration**—the combination of prior expectations with new information to evaluate how well the prediction succeeded.

Integration is not a single mechanism but a broad term for all the processes that align prediction with reality. These include error detection, sensory updating, attentional selection, and the reorganization of internal models. What matters for the present theory is that these operations must be completed **within limited windows of time**. If integration falls behind, experience becomes fragmented, and action loses coherence.

Psychologists studying time consciousness have long recognized these short windows. Some describe them as the "specious present"—a temporally extended moment within which the mind groups events into a coherent whole. Work by David Eagleman and others has shown that these windows are neither instantaneous nor fixed, but they serve as functional constraints on how quickly the mind can update itself.

3.4 The Accumulation of Predictive History

When predictions succeed or fail, they leave traces in the system. These traces accumulate into a record of what the organism can expect of itself and the world. Over time, this record becomes the basis of the self: a structure of tendencies, expectations, memories, and habitual interpretations.

This view of the self aligns with philosophical accounts that describe the self as a model rather than a metaphysical entity. But the IIM adds a specific mechanism: **the stability of the self depends on the balance between prediction and integration across time**.

If predictions are consistently resolved, the self-model becomes coherent, stable, and reliable. If predictions fail faster than they can be resolved, unresolved errors accumulate, and the self-model becomes unstable or fragmented.

3.5 The Emergence of Anxiety

Within this framework, anxiety appears as a natural consequence of the system falling behind. When integration cannot keep pace with predictive demand, the mind faces an excess of unresolved error. These unresolved errors accumulate not only at the moment they occur but also within the long-term record that constitutes the self.

This creates a feedback loop:

1. Delayed information demands prediction.

2. Prediction demands integration.
3. If integration is insufficient, unresolved errors accumulate.
4. These accumulated errors destabilize the self-model.
5. A destabilized self-model increases uncertainty, raising predictive demand even further.

The subjective feeling of anxiety may be the phenomenological signature of this loop becoming unstable. When prediction overextends itself and integration cannot resolve the resulting error, the mind begins to experience time as thin, fragmented, or accelerated. The future feels too close; the present feels too narrow; the self feels unreliable.

3.6 Summary of the Foundations

To summarize, the IIM is built on three unavoidable facts:

1. **Information arrives late.**
2. **Organisms must predict to compensate for delay.**
3. **Predictions must be integrated with delayed input to remain accurate.**

From these facts follow three experiential phenomena:

- **Consciousness** as the ongoing work of integrating prediction and delayed sensation.
- **The self** as the historical accumulation of this work.
- **Anxiety** as the sign that predictive demand exceeds integrative capacity.

This foundation motivates the formal model introduced in the next section.

4. The Information Integration Model (IIM)

The theoretical foundations described in the previous section lead to a simple idea: an organism must continuously reconcile two streams of information—its predictions about the near future and the sensory evidence that arrives too late. Consciousness, on this view, is not a static state but the ongoing activity of managing this reconciliation. The Information Integration Model (IIM) attempts to make this idea explicit by introducing two conceptual quantities:

1. **Integration Efficiency (η_c)** – describing the real-time performance of the prediction–integration cycle.
2. **Self-Entropy (H_{self})** – describing the long-term stability or instability of the self-model formed through accumulated predictive outcomes.

These concepts are not introduced as final theoretical entities. They are deliberately simple and meant to serve as scaffolding for empirical and philosophical refinement.

4.1 Consciousness as Integration Efficiency (η_c)

When people speak about being "present," "clear-headed," or "in sync," they are describing a subjective condition that corresponds closely to how efficiently the mind is integrating new evidence with ongoing predictions. Moments of clarity often feel unified and continuous; moments of cognitive strain feel fragmented or delayed.

In the IIM, **Integration Efficiency (η_c)** is introduced as a way of describing this moment-to-moment capacity. At a functional level, η_c captures how much prediction error the system can detect and resolve within a behaviorally meaningful time window. The relevant time scale is not arbitrary: decades of research suggest that perceptual systems bind events into windows of roughly 100–300 milliseconds. These windows appear across sensory modalities and behavioral contexts.

η_c is not proposed as a measure of consciousness itself but as the **operational core** of conscious processing. A high η_c does not guarantee reflective awareness, but consciousness in everyday experience seems to depend on a baseline level of integration efficiency. When η_c collapses—during panic, exhaustion, or neurological disturbance—experience becomes disjointed, sped up, or strangely distant.

An intuitive way to think about η_c is through a simple ratio:

- **How much predictive error arises within a temporal window?**
- **How much of that error can be resolved within the same window?**

If the system resolves most errors quickly, experience feels fluent. If errors accumulate faster than they can be handled, consciousness becomes stretched and unstable.

> *I do not claim that η_c is the final answer to the "hard problem." The claim is modest:* ***conscious experience depends on a system's capacity to close the gap between prediction and delayed sensation fast enough to sustain a coherent present.***

4.2 The Self as Self-Entropy (H_{self})

If η_c describes what the mind is doing in the moment, **Self-Entropy (H_{self})** describes what the mind has become through time.

The self, in this model, is not an inner object or metaphysical entity. It is the record of how predictions have succeeded or failed across the organism's history. Each predictive event leaves a trace:

- some predictions are confirmed,
- some are contradicted,
- some never fully resolve,
- and some resolve only under specific contexts.

These traces accumulate into what we call a personality, a worldview, or a sense of identity.

H_{self} is introduced to capture the **disorder** within this accumulated record. It reflects how predictable or unpredictable the system's own reactions and interpretations have become. When predictive outcomes are consistent and reliable, H_{self} is low. When predictive outcomes vary widely and fail to cohere, H_{self} is high.

A self with low entropy is not necessarily "better." It is simply more stable. A self with high entropy is not necessarily pathological. But an extremely high H_{self} means the organism cannot rely on its own patterns to shape future predictions. This leads to uncertainty about one's own tendencies—a condition that many clinical descriptions of anxiety, trauma, and identity disturbance capture.

In everyday terms:

- **Low H_{self}** feels like "I know how I usually respond; I can trust my expectations about myself."

- **High H_{self}** feels like "I'm not sure who I am anymore; I don't know what I will do or how I will feel."

This perspective sidesteps debates about the metaphysical reality of the self by grounding the concept in the statistical stability of predictive history.

4.3 Anxiety as a Mismatch between η_c and Predictive Demand

While η_c and H_{self} can be considered independently, their interaction becomes most important in explaining affective states—especially **anxiety**.

From the IIM perspective, anxiety emerges when the system's predictive demand (the amount of near-future inference required to navigate the situation) exceeds the available integration capacity (η_c). When this happens, prediction errors accumulate faster than they can be resolved. These unresolved errors then feed into the self-model, raising H_{self}.

This produces a loop with a characteristic structure:

1. Uncertainty or threat increases predictive demand.
2. η_c proves insufficient to resolve the resulting errors in time.
3. Unresolved errors accumulate as high H_{self} (instability of the self-model).
4. A destabilized self-model increases uncertainty about one's own responses.
5. This further increases predictive demand, which again exceeds η_c.

Subjectively, this maps onto familiar features of anxiety: racing thoughts about the future, an inability to settle in the present moment, a sense that time is accelerating or thinning, the feeling of being "not oneself," or being unreliable to oneself.

> *The IIM does not treat anxiety as a purely emotional phenomenon but as an informational imbalance: prediction is outrunning integration.*

4.4 Summary of the Model's Structure

The IIM centers on a simple but powerful pressure:

> *In a delayed universe, organisms must predict; predictions must be integrated; the quality of this integration shapes both consciousness and the self.*

The model's core propositions are:

- **Consciousness** = the ongoing work of timely integration (high η_c).
- **Self** = the long-term record of predictive outcomes (H_{self}).
- **Anxiety** = predictive demand > integrative capacity, leading to rising self-entropy.

The next section introduces formal expressions that provide clarity—not to over-engineer the model, but to make its logic explicit and testable.

5. Mathematical Formalization

The mathematical expressions introduced in this section are not meant to describe the full computational detail of the mind. Instead, they offer a minimal structure that makes the logic of the model explicit. Each equation is deliberately simple. The point is not precision but clarity: to show how prediction, integration, and self-entropy can be linked under a single formal scheme.

5.1 Representing Prediction Error Over Time

Let $e(t)$ denote the prediction error arising at time t. Because sensory evidence arrives with delay, error does not appear uniformly but in discrete pulses tied to the timing of incoming signals.

We can write the instantaneous influx of prediction error as:

$$\frac{de}{dt} = D(t) - R(t)$$

where:

- $D(t)$ is the demand created by prediction (how much the system must infer about the near future).

- $R(t)$ is the resolution process (how much error can be integrated at that moment).

When $D(t) > R(t)$, unresolved error accumulates. This is the fundamental imbalance that produces strain in the system.

5.2 Integration Efficiency (η_c)

Integration Efficiency describes the fraction of prediction error that the system manages to resolve within a time window Δt. Rather than treat η_c as a static property, it is defined relative to the current conditions.

$$\eta_c(t) = \frac{\int_t^{t+\Delta t} R(\tau)\, d\tau}{\int_t^{t+\Delta t} D(\tau)\, d\tau}$$

This expression has a direct interpretation:

- **$\eta_c \approx 1$** → the system keeps pace with prediction demands.
- **$\eta_c < 1$** → unresolved errors accumulate.
- **$\eta_c \to 0$** → integration collapses; experience becomes fragmented.

Because the time window Δt can be empirically estimated (often around 100–300 ms), η_c is not merely conceptual; it can, in principle, be measured through behavioral or neurophysiological markers.

5.3 Accumulation of Unresolved Error

Unresolved prediction error forms a pool that persists across time. Let this accumulation be $E(t)$:

$$\frac{dE}{dt} = \big(1 - \eta_c(t)\big) \cdot D(t) - \lambda E(t)$$

where:

- $\big(1 - \eta_c(t)\big) \cdot D(t)$ is the inflow of unresolved predictive error,

- $\lambda E(t)$ represents natural decay, forgetting, or successful delayed resolution.

A stable system requires:

$$\lambda > 0 \quad \text{and} \quad \eta_c \text{ not too low for extended periods.}$$

If λ fails to counterbalance unresolved accumulation, the self-model becomes progressively unstable.

5.4 Self-Entropy (H_{self})

The concept of Self-Entropy describes how disordered the system's predictive history has become. A high H_{self} means the system cannot reliably anticipate its own reactions or interpret its own signals.

Let the system categorize unresolved errors into types e_i. The probability of encountering each type depends on the relative frequency with which that unresolved error has occurred.

$$H_{\text{self}} = -\sum_i p(e_i) \log p(e_i)$$

The important point is that $p(e_i)$ is **not** fixed. It depends on the history encoded in $E(t)$. As unresolved error accumulates, the distribution of error types becomes more variable and less predictable, driving H_{self} upward.

We can express this dependence as:

$$p(e_i) = \frac{E_i(t)}{\sum_j E_j(t)}$$

where $E_i(t)$ is the portion of accumulated error belonging to category i.

Increases in H_{self} reflect an increasingly disorganized predictive landscape.

5.5 Anxiety as a Dynamical Instability

Anxiety arises when predictive demand outstrips integration capacity for long enough that $E(t)$ and H_{self} begin reinforcing each other.

We can express predictive demand as a function of uncertainty:

$$D(t) = f\big(U_{\text{world}}(t) + U_{\text{self}}(t)\big)$$

where:

- $U_{\text{world}}(t)$ is environmental uncertainty,
- $U_{\text{self}}(t)$ is the uncertainty generated by a destabilized self-model.

But since $U_{\text{self}}(t)$ increases with $H_{\text{self}}(t)$, we get:

$$D(t) = f\big(U_{\text{world}}(t) + g(H_{\text{self}}(t))\big)$$

When H_{self} increases, predictive demand increases; this reduces η_c further; unresolved error increases; and H_{self} rises again. This forms a positive feedback loop:

$$H_{\text{self}} \uparrow \;\Rightarrow\; D(t) \uparrow \;\Rightarrow\; \eta_c \downarrow \;\Rightarrow\; E(t) \uparrow \;\Rightarrow\; H_{\text{self}} \uparrow$$

Mathematically, anxiety is a regime shift—a point at which this loop destabilizes.

5.6 Conditions for Stability

The system remains stable when the following inequality holds:

$$\eta_c(t) \cdot \lambda > \frac{dD(t)}{dt}$$

This condition has a simple interpretation:

> *The system's ability to resolve and decay predictive error must outpace the speed with which new demand increases.*

When this inequality is violated for long periods, the system enters an anxious—or in extreme cases, dissociative—state.

5.7 What the Formalization Achieves

These expressions do not attempt to fully capture neural dynamics. Instead, they clarify the structure of the theory:

- **Consciousness** corresponds to the timely resolution of prediction error.
- **Selfhood** corresponds to the long-term distribution of unresolved error.
- **Anxiety** corresponds to the instability created when predictive demand exceeds integrative capacity.

The next section shows how these ideas can be tested empirically.

6. Experimental Framework

The Information Integration Model was not developed to remain purely theoretical. While the model is simple, it makes several predictions that—if correct—should be observable through behavioral measures, subjective reports, neural markers, and even artificial systems. This section outlines a set of experimental approaches intended to test the model's central claims. Each experiment is intentionally modest in design: the aim is to show how the model might be falsified or supported, not to prescribe a specific methodology.

> *The guiding principle is straightforward:* ***If consciousness depends on η_c, and if the self depends on H_{self}, then we should be able to manipulate these quantities and observe their consequences.***

6.1 Behavioral Experiments on Temporal Integration

If consciousness relies on Integration Efficiency (η_c), then tasks that strain temporal integration should produce measurable disruptions in subjective experience, accuracy, or reaction time.

6.1.1 Temporal Load Manipulation

Participants could perform rapid prediction tasks in which stimuli appear at frequencies approaching or exceeding typical perceptual binding windows (e.g., ~100-300 ms). When the rate of incoming input surpasses the system's integration capacity, the model predicts:

- increased error rates,
- subjective reports of time "speeding up" or becoming thin,
- mild dissociation-like experiences,
- decreased confidence in one's responses.

Existing research on temporal binding and attentional overload shows effects in this direction, but the IIM predicts a specific form: a drop in η_c should precede or coincide with these phenomenological disruptions.

6.1.2 Predictive Load Challenges

Participants can be asked to anticipate rapidly changing patterns. As predictive uncertainty increases, η_c should decline first, followed by subjective strain. The model predicts:

- prediction makes experience feel more stable only when η_c remains above a certain threshold,
- predictive overload destabilizes the sense of presence.

These experiments would operationalize η_c as the ratio of resolved-to-unresolved prediction error, estimated indirectly through behavioral accuracy and timing.

6.2 Self-Entropy Experiments Using Narrative Reconstruction

If H_{self} is the entropy of the self-model, then conditions that destabilize predictions about oneself should increase subjective instability and anxiety.

6.2.1 Narrative Disruption Paradigm

Participants could be asked to describe a personal pattern (“I usually react calmly in arguments”) and then be shown carefully selected counterexamples from their own prior behavior (collected beforehand). The feedback would not be judgmental but merely accurate.

The IIM predicts:

- increased variability in self-descriptions,
- lower confidence in personal traits,
- a measurable rise in H_{self} (approximated by dispersion in responses),
- and—if the effect is strong enough—an increase in predictive demand in subsequent tasks.

Crucially, individuals with already elevated anxiety should be more sensitive to such disruptions.

6.2.2 Longitudinal Error-Tracking

Over weeks, participants could log predictive successes and failures in everyday life (e.g., "I predicted X would happen, but Y happened instead"). The entropy of these records should correlate with:

- baseline anxiety,
- subjective identity stability,
- and performance on predictive tasks.

The model predicts that rising variability in this log should precede increases in anxiety, not merely accompany them.

6.3 Neural Experiments on Integration Collapse

The model implies that η_c should have neurophysiological correlates related to prediction error resolution.

6.3.1 EEG/MEG Error-Resolution Window

Using fast neuroimaging (MEG or EEG), researchers could examine the latency and amplitude of error-related signals during high-speed prediction tasks. The model predicts:

- when η_c is high → shorter latencies, more synchronized error-resolution patterns,

- when η_c is low → prolonged, disorganized error signals, reduced phase coherence.

These effects should correlate with subjective ratings of clarity or strain.

6.3.2 Induced Delay or Noise

By adding controlled delays or noise to sensory input, researchers could artificially widen the gap between prediction and input. The IIM predicts a **non-linear** decline in η_c once delays exceed typical integration windows. This would provide strong evidence that consciousness depends on time-sensitive integration, not merely on information volume.

6.4 Anxiety as a Dynamical Phenomenon

Because the model defines anxiety as a runaway feedback loop between η_c and H_{self}, the transition into anxious states should be observable as a **phase shift**.

6.4.1 Detecting the Tipping Point

By gradually increasing predictive demand in tasks and simultaneously recording physiological and subjective markers, researchers should observe:

- a stable region where η_c adjusts smoothly,
- a threshold region where small increases in demand produce disproportionate increases in unresolved error,
- and finally a regime shift into unstable performance and high-anxiety phenomenology.

This kind of sudden transition is characteristic of dynamical-system instabilities and should be empirically measurable.

6.4.2 Recovery Dynamics

If the model is correct, stabilizing η_c (e.g., via slower task pacing, grounding techniques, reduced uncertainty) should reduce **H_{self} only gradually**, because entropy reflects accumulated history. The asymmetry between destabilization (fast) and recovery (slow) is a strong prediction of the IIM.

6.5 Artificial System Experiments

Because the model is grounded in information integration rather than biological details, artificial agents can be used to test it.

6.5.1 Simulated Agents with Finite Communication Delay

Agents with delayed input channels (simulating finite causal speed) can be tested under increasing predictive load. The model predicts:

- systems that maintain a high ratio of resolved-to-unresolved prediction error behave more coherently,
- systems whose integration windows are exceeded enter unstable or "anxious" dynamics,
- the tipping point occurs when predictive demand outpaces integration capacity.

6.5.2 Self-Entropy in Artificial Memory Systems

Agents that track prediction outcomes will develop distributions of error types. By manipulating input uncertainty, researchers can observe:

- entropy increases mirroring instability,
- reduced prediction accuracy,
- and feedback loops similar to those described in human anxiety.

This would provide a non-biological test of whether the IIM describes a general property of delayed information-processing systems.

6.6 Summary

Across all methodologies—behavioral, neural, narrative, and computational—the model predicts a consistent pattern:

1. **Consciousness** depends on maintaining η_c above a threshold.
2. **Selfhood** becomes unstable when unresolved error accumulates and increases H_{self}.
3. **Anxiety** appears when predictive demand grows faster than the system's capacity to integrate.

These predictions are intentionally simple, but they offer concrete opportunities for falsification. If the model is wrong, empirical testing should expose inconsistencies quickly.

7. Implications

The Information Integration Model (IIM) reframes several long-standing questions about consciousness, selfhood, and anxiety. These implications do not rely on the mathematical formalism alone; rather, they follow from the conceptual structure developed throughout the paper. The strength of the model is that it ties together three domains that are often treated separately: the nature of conscious experience, the construction of the self, and the emergence of affective instability. Here I explore several consequences that follow if the model's assumptions are correct.

7.1 Consciousness as a Time-Bound Activity

One implication is that consciousness is not simply "what the brain does" but specifically **what the brain must do in real time** to close the unavoidable gap between prediction and delayed sensation. This positions consciousness as an active process rather than a passive state. It suggests that:

- consciousness is fundamentally **temporal**,
- its stability depends on how fast the system can integrate new evidence,
- and disruptions in consciousness are often disruptions in timing rather than in content.

This provides a fresh perspective on phenomena like:

- the feeling of "being ahead of oneself,"
- the subjective slowing or speeding of time,
- moments when perception feels out of sync.

Instead of treating these as purely psychological, the IIM suggests they reflect fluctuations in η_c —the system's real-time integration efficiency.

7.2 The Self as a Predictive Record, Not an Entity

A second implication is that the self is not an inherent thing but a **statistical structure**: the long-term distribution of predictive outcomes. This approach resonates with several philosophical traditions, but it also provides a practical frame:

- The self is stable when predictive patterns are stable.
- The self becomes unstable when unresolved error accumulates.
- Identity crises, trauma, or chronic uncertainty reflect increases in H_{self}.

This shifts the discussion away from metaphysics and toward dynamics. Instead of asking "what is the self made of?" the more productive question becomes:

> *"How stable is the organism's long-term pattern of predictive successes and failures?"*

That reframing has interpretive power. Many experiences people describe as "losing oneself," "becoming someone else," or "not knowing who I am anymore" can be seen not as mysterious phenomenological states but as consequences of the rising entropy of the predictive record.

7.3 Anxiety as an Informational Imbalance

A third implication is that anxiety becomes the natural, even predictable, outcome of a mismatch between predictive demand and integrative capacity. Anxiety is not treated here as:

- an irrational response,
- a chemical imbalance,
- or a mere cognitive bias.

Instead, anxiety is the phenomenological signature of a system that is falling behind:

- prediction outpaces integration,
- unresolved error accumulates,
- self-entropy rises,
- and the organism loses confidence in its own internal models.

This account explains several features of anxiety:

- the sense of urgency,
- the feeling that time is slipping or thinning,
- the difficulty trusting one's own reactions,
- the tendency toward rumination.

Each emerges naturally from the dynamics of η_c and H_{self}. Anxiety is not a "malfunction," but a **signal** that the system is confronting demands it cannot resolve in time.

7.4 The Relationship Between Consciousness and the Self

Because consciousness (η_c) and the self (H_{self}) share prediction error as their common substrate, the IIM suggests that consciousness and selfhood are not independent domains. Instead, they form a layered system:

- η_c governs the now,
- H_{self} governs the long-term structure of expectations,
- and each influences the other.

This offers potential explanations for:

- why self-disturbances affect the immediacy of experience,
- why disruptions in presence affect long-term identity,
- and why high-anxiety states often include both momentary instability and persistent self-doubt.

If the self-model becomes highly entropic, predictive demand rises, compressing the present window and destabilizing consciousness. In turn, low η_c increases unresolved error, feeding back into the self-model.

7.5 Mental Health as a Balance of Error Dynamics

Another implication is that many psychological difficulties may be better understood in terms of error dynamics. For example:

- Depression could be interpreted as a regime where predictive demand is low but η_c is also low, leading to flattened experience.
- Anxious states arise when predictive demand rises faster than η_c.
- Dissociative states might occur when η_c collapses completely, forcing the system to "drop" the present moment.
- Trauma could be seen as a sudden, large injection of unresolved error that permanently alters H_{self}.

This approach does not reduce psychological states to equations, but it does provide a coherent framework that connects subjective experience to measurable dynamics of prediction and integration.

7.6 Consciousness in Artificial Systems

The IIM also has implications for artificial intelligence. If consciousness depends on time-sensitive integration rather than on biological substrate, then any system with:

- delayed input channels,
- predictive mechanisms,
- and a limited integration window,

would exhibit qualitatively similar dynamics. The model predicts that such systems might experience:

- stable or fragile integration regimes,
- rising or falling self-entropy (depending on design),
- anxiety-like instability when demand exceeds capacity.

While this does not imply that artificial systems currently have conscious experience, it suggests that certain patterns of internal instability may be universal in predictive architectures.

7.7 A Unified View

The most important implication is that consciousness, selfhood, and anxiety can be understood as different expressions of the same underlying structure:

> *- a physical world with finite causal speed,*
>
> *- an organism forced to predict,*
>
> *- and a need to integrate those predictions with delayed input.*

If this view is correct, the boundary between cognitive science, phenomenology, and affective psychology becomes less rigid than it appears. These domains may all reflect different aspects of a single control problem.

8. Limitations

No theoretical model is complete, and the Information Integration Model (IIM) is no exception. The framework offered here is intentionally simple and may overlook important biological, computational, or phenomenological aspects of consciousness. The value of acknowledging limitations is not to weaken the model but to clarify where it can be improved or challenged.

8.1 Simplicity of the Mathematical Structure

The mathematical expressions used in this paper are deliberately minimal. They describe:

- how prediction error accumulates,
- how integration efficiency (η_c) fluctuates,
- and how self-entropy (H_{self}) evolves.

However, these equations:

- do not capture the full richness of neural dynamics,
- do not engage with detailed cortical architectures,
- and do not incorporate probabilistic inference in a rigorous formal sense.

This simplicity is both a strength and a weakness. It makes the model transparent and easy to test, but it risks leaving out relevant mechanisms that more sophisticated computational models capture.

8.2 Lack of Neurobiological Specificity

While the model builds on empirical findings in time perception and prediction, it does not attempt to map η_c or H_{self} onto specific neural circuits. The model assumes:

- the existence of prediction mechanisms,
- temporal integration windows,
- and error-monitoring systems,

but it does not specify:

- which cortical regions compute η_c,
- how unresolved error is represented biologically,
- or how long-term error distributions are encoded.

More biologically detailed models—such as those inspired by predictive coding or dynamical systems neuroscience—could refine these aspects.

8.3 Conceptual Tension Between Phenomenology and Computation

The IIM attempts to bridge subjective experience with computational dynamics. This is an ambitious move and carries risks:

- subjective experience may not always align neatly with measurable quantities,
- η_c is a functional construct, not a directly observable property,
- H_{self} is a conceptual entropy, not a thermodynamic one.

Although I have tried to ground phenomenological claims in functional dynamics, this bridge remains provisional.

8.4 Dependence on Coarse Error Categories

The definition of H_{self} relies on categorizing unresolved errors into discrete types. This is a simplification. In reality:

- prediction errors may exist on continuous or multi-dimensional scales,
- categories may overlap or shift over time,

- and the system may construct new categories that the model does not anticipate.

This means the current form of H_{self} should be treated as a first approximation, not a definitive measure.

8.5 Limited Scope of the Experimental Predictions

The experiments proposed in Section 6 are designed to test broad predictions of the IIM. However:

- they may not isolate η_c or H_{self} cleanly,
- behavioral measures may be influenced by confounding factors,
- and subjective reports are notoriously variable.

The experimental framework is therefore best viewed as a starting point rather than a complete empirical program.

8.6 Possible Alternative Explanations

Many of the phenomena described by the IIM can also be explained by other theories:

- Predictive processing offers mechanisms for prediction and error correction.
- Global workspace theories explain access and integration through broadcasting.
- Integrated information theory describes consciousness in terms of causal structure.
- Cognitive load theories explain anxiety through resource limits.

The IIM differs in emphasizing **finite causal delay** and **error dynamics** as foundational, but whether this foundation is necessary—or merely one possibility—remains open for debate.

8.7 Philosophical Challenges

Although the model attempts to remain naturalistic and non-metaphysical, several philosophical issues remain unresolved:

- Does η_c describe consciousness, or only the conditions under which consciousness is stable?

- Does H_{self} capture the self, or only its functional scaffolding?
- Does anxiety emerge from error dynamics alone, or do affective systems play an independent causal role?

These questions require further philosophical analysis and may expose weaknesses in the model's assumptions.

8.8 Summary

The IIM is best understood as a conceptual framework rather than a fully developed theory. Its strength lies in unifying prediction, integration, and self-entropy under a physically grounded constraint. Its weaknesses lie in its simplicity, lack of neural detail, and dependence on constructs that require further empirical validation.

> *Acknowledging these limitations clarifies the model's role: not as a final answer to the nature of consciousness, but as an invitation to explore how finite causal speed shapes the structure of experience.*

9. Conclusion

The central idea of this paper is simple: the world does not arrive on time. Every signal that reaches an organism—light, sound, touch, internal feedback—comes with inherent delay. This small physical fact forces a large cognitive consequence: to survive, an organism must continuously predict the near future and integrate those predictions with evidence that always comes too late. From this constraint the core components of the Information Integration Model (IIM) arise.

If the present is always slightly beyond reach, consciousness becomes the activity of repairing that gap. It is not an internal light or a hidden essence but the ongoing work of keeping perception and prediction aligned within tight temporal limits. The self, in turn, becomes the long-term record of this work—a history of predictive successes and failures that forms the backbone of the organism's expectations about itself. And anxiety appears when the system begins losing the race between predictive demand and integrative capacity, producing a feedback loop that destabilizes both experience and identity.

The conceptual quantities introduced in this paper—**Integration Efficiency (η_c)** and **Self-Entropy (H_{self})**—are not proposed as final theoretical primitives. They serve as placeholders that allow us to speak about real-time performance and long-term stability in a way that is testable and falsifiable. Their value lies in revealing common structure beneath phenomena typically studied separately: temporal distortions, self-instability, and affective tension.

The model also invites a shift in perspective. Instead of treating consciousness as something added by the brain, we might view it as something required by the universe. In a world of instantaneous signals, organisms would not need to predict, integrate, or maintain a coherent self-model. But in the world we inhabit, where every action depends on anticipating a moment that has not yet arrived, consciousness becomes less a mystery and more a functional necessity.

> *This does not diminish the richness of subjective life. Instead, it roots that richness in a fundamental constraint: we experience because we must keep up with a world that is always ahead of us.*

The IIM offers only a starting point for understanding this relationship. Much remains to be clarified, expanded, or corrected. But if the view developed here is even partly right, then the structure of experience is not arbitrary. It reflects the physics of information flow, the dynamics of prediction and error, and the fragile balance through which organisms remain aligned with a world they can never perceive in real time.

10. References (Prose Style)

The ideas in this paper draw on several research traditions. The discussion of predictive processing relies on work by Karl Friston and colleagues, who have argued over many years that the brain continually anticipates sensory signals and corrects its predictions when errors occur. Although the present model approaches prediction from a different angle, the general insight—that perception and action depend on anticipation—owes much to this line of thinking.

The account of temporal integration is influenced by the work of David Eagleman, whose experiments on simultaneity, delay, and subjective time have shown that the brain stitches events together within short temporal windows to generate a coherent sense of the present.

These observations support the idea that consciousness depends on the timely integration of incoming information.

The view of the self as a constructed model draws partly on the work of Thomas Metzinger, who has argued that the self is not a metaphysical entity but a transparent, functional representation created by the brain. While the version of the self proposed here differs in its emphasis on predictive history and entropy, the general idea that the self is built rather than discovered is shared.

The discussion of anxiety reflects themes found in psychological research on uncertainty, cognitive load, and anticipatory threat. Many authors have noted that anxiety is tied to a mismatch between environmental demands and internal resources. The present model extends this by suggesting a specific mechanism for how such mismatches destabilize both momentary experience and the long-term self-model.

Finally, the broader project of explaining consciousness in terms of integration owes something to Giulio Tononi's work on integrated information. Although the IIM takes a different direction by focusing on temporal rather than spatial integration, Tononi's effort to link consciousness with structural constraints has been an important inspiration.

Correspondence concerning this article should be addressed to Qinghong Yu at **Contact: yuqinghong2002@gmail.com | Phone: +86 18033660513**

This is a preprint manuscript. Final version may differ.

Generated on December 4, 2025