# Universal Collapse Theory—Conscious Collapse: Mind as a Phase of Constraint-Guided Collapse (WP04 v1.0)

## Perception, Belief, and Free Will as Collapse under Neural and Cognitive Constraints

Jeremy C. Jones (ORCID 0009-0007-2515-3774) — HoldingLight LLC — contact@universalcollapse.com



Part of the Universal Collapse Theory White Papers Series

Companion volume: Universal Collapse Theory (2025), ISBN 978-1-969095-01-6.

Version: v1.0—Prepared 2026–08

# Conscious Collapse: Mind as a Phase of Constraint-Guided Collapse


## Abstract

This paper extends Universal Collapse Theory (UCT) into the mind-phase, treating human consciousness as a proposed regime in which some neural resolutions become globally available and recursively usable for report, deliberate control, memory, and possible later constraint update. The neural claim is a defeasible access-level identification hypothesis: broadcasting, recurrence, integration, and differentiation are candidate third-person signatures, not sufficient conditions by definition. Interpretively, UCT advances a non-bifurcation thesis: first-person presence and third-person neural organization are non-interchangeable modes of access to one conscious occurrence, not ontologically separate products. This does not settle phase criteria, episode unity, or content correspondence. Non-conscious processing remains locally effective; conscious processing is provisionally modeled as system-wide, globally available, recursively usable resolution.

Perception is modeled as hypothesis resolution under constraints supplied by priors, context, goals, bodily state, and neural architecture. The FRLB schema—Faith, Reason, Logic, Belief—is presented as a proposed functional-role decomposition of cognitive resolution and update. A revised heuristic update model distinguishes Reason from the kernel record symbol and separates Event Belief from standing Belief. Action and agency are modeled as constrained behavioral selection; free will is treated operationally as graded reasons-responsive participation in revising selected later constraints, not as an escape from causation or a settlement of the metaphysical free-will debate.

Across histories of certified records and constraint updates, learning and memory can stabilize an identity architecture; rigidity or misalignment becomes a testable hypothesis about update dynamics rather than a clinical diagnosis. Consciousness-Induced Material (CIM) is retained as the consciously shaped, record-bearing layer through which cognition becomes durable and can non-trivially constrain later minds. The individual neural, perceptual, FRLB, and institutional predictions are not unique to UCT. The discriminating burden lies in preregistered transfer, for each S-signature, of a restricted model across physics, biology, and mind. Successful transfer would support the shared-schema claim; qualitative analogy or domain-specific refitting would not.



**Keywords:** conscious access; collapse under constraint; global workspace; predictive processing; Faith–Reason–Logic–Belief (FRLB); agency and constraint management; identity and update rigidity; Consciousness-Induced Material (CIM); cross-domain signature transfer

## Contents

## §1 Introduction: Conscious Collapse and the Mind-Phase

Universal Collapse Theory (UCT) treats reality as **collapse under constraint**: structured potential resolving into realized form. WP01 introduced the series schema. Under the current kernel contract, a transition emits a realized resolution, a candidate trace, and an open residue carrier; recordhood is certified separately, and the update map runs whether certification succeeds or returns None. Appendix A states the interface used in this revision.

$$(x_t^*, q_t, S_t) \sim C_t^K(\omega_t, \cdot),$$
$$R_t := \mathrm{cert}_P^d(\mathrm{out}(x_t^*), q_t), \qquad K_{t+1} := U(K_t, x_t^*, R_t, S_t)$$

WP02 applied the schema to physics, and WP03 extended it to Biological Collapse: organization maintained under sustained gradients and inherited constraints (Jones 2025b, WP01 v2.0; Jones 2025a, WP02 v1.0; Jones 2026l, WP03 v1.0). WP04 begins at the handoff from life to mind and asks:

What changes when collapse under constraint is implemented in systems that can make some of their own resolutions globally available, represent them as content, and use them in later control and update?

We call this proposed regime Conscious Collapse or the mind-phase of UCT. At this level, the mind's possibility space includes hypotheses, percepts, intentions, and actions; its constraints include priors, values, goals, self-models, biological conditions, and cultural context; and typed mind-domain kernels yield neural and cognitive resolutions. Neural traces, memories, narratives, and externalized structures are candidate record placements. They count as certified records only after a declared mind-domain test establishes durability, inspectability, capacity to constrain a later update, and material dependence on the realized resolution. The durability and inspectability requirements inherit the record semantics of Records Across Nature, Life, and Mind v2.0 (Jones 2026i); the Option-typed certification interface used here follows the current kernel contract. No complete mind-domain test is claimed here.

Claim status. WP04 combines four claim levels that must not be conflated: established findings imported from cognitive neuroscience; UCT-specific structural mappings; a program-level non-bifurcation thesis according to which first-person presence and third-person organization are non-interchangeable modes of access to one conscious occurrence; and open empirical hypotheses about which structures identify that phase. The non-bifurcation thesis is not derived from the generic kernel, and success of an access marker would not by itself establish it. Evidence for global availability, predictive processing, complexity, therapy, or institutional hysteresis supports the source theories on its own terms; it becomes evidence for UCT only when a restricted cross-domain model transfers without ad hoc redefinition. Section 9 makes that burden explicit.

**Reader orientation**

WP04 is a spine paper whose full argument is cumulative, and sequential reading is recommended for readers new to UCT. Readers already familiar with global-workspace, predictive-processing, and related consciousness literature may treat §§2–3 as the neural and perceptual alignment layer and move first to §2.4 and §§4–7, where the paper's mind-phase commitments, update architecture, and record-bearing extension are developed. Readers principally evaluating whether UCT offers more than domain redescription may move from §1 to §§8.4–10, which state the proposed contribution, discriminating transfer test, failure conditions, and open obligations. Appendix C isolates the paper's metaphysical position. These are alternate entry points into the argument, not substitutes for the dependencies established in the intervening sections.

**Scope: human consciousness as a specific mind-phase**

Human Conscious Collapse, as modeled in WP04, is a proposed phase of Biological Collapse implemented in living nervous systems. The more general conscious-phase category is realization-neutral in principle; whether any non-biological architecture satisfies its still-open phase criteria is not established here. Many non-human animals are plausible biological candidates on graded dimensions of perception, learning, affect, and flexible behavior (Edelman & Seth 2009; Griffin & Speck 2004; Birch et al. 2020). This paper focuses on humans because behavioral report, neuroscience, and phenomenology are most richly linked there. It does not treat biological directionality or flexible behavior alone as proof of conscious experience, and it does not presume that all candidate mind-phases share the same constraint architecture.

Roadmap

Sections 2–3 develop the provisional phase-identification and perception models. Sections 4–6 treat FRLB, action, identity, and update rigidity. Section 7 protects the CIM contribution while distinguishing cognitive process from record-bearing output. Section 8 locates the proposal among GWT/GNWT, predictive processing, IIT, and quantum approaches. Section 9 gives the empirical program and cross-domain transfer criterion; Section 10 states limitations and open obligations; Section 11 concludes and hands the law-level burden to WP05.

## §2 The Conscious Phase: Neural Substrate and Thresholds

Biological Collapse supplies the neural substrate: self-organizing networks shaped by evolution and development. WP04 proposes that an access-conscious regime begins when some neural resolutions become:

1. **globally available** within the system, and
2. recursively usable as reportable content for memory, deliberate control, and later update.

This section sketches how that threshold can be understood in the collapse-kernel language without committing to a specific microphysical story.

## 2.1 From life-phase to mind-phase

In WP03, Biological Collapse modeled life as networks that capture gradients, preserve organization, and reproduce under biological constraints (Jones 2026l, WP03 v1.0). Nervous systems are specialized implementations of this phase: highly plastic networks that

- rapidly integrate signals from many modalities,
- form and break connections,
- and route activity through recurrent loops (Kandel et al. 2013; Dayan & Abbott 2001).

From the kernel's perspective, a brain defines:

- a structured possibility space $\Omega_{\text{neural}}$: patterns of spiking, synaptic strengths, field dynamics;
- a constraint set $K_{\text{neural}}$: anatomy, connectivity, neuromodulators, receptor distributions, metabolic limits;
- a neural-domain transition kernel: neural dynamics that resolve particular activity patterns from the live neural state;

Most of this is simply **non-conscious** processing: reflexes, local loops, routine sensorimotor transformations can all be described as collapse in $\Omega_{\text{neural}}$ under relatively local $K$. Conscious Collapse, however, appears when neural architecture gains the capacity to:

- integrate information across diverse subsystems, and
- broadcast the result as a **global constraint** on further collapse.

WP04 provisionally classifies a winning workspace pattern as access-conscious when it is globally available for report, deliberate control, working-memory use, and possible participation in later constraint update (Baars 1988; Dehaene & Changeux 2011; Dehaene 2014). That classification is a phase-identification hypothesis, not an entailment from broadcast alone.

## 2.2 Global broadcasting and integration–differentiation

Global Workspace Theory (GWT) and related empirical work suggest a simple structural story: in conscious states, certain neural activity patterns are broadcast across many specialized systems, while in non-conscious states, processing remains local. Integration-differentiation measures (e.g., perturbational complexity indices) similarly indicate that conscious states occupy a regime of:

- high **integration** (information from many areas is combined), and
- high **differentiation** (activity patterns are richly structured, not uniform) (Dehaene & Changeux 2011; Casali et al. 2013; Massimini et al. 2005).

UCT can recast this in kernel terms. Consider:

- $\Omega_{\text{local}}$: hypotheses and activity patterns in local modules (sensory cortices, motor areas, etc.),
- $\Omega_{\text{GW}}$: patterns that involve coordinated activity across many modules (a "workspace"),
- $K_{\text{mind}}$: constraints that determine which patterns are accessible and how they influence downstream processing.

We can then distinguish two regimes of collapse:

- **Non-conscious collapse:**

$$(x^*_{\text{local,t}}, q_{\text{local,t}}, S_{\text{local,t}}) \sim C^{K_{\text{local}}}_{\text{local,t}}(\omega_t, \cdot)$$

Local circuits resolve ambiguity, drive behavior, or update internal state, but these transitions are not globally broadcast; they do not become part of the system's current experiential field.

- **Conscious collapse:**

$$(x^*_{\text{GW,t}}, q_{\text{GW,t}}, S_{\text{GW,t}}) \sim C^{K_{\text{mind}}}_{\text{mind,t}}(\omega_t, \cdot)$$

A particular pattern wins access to a workspace-like configuration and is broadcast, becoming available to multiple subsystems for report, deliberate control, and possible memory encoding. WP04 calls that pattern access-conscious content for the episode (Dehaene 2014; Mashour et al. 2020).

On this view, the proposed access threshold is a network-level regime in which activity becomes globally influential, differentiated, and recurrently available. Perturbational complexity and related measures are possible indicators of that capacity, not definitions of consciousness and not by themselves sufficient to identify a conscious occurrence (Casali et al. 2013; Massimini et al. 2005).

## 2.3 Conscious vs non-conscious collapse

With this distinction in hand, we can define Conscious Collapse more precisely:

- **Non-conscious collapse**
    - occurs whenever a local neural subsystem resolves its own structured potential under local constraints;
    - can be fast, automatic, and behaviorally effective;
    - may leave local candidate traces in synapses, priming, or reflex gains; whether a trace qualifies as a record depends on the declared certification test, not on conscious access.
- **Conscious collapse**
    - occurs when a pattern of neural activity wins access to a global workspace-like configuration;
    - is shaped by a richer constraint set $K_{\text{mind}}$ that includes priors, goals, self-models, and social context;
    - may leave candidate traces associated with memories, narratives, or explicit beliefs; qualified records can later be taken up by U, but neither certification nor uptake is guaranteed by global availability (Baars 1988; Dehaene & Changeux 2011; Dehaene 2014).

Formally, local and workspace-level transitions are placements in a typed mind-domain kernel family. They share an interface but differ in state space, constraints, candidate traces, and access consequences:

- Non-conscious: a draw from a local kernel remains locally effective and is not globally available as current content.
- Access-conscious: a workspace-level draw becomes globally available and recursively usable. UCT does not model phenomenal presence as an additional output appended after this draw. Whether global availability and recursive use correctly identify the conscious phase remains an empirical and phenomenological question.

This distinction supplies a candidate functional discriminator without introducing a second substance. It does not make broadcasting sufficient by definition, establish the non-bifurcation thesis empirically, or specify the episode's phenomenal character.

The rest of the paper uses this neural threshold provisionally and asks what follows if the proposed phase criterion holds. Perception, FRLB, action, learning, and identity are treated as mind-domain specializations of a shared kernel schema, not as applications of one concrete operator over unlike state spaces.

### 2.4 The inside–outside relation and a falsifiable access hypothesis

**Non-bifurcation thesis.** For a genuinely conscious occurrence, first-person presence and third-person neural organization are non-interchangeable modes of access to one process, not two ontologically alien products connected by a production step. WP04 therefore rejects the picture in which a complete non-experiential neural process subsequently emits an additional phenomenal object. This is a philosophical commitment of the mind-phase account, not a conclusion derived from the generic kernel. In the kernel's technical vocabulary, event remains reserved for a transition whose candidate trace certifies; conscious occurrence and formal eventhood must therefore be distinguished until a mind-domain record test is supplied (Chalmers 1995).

**Empirical phase-identification hypothesis.** Global availability for report, flexible control, memory, and self-referential update—together with recurrence, integration, and differentiation—are proposed third-person indicators of conscious access (Block 1995). They are not sufficient by definition. The mapping would be weakened by conscious episodes lacking the predicted signatures, by matched non-conscious processing exhibiting them, or by failure to distinguish content availability from attention, report preparation, ordinary evidence accumulation, and post-resolution maintenance.

Independently, the sustained-resolution hypothesis proposes that recurrent coordination across completed local transitions may keep competing or partially settled content globally influential long enough for multiple systems to operate on it. This is a candidate temporal mechanism, not the basis of the non-bifurcation thesis and not a result derived from the

kernel. It predicts measurable temporal persistence and recurrent causal influence beyond matched local processing.

The hypothesis fails if conscious access systematically occurs without the proposed recurrent interval, if the interval tracks task or report demands rather than content availability, or if adding interval measures yields no predictive gain over established workspace, recurrent-processing, or higher-order accounts. $S_2$ remains a signature of ambiguity and delayed resolution, not a constitutive feature of every conscious state.

**Remaining obligations.** This reframing does not remove WP04's explanatory debts. The paper still owes discriminating criteria for the conscious-phase boundary; the temporal unity, limits, and cessation of an episode; the mapping from neural and constraint organization to particular phenomenal contents; and the relation between first-person reports and third-person measures. Under the non-bifurcation thesis, a completed mind-domain model must also specify enough relevant structure that a phenomenal difference between systems identical on every declared variable would expose a missing variable or an incomplete mapping. These are problems of phase identification, phenomenological differentiation, structural completeness, and first-person/third-person correspondence—not, within UCT, a demand to manufacture experience from an ontologically alien substrate. Section 8 states this position in full.

## §3 Perception as Hypothesis Collapse: The Active Observer

Perception is modeled here as constrained hypothesis resolution. A perceptual hypothesis becomes access-conscious when it wins global availability and can guide report, further thought, and action. This section rewrites predictive-processing and workspace ideas in the current domain-kernel language without claiming that the access model exhausts phenomenal structure.

- $\Omega_{\text{percept}}$ — a structured space of possible interpretations of current sensory input,
- $K_{\text{percept}}$— constraints shaped by priors, context, goals, and bodily state,
- Perceptual transition kernel — the provisional domain kernel that resolves a hypothesis under the current perceptual constraints.

Access-conscious perception is the global availability and recursive use of the selected hypothesis.

### 3.1 Perceptual Ω and constraint architecture

At any moment, the sensory surfaces and early processing stages deliver a torrent of ambiguous signals. For a given sensory scene, there is not one but many candidate interpretations:

- object identities and boundaries,
- distances, orientations, motions,
- inferred causes (what is happening, who is doing what, why).

Let $\Omega_{\text{percept}}$ denote the space of such **perceptual hypotheses**. This space is structured: nearby hypotheses in $\Omega_{\text{percept}}$ share features (e.g., "tilted line" vs "slightly more tilted line"), while distant ones disagree qualitatively ("face" vs "tree") (Goldstein 2014).

The constraints that shape which hypotheses are admissible and how strongly they are favored can be grouped as:

- **Priors and learned models** $K_{\text{prior}}$: expectations built from past experience (what objects usually look like, how they move, what is likely in this context).
- **Immediate context** $K_{\text{context}}$: current scene, task, bodily posture, and environmental cues.
- **Goals and values** $K_{\text{goal}}$: what the organism currently cares about or is attending to (threats, opportunities, instructions).
- **Neural implementation** $K_{\text{neural}}$: architecture of sensory cortices and higher areas, recurrent connectivity, and neuromodulatory state (Clark 2013; Friston 2010).

Together these define a perceptual constraint set. Compose is used below as a paper-local placeholder for domain-specific constraint composition; it is not an additional universal operator. A runnable model must specify the types, weighting, interaction rules, and admissibility conditions for the listed components:

$$K_{\text{percept}} := \text{Compose}(K_{\text{prior}}, K_{\text{context}}, K_{\text{goal}}, K_{\text{neural}})$$

Given sensory input and this $K_{\text{percept}}$, the brain does not passively "receive" a picture. It uses recurrent dynamics and feedback to explore $\Omega_{\text{percept}}$, then **collapses** onto a particular hypothesis:

$$(x^*_{\text{percept,t}}, q_{\text{percept,t}}, S_{\text{percept,t}}) \sim C^{K_{\text{percept}}}_{\text{percept,t}}(\omega_t, \cdot)$$

Under the phase-identification hypothesis, the selected interpretation becomes access-conscious when it is broadcast and made available to guide later cognition and action (Dehaene 2014; Clark 2013).

## 3.2 Predictive processing as collapse

Predictive-processing models fit naturally into this picture. They treat perception as:

- maintaining a generative model of the world,
- predicting incoming sensory signals,
- minimizing prediction error by updating internal hypotheses or acting on the world (Rao & Ballard 1999; Friston 2010; Clark 2013).

UCT rephrases this as:

- $\Omega_{\text{percept}}$ is implicitly encoded by the generative model;
- $K_{\text{percept}}$ includes the model's parameters, precision weights, and error-coding structure;
- dynamics in cortical hierarchies approximate the search for a hypothesis in $\Omega_{\text{percept}}$ that minimizes a constraint-shaped error functional.

On the present functional model, access-conscious perception corresponds to a reportable resolution after further error reduction yields diminishing returns and the system settles on an interpretation available for broadcast.

This makes the "active observer" role explicit. The system is not a passive recipient of stimuli. It is an agent that:

- proposes hypotheses in $\Omega_{\text{percept}}$,
- uses constraints embedded in $K_{\text{percept}}$ to prune them,
- and commits to one $x^*_{\text{percept}}$ as its current model of the world.

The reportable percept is the functional endpoint of this model; the model does not yet derive experiential differentiation or validate the correspondence between modeled structure, report, and first-person evidence.

### 3.3 Ambiguity, rivalry, and $S_2$: when collapse is delayed

Ambiguous and rivalrous stimuli—bistable images, Necker cubes, binocular rivalry—provide clean demonstrations that perception is not a simple feedforward mapping from input to content. The same stimulus can give rise to **multiple, mutually exclusive percepts** over time (Leopold & Logothetis 1999; Sterzer, Kleinschmidt & Rees 2009).

In kernel terms:

- The sensory input constrains $\Omega_{\text{percept}}$ but does not uniquely determine $x^*$: there are at least two comparably coherent hypotheses compatible with $K_{\text{percept}}$.
- The system occupies a near-neutral region in $\Omega_{\text{percept}}$ in which the declared outcome weights do not yet strongly favor one candidate.
- Over time, small fluctuations in internal state and attention change $K$ just enough to bias collapse, and the system alternates between attractors: first one interpretation becomes $x^*$, then the other (Devia et al. 2022).

Bistable perception is a candidate $S_2$ placement and a suitable domain for an audited $S_2$ test; ambiguity and longer latency alone do not establish the signature:

- **Neutral plateau:** for an extended period, the system hovers in a state where multiple perceptual outcomes are compatible; response times and uncertainty are elevated.
- **Resolution:** a small shift in $K$ (e.g., a micro-change in attention or neural noise) tips the system, and collapse selects a specific $x^*_{\text{percept}}$, which then dominates conscious content.
- **Alternation:** when conditions for the plateau persist, collapse may oscillate between attractors rather than stabilizing, leading to the characteristic flipping of conscious experience.

Ambiguity effects in reaction-time tasks, decision conflict, and bistable perception are therefore candidate mind-phase placements for neutrality → delayed resolution. An audited

$S_2$ claim requires an independently specified asymmetry or neutrality variable, a directional latency prediction, and comparison with relevant confounds rather than a neutrality score fitted from the same latency distribution it is used to explain.

### 3.4 Redundancy and consensus: perceptual $S_1$

Perception also offers a candidate $S_1$ placement (redundancy → consensus). When sensory input and constraints are clear:

- Repeated presentations of the same stimulus under similar $K_{percept}$ can establish response reliability, but repetition alone is amplification rather than independent redundancy.
- Independently informative sensory channels and cues may converge on the same $x_{percept}^*$: visual, auditory, and proprioceptive information can align to a single coherent scene (Ernst & Banks 2002).
- Across individuals with similar $K$ (similar priors, context, and training), there is substantial agreement about what is seen.

UCT treats this as a candidate $S_1$ placement in a well-constrained regime, subject to a declared independence audit:

- Redundancy in independently informative channels or records, together with distributed internal representation; repeated copies or strongly correlated cues count as amplification, not independent redundancy
- → **consensus in collapse outcomes**: the system repeatedly lands in the same region of $\Omega_{percept}$ because it is a **deep coherence pocket** under $K_{percept}$.

Where independent redundancy is low or K is unstable, consensus may break down and candidate $S_2$-like patterns of ambiguity or conflict may appear. Robust, rapid perception reflects not only sensory coding but a well-shaped constraint architecture.

### 3.5 Perception as input to FRLB

Perception does not operate in isolation. The hypotheses that win collapse in $\Omega_{percept}$ become inputs to the higher-order collapse processes described by FRLB (Faith–Reason–Logic–Belief):

- Some perceptual priors may function as phase-Faith inputs when they are received as premises within the declared inquiry episode. Other priors remain ordinary perceptual constraints without occupying the Faith role.
- The Reason role and the episode's admission rules operate over sequences of perceptual resolutions, inferring structure and relationships from selected interpretations.
- Standing Belief (B) is retained commitment-bearing structure within $K_{mind}$ that may result from integrating perceptual evidence over time (Jones 2026g, How Minds Resolve).

Each perceptual resolution may co-produce candidate traces. Where certification succeeds, a resulting record becomes available for later uptake; where relevant uptake also succeeds, that record can contribute to longer-lived perceptual, conceptual, or behavioral constraints. Where certification does not succeed, local processing can still proceed without producing a certified mind-level record. In that bounded sense, perceptual Conscious Collapse supplies hypothesis-shaped content that FRLB may use in building or revising a model of the world.

## §4 FRLB: Faith, Reason, Logic, and Belief as Functional Roles

Perception supplies momentary resolutions, but conscious cognition also retains and revises longer-lived commitments. How Minds Resolve names four functional roles in this cycle: Faith, Reason, Logic, and Belief (Jones 2026g). The roles are not serial modules, and the same content can occupy different roles in different inquiry episodes.

The FRLB roles are:

- Faith — cycle-relative starting trust: premises received by the present act rather than established within it.
- Reason — recursive generation, comparison, and refinement of candidate interpretations against feedback.
- Logic — coherence, admission, and update constraints distributed across the standing architecture and the update rule.
- Belief — Event Belief is the certified commitment payload for an episode; standing Belief is a later commitment-bearing structure after update. They are linked by provenance, not identical quantities.

FRLB is a proposed functional-role decomposition, not a religious taxonomy, four-box cognitive module model, clinical classification, or validated psychometric scale.

### 4.1 Mapping FRLB onto the collapse kernel

We can write FRLB as a proposed mind-domain specialization of the shared kernel schema:

- $\Omega_{mind}$: space of possible interpretations, explanations, plans, and self-models—larger and more abstract than $\Omega_{percept}$, but informed by it.
- $K_{mind}$: constraint architecture encoded as current beliefs, values, priors, and structural rules of inference.
- Mind-domain transition kernel — the provisional kernel for thoughts, judgments, or decisions.
- $x_{mind,t}^{*}$: realized mind-domain resolution (a specific judgment, interpretation, or decision), which is classified as access-conscious only when the declared phase criterion is met.
- $q_{mind,t}$: candidate traces associated with memories, narratives, external notes, or social feedback. When a declared certification test succeeds, $R_{mind,t} = Some(r_{mind,t})$; otherwise $R_{mind,t} = None$.
- Mind-domain update map — consumes the realized resolution, the Option-typed record, and the open carrier; it can modify later constraints even when certification returns no record.

Within this frame:

- Phase Faith is the commitment-bearing subset of $K_{mind}$ received as starting trust within a defined inquiry episode. It is not all of $K_{mind}$: biological conditions, context, inferential rules, and other constraints remain distinct. A premise may be historically well supported and still function as Faith if the present act receives rather than re-derives it.
- Reason is candidate generation and feedback-sensitive refinement within the live resolution process. The role name is spelled out so that the reserved kernel symbol R continues to mean certified record.
- Logic constrains what may stabilize and what may be admitted into later update. Standing consistency requirements live in $K_{mind}$; transition and admission discipline are expressed through $U_{mind}$.

- Following How Minds Resolve v1.0, Event Belief is the cognitive record of resolution $x_{mind,t}^{*}$. Under the Option-typed interface used here, this reads as the certified payload $r_{mind,t}$ when $R_{mind,t} = Some(r_{mind,t})$: the wrapper is interface typing, not content. There is no Event Belief when $R_{mind,t} = None$. Standing Belief is the commitment-bearing structure $B \subset K_{mind,t+1}$ into which an event record may sediment after U. The two have continuous provenance but different types and functions.

The FRLB cycle can be summarized as:

1. Starting from $K_{mind,t}$—which includes phase-Faith inputs, search habits, admission rules, and standing Belief—perception and thought explore $\Omega_{mind}$.
2. A mind-domain draw yields a specific judgment, conclusion, or commitment.
3. The resolution co-produces a candidate trace $q_{mind,t}$; certification may return $Some(r_{mind,t})$; otherwise it returns None.
4. The update map $U_{mind}(K_{mind,t}, x_{mind,t}^{*}, R_{mind,t}, S_{mind,t})$ runs in either case and produces $K_{mind,t+1}$.

FRLB is one proposed decomposition of the cognitive roles surrounding this mind-domain update cycle. It does not add a second collapse operator.

Two grains must remain separate. Phase Faith is a task-bounded input role; standing Belief is a retained part of later $K_{mind}$. UCT also advances a philosophical thesis of constitutive commitment: finite inquiry begins from premises it cannot all justify within the same act. That thesis motivates the vocabulary but is not required to validate the empirical FRLB distinctions, and it should not be inferred from the sustained-resolution hypothesis.

Belief likewise changes role across the update seam. Under that Option-typed reading, if $R_{mind,t} = Some(r_{mind,t})$, the certified payload $r_{mind,t}$ is Event Belief for that episode; if $R_{mind,t} = None$, there is no Event Belief. After U, related content may become standing Belief $B \subset K_{mind,t+1}$. Event record and standing constraint are not one quantity at two times; they are linked by provenance through an update that can transform function and type.

This mapping follows How Minds Resolve, the deposited FRLB role-and-grain authority (Jones 2026g). WP04 owns the working ansatz below. Human Interface Laws develops the interface-level pressures that modulate bandwidth, signal weighting, identity protection, record retention, self-sealing, and reopening (Jones 2026h). Biological Faith Systems supplies the preconscious commitment-before-certainty lineage (Jones 2026d).

The Tether §61 provides the corpus-level Peircean anchor (Jones 2026k), while How Minds Resolve §7 develops the mind-domain comparison (Jones 2026g). WP04's additional obligation is narrower: state how researchers could distinguish the four roles without pretending that a validated measure already exists.

## 4.2 Operational boundary and current measurement status

The roles are episode-relative and must be coded against a preregistered unit of inquiry. Before exposure to the focal evidence, researchers identify the premises received as starting trust; during the task they log candidate generation and search; after the task they code admission rules and the commitment retained. This prevents the content of a proposition from determining its FRLB label by itself.

**Faith versus standing Belief.** Faith is a premise functioning as input to the current act. Standing Belief is a retained commitment already constraining that act. The same proposition can be standing Belief at t and phase Faith at t+1; temporal role, not wording or confidence, separates them.

**Motivated Reason versus distorted Logic.** Motivated Reason appears as asymmetric search, sampling, elaboration, or counterexample generation while stated admission standards remain stable. Distorted Logic appears when consistency or evidential standards themselves change, are selectively applied, or admit conclusions that do not follow from the acknowledged material.

Replication requirement. Independent coders should receive time-ordered process traces, declared standards, and held-out cases; the research program must report inter-rater reliability and incremental prediction beyond established measures of motivated reasoning, cognitive rigidity, and belief updating. The current unvalidated working codebook (FRLB-RC v0.1) supplies development-stage coding rules; WP04 does not present it as a finished public instrument.

**Failure condition.** If trained coders cannot distinguish the roles above chance, if the categories collapse into existing constructs without incremental value, or if the distinctions do not predict update trajectories, the operational FRLB claim should be narrowed or rejected.

## 4.3 A heuristic update ansatz

WP04 adopts the following working ansatz. It is not a derivation from the kernel, a microscopic synaptic rule, a universal learning law, or a finished observation model. Here, $B_t$ is one task-bounded standing-Belief coordinate, $F_t$ is the task-bounded phase-Faith input,

1 The Tether centralizes the corpus-level inheritance from Peirce—consequence-legibility, external resistance, and fallibilism—while How Minds Resolve applies that inheritance to FRLB, including its semiotic and abductive seams.

$\text{Reason}_t$ is the output of candidate search, $e_t$ is acknowledged evidence, and $\text{Lic}_t$ is the episode's task-local licensing function rather than the whole Logic role.

The working ansatz is:

$$\Delta B_t := B_{t+1} - B_t$$

$$\hat{B}_t = \text{Lic}_t(F_t, \text{Reason}_t, e_t)$$

$$B_{t+1} = B_t + \alpha_t(\hat{B}_t - B_t)$$

where:

- The first line defines observed standing-Belief change; the second defines the task-local licensed target from starting premises, search output, and acknowledged evidence under the declared licensing rule.
- The baseline and licensed target must occupy the same task-defined scale or representation.
- The effective update parameter is episode- and belief-specific. When monotonic movement without overshoot is assumed, it is constrained to the unit interval; values outside that interval require an explicit overshoot or reversal model.
- The fitted update parameter is descriptive uptake at the chosen observation grain—not willingness, intelligence, virtue, or recordhood (Jones 2026g, How Minds Resolve; Knill & Pouget 2004; Oaksford & Chater 2007).

This captures several intuitive phenomena:

- When the licensed task-bounded target differs from the baseline and the effective update rate is nonzero, the observed belief representation moves toward that target.
- A larger effective update rate yields faster movement in the fitted ansatz; a smaller value yields slower movement, conditional on the observation model and evidence direction.
- Mis-specified premises, search, evidence coding, or admission rules can license an inaccurate target even when the fitted update rate is large (Festinger 1957; Hahn & Oaksford 2007).

In kernel terms, this is a task-level observation ansatz for one selected standing-Belief coordinate of the output of the mind-domain update:

$$K_{\text{mind,t+1}} = U_{\text{mind}}(K_{\text{mind,t}}, x^*_{mind,t}, R_{mind,t}, S_{mind,t})$$

It neither parameterizes nor decomposes the full update map. FRLB labels functional roles surrounding U so researchers can ask which premises, search processes, admission rules, prior commitments, and records contributed to the observed update.

## 4.4 Generative and supplanted FRLB loops

The FRLB cycle distinguishes two ideal-typical modes. Generative means corrigible, not correct; supplanted means that revisability has been revoked, not that the protected belief is false or that the person lacks intelligence. This corrigibility distinction draws on the update-integrity framework formalized in UIS (Jones 2026n).

- **Generative FRLB loops**
    - Faith is held as revisable and answerable to relevant evidence: starting assumptions remain open to correction under declared standards.
    - Reason is active and exploratory: the system entertains alternative explanations and counterfactuals.
    - Logic is used to test and constrain, not merely to defend.
    - Beliefs remain corrigible: matched evidence can change the task-bounded target or the retained commitment under declared admission rules.

In this mode, $K_{mind}$ remains plastic but structured. The system can revise protected and unprotected commitments across successive certified records and updates.

- **Supplanted FRLB loops**
    - Faith is absolutized: certain assumptions (about self, world, group, authority) are taken as non-negotiable.
    - Reason becomes instrumental: it searches for arguments and interpretations that support pre-existing commitments.
    - Logic is selectively applied: inconsistencies favorable to core Faith are tolerated or ignored.

- Belief updating becomes selectively insensitive: comparable disconfirming and confirming evidence produces directionally asymmetric change (Festinger 1957; Kunda 1990; Nickerson 1998).

Supplanting is detected at the update seam when relevant diagnostic records fail to produce proportionate change in a protected commitment, but it need not originate in the update map alone. Proximal causes may include protected constraints in K, restricted candidate search, corrupted or excluded records, counterfeit redundancy, asymmetric admission rules, or U itself. The unit of analysis is a commitment-specific loop, not a whole person; the claim concerns selective update, not intelligence or moral worth.

The distinction is structural and non-diagnostic. A generatively held belief can be false, and a supplanted belief can be true. Clinical, ideological, developmental, and ordinary cases require separate operational definitions and comparisons; FRLB should not be used to label persons from disagreement alone.

## 4.5 Meaning and intentionality as a functional account

FRLB supplies a functional vocabulary for how reportable contents acquire predictive, inferential, and action-guiding roles. It does not derive original intentionality from the kernel. Across repeated episodes:

- recurrent resolution patterns can stabilize as ways of organizing concepts, roles, and narratives;
- those regularities can become reusable constraint structures in $K_{mind}$, including schemas, frames, and categories;
- standing beliefs can then use those structures to interpret new input and anticipate later outcomes.

On this functional account, meaning is the predictive, inferential, and action-guiding role a pattern plays within a network of constraints and updates. This account does not settle every metaphysical question about intentionality.

- Within that limited account, a state's content is partly characterized by its place in a network of premises, search processes, and admission rules, together with reliable relations to inputs, expectations, and actions (Dretske 1988; Millikan 1984; Hohwy 2013).

This is a functional hypothesis about learned targeting within $K_{mind}$. It explains how a state can participate in stable world- and self-directed relations, but it does not establish that intentionality is exhausted by those relations or that FRLB is its unique engine.

The next section turns from internal resolution to action: how perceptual and FRLB outputs constrain behavior in $\Omega_{act}$, and how agency can be modeled as graded reasons-responsive constraint management rather than escape from prior causes.

## §5 Action, Agency, and Free Will as Behavioral Collapse

Perception and FRLB describe **internal** collapse: hypotheses, explanations, and beliefs resolving under constraints $K_{mind}$. But conscious beings do not only update their maps; they **move**. They act on the world in ways that can maintain or reshape their own coherence. In UCT language, action is collapse in a **behavioral possibility space** under constraints shaped by perception, FRLB, and bodily/environmental realities.

Free will, on this view, is not an escape from the collapse kernel. It is the graded capacity of a conscious system to represent, evaluate, and causally participate in revising selected constraints that shape later perceptual, cognitive, and behavioral resolutions.

### 5.1 The action space and constraints on behavior

At any decision point, an organism faces a structured space of possible behaviors: $\Omega_{act}$. For a human, this space is vast and hierarchically organized:

- immediate motor outputs (where to look, how to move),
- proximal actions (what to say next, whether to accept an invitation),
- long-horizon plans (whether to stay in a job, how to raise a child).

This action space is constrained by several layers:

- **Physical constraints** $K_{phys}$: body structure, environmental affordances, available tools.
- **Biological constraints** $K_{bio}$: needs, drives, metabolic limits, fatigue.
- **Mind-level constraints** $K_{mind}$: beliefs, values, goals, self-image, perceived norms.
- **Social constraints** $K_{soc}$: laws, expectations, obligations, roles.

Combined, these define:

$$K_{\text{act}} := \text{Compose}(K_{\text{phys}}, K_{\text{bio}}, K_{\text{mind}}, K_{\text{soc}})$$

Given a current perceptual resolution and $K_{\text{mind}}$, the organism does not evaluate every element of $\Omega_{\text{act}}$. Neural decision circuits—including basal-ganglia and prefrontal networks—can be modeled as realizing a provisional action-domain transition:

$$(x^*_{\text{act,t}}, q_{\text{act,t}}, S_{\text{act,t}}) \sim C^{K_{\text{act}}}_{\text{act,t}}(\omega_t, \cdot)$$

selecting a particular behavior or policy from among many possibilities (Redgrave, Prescott & Gurney 1999; Gold & Shadlen 2007).

## 5.2 Neural decision circuits as collapse machinery

Neuroscience suggests a concrete implementation of this picture. Across species, action selection appears to be mediated by:

- **competition** among candidate actions (e.g., in basal ganglia loops),
- **accumulation of evidence** or value signals toward thresholds (drift-diffusion models),
- **inhibitory gating** that disinhibits one motor program at a time.

In kernel terms:

- $\Omega_{\text{act}}$ is implicitly encoded in a set of competing action channels.
- $K_{\text{act}}$ is implemented by weights, thresholds, neuromodulation, and contextual inputs that bias which channels are favored.
- The dynamics of these circuits approximate a **stochastic search** over $\Omega_{\text{act}}$ that ends when one channel reaches a bound—collapse (Bogacz 2007; Gold & Shadlen 2007).

Access-conscious perceptual and FRLB outputs can contribute to $K_{act}$ by representing what the system currently classifies as desirable, acceptable, or forbidden. The winning action is therefore hypothesized to track current values, goals, and self-model constraints rather than being merely reflex-driven.

### 5.3 Agency as reasons-responsive constraint–action alignment

We typically ascribe **agency** when actions can be:

- explained in terms of the agent's beliefs, desires, and values, and
- predicted or influenced by changing those beliefs, desires, and values.

In the kernel, this corresponds to:

- a relatively stable $K_{\text{mind}}$ and $K_{\text{act}}$ that shape which regions of $\Omega_{\text{act}}$ are admissible and attractive, and
- a consistent mapping from changes in $K_{\text{mind}}$ (e.g., adopting a new value, recognizing a mistaken belief) to changes in $x^*_{\text{act}}$.

Agency is modeled as coherent, reasons-responsive coupling among:

- perception → $x^*_{\text{percept}}$,
- FRLB → $updated K_{\text{mind}}$,
- action selection → $x^*_{\text{act}}$,

such that the agent's actions track its own constraint architecture in a reasons-responsive way. Reported agency may be impaired when this coupling is disrupted, including in some compulsive, akinetic, or dissociative presentations; this is a correspondence hypothesis that must be compared with motor, clinical, and contextual alternatives, not a complete etiological account (Haggard 2008; Synofzik, Vosgerau & Newen 2008).

## 5.4 Free will as graded constraint management

Traditional debates about free will often imagine a choice between:

- strict determinism (no real choice), or
- metaphysical freedom (actions uncaused by prior states).

UCT offers a third angle. Within the collapse kernel, the question is not "are actions caused?" (they are) but:

To what extent can a system represent, evaluate, and causally participate in revising selected constraints under which its later collapses occur?

Free will, in this operational framing, is a graded capacity for reasons-responsive constraint management:

- to represent and reflect on accessible portions of $K_{mind}$ (beliefs, values, habits),
- to simulate alternative configurations in $\Omega_{mind}$,
- to formulate, rehearse, or enact candidate revisions to selected commitments, rules, or identities,
- and to let those interventions participate in later updates of $K_{mind}$ and $K_{act}$, altering how $\Omega_{act}$ resolves in future situations (Dennett 2003; Haggard 2008).

Nothing in this view violates physical law or requires neural indeterminism: the transition is typed as a stochastic kernel, with deterministic dynamics included as Dirac cases. The novelty lies in the system's **recursive organization**: conscious minds are life-phase networks that can:

1. represent some of their own constraints,
2. evaluate them against perceived coherence (internal and external), and
3. intervene on some represented constraints in ways that causally participate in later updates of $K_{mind}$ and $K_{act}$.

Under FRLB, a free-will exercise looks like:

- Faith: “I could live by this value instead of that one.”
- Reason: exploring consequences and counterfactuals.
- Logic: checking for consistency with other commitments and evidence.
- Belief: the episode's decision or commitment resolution—for example, deciding to stop a harmful habit, adopt a practice, or reinterpret an experience—and, where the associated candidate trace certifies, the resulting Event Belief record. U consumes the resolution and the Option-typed certification result and runs on either Some(r) or None. If a certified record is taken up through U, related content may become standing Belief or another later constraint.

Subsequent behavior remains caused and constrained by the updated architecture. The act of revising selected constraints is itself conditioned by earlier K, so this account does not create an uncaused chooser or imply transparent control of the full architecture. It offers a graded compatibilist proposal: agency increases when a system can represent some constraints, evaluate counterfactual alternatives, and revise later behavior in response to reasons and feedback.

**Empirical burden.** If explicit representation and deliberate intervention on a target constraint do not improve prediction of later behavior beyond matched non-conscious training or ordinary conditioning, the proposed distinct role of conscious constraint management loses empirical content.

## 5.5 S-signals in agency and habit

WP04 proposes candidate S-signature tests for agency and habit:

- **$S_1$ (redundancy → consensus):**
    - Repeated choices under similar constraints establish response reliability or policy stability, not $S_1$ by themselves.
    - A candidate agency-level $S_1$ test additionally requires independently informative reasons, cues, or records whose convergence predicts policy stability beyond correlated restatement, repeated exposure, or a matched single-source condition.
- **$S_2$ (neutrality → delayed resolution):**
    - In difficult decisions or conflicts of value, the system can linger in indecision: reaction times lengthen, ambivalence is felt, and conscious oscillation between options occurs.

- This reflects a near-neutral region in $\Omega_{act}$ where multiple actions remain similarly weighted under the declared constraints and small changes in evidence or internal weighting can tip resolution (Gold & Shadlen 2007).

- **$S_3$ (sweeps → hysteresis):**

  - Habits and commitments may exhibit hysteresis-like path dependence, but an audited $S_3$ claim requires a declared control parameter, forward and reverse sweeps, a state or outcome coordinate, loop-area or threshold comparison, and a sweep-rate audit.

  - For example, gradually increasing stress may push an agent toward avoidance while later stress reduction fails to restore prior engagement because associations, expectations, and self-models have changed within $K_{mind}$. This is a candidate $S_3$ placement; where record-driven hysteresis is claimed, the record-state relation must also be specified independently (Wood & Neal 2007; Ouellette & Wood 1998).

These patterns are candidate cross-domain analogues. Choices and actions may produce durable traces and modify later constraints, but not every action yields a certified record and not every record is taken up by the immediate update.

### 5.6 Bridge to learning, identity, and pathology

Understanding action as Behavioral Collapse under $K_{act}$ sets up the next layer of the mind-phase story. Over time:

- repeated actions can stabilize habits and self-models within $K_{mind}$;
- FRLB cycles entrench certain self-narratives and expectations,
- and the system moves into coherence pockets in $\Omega_{self}$: patterns of identity, style, and habit that are hard to leave (McAdams 2001; Wood & Neal 2007).

The next section turns to learning, memory, identity, and structural pathology as long-term dynamics of $K_{mind}$: how record-mediated updates support a self-model, how that model can become rigid or maladaptive, and how generative FRLB loops and therapeutic interventions may reshape later constraints.

## §6 Learning, Memory, Identity, and Structural Pathology

Perception, FRLB, and action describe **moment-to-moment** Conscious Collapse. Over longer timescales, these collapses accumulate into something we call a "self": a relatively stable pattern of beliefs, values, habits, and narratives that shapes how future collapses

unfold. In UCT language, this is simply the long-term evolution of $K_{\text{mind}}$ under repeated updates:

$$K_{\text{mind,t+1}} = U_{\text{mind}}(K_{\text{mind,t}}, x^*_{mind,t}, R_{mind,t}, S_{mind,t})$$

Learning and memory concern how the mind-domain update map uses a realized resolution, an Option-typed record, and the open carrier to produce later constraints. Identity is the emergent regularity of that evolving architecture, and structural pathology is a hypothesis about rigid or misaligned update.

## 6.1 The self as record-mediated constraint architecture

At any given time, a person's mind-phase can be described by:

- a **constraint architecture** $K_{\text{mind}}$ capturing:
    - beliefs and models of the world,
    - values and priorities,
    - self-concepts and narratives,
    - procedures and habits for reasoning and acting;
- a history of realized resolutions and certified records that have shaped later K.

The self is modeled here as a relatively stable pattern in the evolving constraint architecture rather than an extra entity behind it (Conway & Pleydell-Pearce 2000; McAdams 2001). Self-descriptions refer to regularities in which perceptual, conceptual, and action outcomes are repeatedly favored under the current architecture.

Identity thus has two faces:

- **Structural:** a constraint field over possible perceptions, thoughts, and actions.
- Phenomenological: a reported continuity of 'me-ness' often associated with operation inside a familiar self-model; the structural account does not yet explain differences in that continuity or validate correspondence between self-report and modeled architecture (James 1950 [1890]; Gallagher 2000).

From the kernel's perspective, identity is modeled as record-mediated constraint architecture: prior resolutions, certified records, and updates shape later perception, action,

and self-narrative. The structural model does not by itself specify how changes in that architecture map to changes in experienced self-continuity.

## 6.2 Learning and memory as record-mediated update

Learning and memory are the mechanisms by which $K$ changes in response to experience. In the FRLB update rule:

$$B_{t+1} = B_t + \alpha_t(\hat{B}_t - B_t)$$

At the standing-Belief grain, B is a task-defined slice of $K_{mind}$: a belief or belief-cluster. Selected neural, behavioral, and psychological learning trajectories may exhibit update patterns that this ansatz approximates at an appropriate observation grain. The ansatz is not asserted as the microscopic synaptic rule or as the universal form of learning. Candidate mechanisms and outcomes include:

- synaptic plasticity: changing neural weights and connectivity in response to activity patterns (Hebbian learning, error-driven learning) (Hebb 1949; Kandel 2001; Squire & Kandel 2009);
- **associative learning:** strengthening or weakening links between stimuli, responses, and outcomes;
- **schema formation:** building higher-order structures (concepts, scripts, frames) that summarize many experiences (Rumelhart & Norman 1978; Ghosh & Gilboa 2014);
- narrative integration: weaving experienced episodes into self-referential stories ('this is what happened to me,' 'this is what it means') (Conway & Pleydell-Pearce 2000; McAdams 2001).
- These grains must not be collapsed. Neural plasticity is modeled through $U_{neural}$ acting on $K_{neural}$; schema, standing-Belief, and narrative revision are modeled through $U_{mind}$ acting on $K_{mind}$. WP04 assumes that these levels are materially coupled but does not yet supply a complete cross-scale map from neural record and update to standing cognitive constraint.

Each modeled mind-domain learning transition can be represented as an update:

$$K_{\text{mind,t+1}} = U_{\text{mind}}(K_{\text{mind,t}}, x^*_{mind,t}, R_{mind,t}, S_{mind,t})$$

where:

- $x_{\text{mind,t}}^*$ is a realized mind-domain resolution, which may or may not be access-conscious depending on the declared phase criterion;
- A certified record may contain an internal or external payload relevant to the declared update. Candidate examples include memory traces, notes, consequences of action, and feedback from others, but recordhood is not assumed from the example alone;
- U is grain-specific: $U_{\text{neural}}$ updates neural constraints, while $U_{\text{mind}}$ updates conceptual or standing-Belief architecture under the declared model.

Over time, repeated updates can create attractor-like regularities in $K_{\text{mind}}$: beliefs, habits, and identity patterns that are more readily reinstated and harder to dislodge. 'Coherence pocket' is a dynamical metaphor here; a domain model must specify the state variables, distance metric, and stability criterion before treating it as a measured object.

### 6.3 Healthy coherence: generative update and flexible identity

In a healthy mind-phase, the FRLB loop and learning dynamics produce **generative coherence**:

- Perception and FRLB remain in contact with external constraints (physical, social, biological).
- *K* updates in ways that preserve internal consistency *and* improve fit with reality.
- Identity remains **structured but flexible**:
    - stable enough to provide continuity and orientation,
    - plastic enough to accommodate new information, changing roles, and growth (McAdams 2001; West-Eberhard 2003).

Structurally, this corresponds to:

- deep but **not over-narrow** basins in $\Omega_{\text{self}}$: the self can absorb perturbations without shattering, but can also move to neighboring states when pressure and evidence demand it;

- FRLB loops that remain corrigible under matched confirming and disconfirming evidence and retain stable, inspectable admission standards;

The proposed healthy regime is structured but corrigible: stable enough for continuity, yet responsive to relevant evidence and changing conditions. 'Coherence' must be disaggregated into internal consistency, predictive accuracy, dynamical stability, and well-being rather than treated as one scalar good.

## 6.4 Structural pathology as update rigidity and misalignment

**Scope caution.** Structural pathology is a hypothesis about update rigidity or misalignment, not a clinical diagnosis, a theory of etiology, or a substitute for established assessment. Trauma, embodiment, neurodevelopment, environment, social power, and biological disease can shape the same observed behavior through different mechanisms. The examples below are candidate research mappings only.

Candidate forms of structural pathology are modeled here as cases in which $K_{\text{mind}}$ is configured so that resolution repeatedly undermines coherence—internally, externally, or both. Within the FRLB framework, this can take several forms:

- **Rigid Faith:**
    - Core assumptions F are fixed as non-negotiable ("I am worthless," "others cannot be trusted," "my group is always right").
    - Reason and Logic are recruited to defend these assumptions, not to test them.
    - Effective update is selectively insensitive to relevant disconfirming evidence (Beck et al. 1979; Beck & Clark 1997).
- **Distorted Logic:**
    - Inferential rules L are biased or broken (catastrophizing, overgeneralization, black–white thinking).
    - Even accurate perceptions and adequate, non-distorting starting premises cannot produce coherent standing beliefs when the operative admission rules are systematically biased (Beck 1976; Burns 1999).
- **Maladaptive Belief architecture:**
    - Beliefs B organize $\Omega_{\text{self}}$ and $\Omega_{\text{world}}$ into patterns that generate chronic threat, shame, distrust, or helplessness.

- Action collapse in $\Omega_{act}$ repeatedly produces behaviors that reinforce these beliefs (self-fulfilling loops) (Clark & Beck 2010).

In each candidate case, the mind-phase is modeled within the proposed kernel schema. Because no complete mind-domain certification and dependence test is yet supplied, WP04 does not claim empirical compliance merely from the qualitative mapping. The proposed problem is a misaligned or selectively insensitive update architecture:

- with external reality (physics, biology, social feedback),
- or with internal needs for viability and well-being.

Some symptom trajectories may exhibit hysteresis-like persistence, but any FRLB interpretation must be compared with disorder-specific, developmental, biological, and contextual models rather than inferred from persistence alone (Ingram, Miranda & Segal 1998; Hayes, Strosahl & Wilson 2011).

Human Interface Laws develops the complementary interface-level account of how bandwidth, signal weighting, identity protection, record retention, self-sealing, and reopening can shape these trajectories (Jones 2026h).

### 6.5 Structural Ethic: an explicit normative extension

A Structural Ethic does not follow from the descriptive kernel alone. It adds value premises: preserve corrigibility, truth-sensitive contact, well-being, and multi-scale viability while recognizing that these goods can conflict. The resulting proposal is:

At the normative level, agents can cultivate practices that participate in revising selected constraints within $K_{mind}$ through ordinary learning, deliberative, social, and embodied update processes. The declared aim is improved coherence across scales—internal consistency, fit with external reality, and alignment with the needs of the broader systems they inhabit (relationships, communities, ecosystems).

These are declared normative commitments, not consequences of mere persistence or internal consistency:

- Generative FRLB loops that remain answerable to relevant records and stable admission standards can move K toward better predictive or relational performance, subject to the value criteria declared for the application.

- Supplanted loops may be associated with poor predictive, relational, or well-being outcomes when selective update insensitivity persists. The descriptive kernel alone does not derive suffering, health, or breakdown.

Therapeutic and reflective practices—cognitive behavioral therapy, mindfulness, philosophical inquiry, honest conversation, scientific reasoning—can be understood as practices intended to influence selected constraints through ordinary cognitive and behavioral pathways: structured attempts to:

- surface hidden Faith and distorted Logic,
- examine them in light of candidate search, declared admission rules, relevant evidence, and candidate records under a declared standard,
- and support revisions that may be taken up into selected portions of $K_{mind}$, shaping later resolution toward the declared aims of the practice (Beck et al. 1979; Kabat-Zinn 1990; Hayes et al. 2011).

UCT therefore recommends an update discipline rather than a prescribed worldview: keep important commitments answerable to relevant records, make admission rules inspectable, and protect the conditions under which revision is possible. Detailed ethical justification belongs in companion work. This update discipline follows the operational governance developed in UIS (Jones 2026n).

### 6.6 Bridge to intersubjective mind and CIM

So far, the paper has treated learning, memory, identity, and update rigidity within one mind. Cognition also becomes durable outside the originating episode. Speech, writing, tools, institutions, and digital traces can enter a record-bearing layer that constrains later minds. Section 7 follows CIM Foundational as the category authority (Jones 2026e).

The next section turns to this **intersubjective mind**: how individual Conscious Collapse interacts with shared symbols, cultural constraints, and institutional structures. There we will treat CIM as the medium through which mind-phase coherence extends beyond the skull into social and technological systems, and through which those systems, in turn, shape the constraints under which future minds collapse.

## §7 Intersubjective Constraint and Consciousness-Induced Material (CIM)

Human cognition is coupled through durable records. Cognition-Induced Collapse (CIC) names the broader process class, including candidate outputs of experience-bearing

cognition; Consciousness-Induced Material (CIM) names a bounded record-bearing output layer shaped by conscious cognition. CIM begins when such cognition is stabilized in physically instantiated, record-bearing form beyond the originating interior episode and can non-trivially constrain later cognition. Interior experience is not yet CIM, and not every artifact a mind touches qualifies.

For WP04, a candidate qualifies as CIM when four conditions are explicit:

it is shaped by conscious cognition; physically instantiated in record-bearing form; stabilized beyond the originating interior episode; and capable of non-trivially constraining later cognition. Externalized CIM can later be internalized through learning; primary and derivative CIM preserve different cognitive lineages.

CIM supports coupled intersubjective dynamics through language, notation, tools, media, and institutions. This does not establish a collective phenomenal subject or a societal mind (Hutchins 1995; Clark & Chalmers 1998; Tomasello 1999).

CIM Foundational supplies the category architecture: CIM as record-bearing output; externalized and internalized forms; primary and derivative lineages; Synthetic Collapse operating on accumulated CIM; and Conscious Synthetic Collapse as an open conscious-phase realization question (Jones 2026e). WP04 uses that architecture and adds only a provisional kernel mapping.

## 7.1 CIM as a record-bearing output layer

Let $\Omega_{ext}$ denote possible externalization outcomes, and let $K_{ext} := \text{Compose}(K_{mind}, K_{soc})$ denote the combined constraints on the externalization act. These are analytic spaces and constraints; CIM itself remains the qualifying record-bearing output layer.

- Candidate external structures include:
    - spoken utterances, written texts, diagrams;
    - tools, buildings, user interfaces;
    - social rules, laws, rituals, protocols;
    - digital artifacts: code, models, datasets, networks.
- $K_{soc}$ denotes the social and technological constraint architecture that shapes production, persistence, circulation, and uptake:
    - existing languages, symbols, and media;
    - institutions and legal frameworks;
    - economic and technological infrastructures;

- cultural norms about what may be said, built, or enforced (North 1990; Ostrom 1990; Henrich 2015).

A coupled mind-to-externalization-domain transition can co-produce a candidate trace $q_{CIM,t}$:

$$(x^*_{ext,t}, q_{CIM,t}, S_{ext,t}) \sim C^{K_{ext}}_{mind \to ext,t}(\omega_t, \cdot)$$

Here $x_{ext,t}^*$ is the realized externalization act or product and $q_{CIM,t}$ is its candidate trace. This is a coupled mind-to-externalization placement within the shared schema, not a second universal operator, a distinct CIM operator, or a new conscious substrate.

Certification may return $R_{CIM,t}$ = Some($r_{CIM,t}$) when the four CIM category conditions above are satisfied and the declared trace-quality and provenance/dependence tests pass. If later uptake occurs, social or individual constraints may update:

$$K_{soc,t+1} = U_{soc}(K_{soc,t}, x^*_{ext,t}, R_{CIM,t}, S_{ext,t})$$

Over time, this produces relatively stable social and technological coherence pockets: languages, legal systems, scientific paradigms, platforms, and cultural traditions (Kuhn 1962; Tomasello 1999; Henrich 2015).

## 7.2 From thought to word to world

At the individual level, the pipeline looks like:

1. Internal resolution
    - Perception and FRLB produce an access-conscious resolution $x_{mind,t}^*$: a thought, judgment, intention, or plan.
2. Externalization
    - Under $K_{ext}$, the agent realizes an externalization $x_{ext,t}^*$ and co-produces a candidate trace $q_{CIM,t}$ through speech, writing, design, or action in the built or digital environment.

3. Certification and distribution
    - When the declared CIM and kernel tests succeed, certification returns $R_{CIM,t}$ = Some($r_{CIM,t}$). The durable output can then be perceived, interpreted, and redistributed by other agents.
4. **Feedback into *K***
    - As others encounter and respond to the externalized record, their own $K_{mind}$ and shared $K_{soc}$ may update; norms and infrastructures can also shift (Hutchins 1995; Clark & Chalmers 1998).

The same thought—say, "we should treat disease X this way"—can, via CIM, become:

- a published paper,
- a clinical guideline,
- a training program,
- an institutional policy,
- a cultural expectation.

Each is a candidate externalization and potential CIM record that can constrain what clinicians expect, what patients request, and what regulators require. Qualification and later uptake must be tested rather than presumed.

Externalization is not a neutral transcript of a settled interior state. Selection, phrasing, medium, audience, institutional rules, and expected response further constrain what becomes durable. This second transition should be studied directly rather than identified with constitutive faith or treated as a transparent copy of private content.

## 7.3 Intersubjectivity as coupled levels

Intersubjectivity is modeled as coupled dynamics among individual minds and shared records. Three analytic levels are useful without reifying a collective experiencer:

1. **Individual minds**
    - Each has its own $K_{mind,i}$ and performs perception, FRLB, action, and learning.

2. Record-bearing CIM layer
    - Qualifying externalized and internalized records produced by cognition and governed by persistence, circulation, and uptake constraints.
3. Societal and institutional coordination
    - Coherence patterns in how populations of agents and CIM interact:
        - scientific communities,
        - markets,
        - legal systems,
        - cultural traditions,
        - online platforms (Kuhn 1962; North 1990; Ostrom 1990; Henrich 2015).

Each level may be modeled by a domain-appropriate kernel within the shared schema, and the levels are materially coupled:

- individual minds resolve thoughts and actions that may generate candidate CIM traces;
- CIM constrains what minds can easily think, say, or do;
- institutions (patterns in CIM + collective behavior) shape both $K_{\text{soc}}$ and $K_{\text{mind}}$ through incentives, norms, education, and enforcement.

The result is a **multi-level coherence pocket** in which mind-phase dynamics and CIM co-evolve. For example:

- legal concepts of personhood and responsibility feed back into how individuals model themselves and others;
- scientific frameworks constrain how data are interpreted and which hypotheses are taken seriously;
- recommendation algorithms and interfaces bias what content is seen and how attention collapses across populations (Pariser 2011; Sunstein 2017).

Individual conscious access is therefore embedded in social and technological constraints, but embeddedness does not establish group consciousness.

## 7.4 CIM, S-signals, and structural inertia

CIM-mediated social dynamics may exhibit candidate S-signatures when the required variables and audits are declared:

- **$S_1$ (redundancy → consensus):**
    - When genuinely independent observers and methods converge on compatible durable records, communities may become more likely to converge on a shared model. The relevant $S_1$ test must audit channel independence rather than count repeated copies of one source.
    - Where an independence audit passes, genuinely distinct channels can reinforce the same coherence pocket in $K_{soc}$; papers, lectures, or media that merely repeat one source remain amplification rather than independent redundancy (Kuhn 1962; Merton 1973).
- **$S_2$ (neutrality → delayed resolution):**
    - Controversies and paradigm disputes may persist when declared evidence and incentives leave competing models similarly supported; this is a candidate $S_2$ placement only after the neutrality assumptions are audited.
    - Collapse into a new consensus may be delayed until decisive constraints (new data, crises, conceptual innovations) tip the landscape (Kuhn 1962; Lakatos 1978).
- **$S_3$ (sweeps → hysteresis):**
    - Social and institutional systems may exhibit hysteresis-like path dependence: once norms, platforms, or paradigms are entrenched, simply reversing a declared driver (for example, a temporary crisis or leadership change) may not restore the previous state.
    - Past CIM and accumulated $K_{\mathrm{soc}}$ (infrastructure, laws, habits) continue to shape collapse trajectories, making change path-dependent (Mahoney 2000; Pierson 2000; North 1990).
    - A formal $S_3$ test still requires a declared control parameter, forward and reverse sweeps, an outcome coordinate, loop-area or threshold comparison, and a sweep-rate audit; where record-driven hysteresis is claimed, the record-state relation must be specified independently.

These are candidate analogues to physical and biological signatures. Consensus is not automatically truth, controversy is not automatically neutrality, and institutional persistence is not automatically a measured hysteresis loop; each claim requires declared variables and alternative models.

### 7.5 Mind, CIM, and responsibility

Because CIM both results from and shapes Conscious Collapse, it is also where **responsibility** becomes structurally meaningful:

- Individual choices about what to say, build, and amplify can generate candidate CIM traces with long-lived effects on $K_{soc}$ and on later minds.
- Structural pathologies in individual $K_{\text{mind}}$ can propagate through CIM into collective pathologies (misinformation, harmful norms, brittle institutions).
- Conversely, generative FRLB loops and CIM aligned with the declared aims of truthful, transparent, and agency-preserving coordination can scaffold better mind-phase coherence for others (Habermas 1984; Sunstein 2017; Pariser 2011).

Responsibility enters as a declared normative extension: agents and institutions differ in their leverage over the record and constraint environments others inhabit. The kernel can trace those dependencies, but it does not derive the moral weighting by itself.

## §8 Conscious Collapse in the Landscape of Theories

Consciousness theories operate at different explanatory levels: functional access and broadcasting, recurrent and higher-order processing, predictive modeling, information integration, and proposed microphysical mechanisms. WP04 is not a replacement implementation theory. It offers a structural mapping whose value depends on restrictions, discriminators, and successful transfer—not on redescribing each theory in kernel vocabulary.

This section briefly situates UCT's mind-phase relative to major existing approaches.

## 8.1 Functional architectures: Global Workspace and Predictive Processing

Global Workspace Theory (GWT) and related global-broadcast models propose that a mental state becomes conscious when it achieves **global availability**: it is "posted" to a workspace that many specialized subsystems can read and write (Baars 1988; Dehaene & Changeux 2011; Dehaene 2014). Predictive-processing (PP) and active-inference frameworks treat the brain as a hierarchical prediction engine: perceptions and actions minimize prediction error under a generative model (Rao & Ballard 1999; Friston 2010; Clark 2013).

UCT is compatible with important elements of GWT/GNWT and predictive processing, but it treats them as candidate implementations rather than settled neural identities:

- In GWT terms, Conscious Collapse says:
    - $\Omega_{\mathrm{GW}}$ is the space of globally broadcast hypotheses;
    - $K_{\mathrm{mind}}$ includes priors, goals, and workspace anatomy;
    - conscious content $x^{*}_{\mathrm{percept}}$ is the hypothesis that has **won collapse into the workspace** under $K_{\mathrm{percept}}$ (Section 3).
    - Global broadcasting is one candidate implementation of conscious access. Recurrent-processing and higher-order theories supply alternatives or additional conditions, and predictive processing by itself is not a complete theory of consciousness.

**Current evidence.** A preregistered 2025 adversarial collaboration found some results aligned with IIT and GNWT while substantially challenging key predictions of both, including the expected GNWT offset ignition and broad prefrontal representation of several content dimensions (Cogitate Consortium et al. 2025). WP04 therefore treats workspace broadcasting and integration as contestable architecture-level hypotheses, not as established proof of consciousness or of UCT.

Series evidence. A pre-specified UCT-affiliated reanalysis of the open COGITATE iEEG data tested three narrower constraint-architecture predictions. Task-relevant non-targets showed later high-gamma onset than the same category when irrelevant (25 of 34 patients in the predicted direction; median +14.1 ms; p = .007); a predicted cross-block hysteresis effect was null (p = .748); and task modulation of duration-tracking was inconclusive at the subject level (Jones 2026f). This mixed result supplies direct evidence that task constraints can shape resolution timing while limiting stronger persistence claims. It does not establish UCT, identify conscious-phase membership, characterize experiential differences, or validate cross-domain transfer.

- In PP terms, Conscious Collapse says:
    - perception is **hypothesis collapse** in $\Omega_{\text{percept}}$ with constraints $K_{\text{percept}}$, as we already wrote;
    - conscious percepts are those prediction-error resolutions that reach the workspace regime;
    - FRLB operates on sequences of resolutions, certified records, and updates to $K_{\text{mind}}$.

GWT/GNWT and predictive-processing models specify candidate mechanisms at a finer neural or computational level. UCT earns explanatory value only if its additional restrictions predict something those models and generic dynamical-systems descriptions do not.

## 8.2 Information-based views: Integrated Information Theory

Integrated Information Theory (IIT) starts from phenomenological axioms and proposes that consciousness corresponds to **maximally integrated information structures**, quantified by measures like Φ (Tononi 2004; Tononi 2008; Tononi et al. 2016). It suggests that systems with high Φ have rich consciousness; low-Φ systems have little or none.

UCT's alignment with IIT is partial:

- **Agreement:**
    - Conscious systems occupy a regime of high **integration and differentiation** (Section 2.2).
    - Measures such as Φ and perturbational complexity quantify different proposed properties and should not be collapsed into a single measure of 'coherence.' WP04 treats them only as candidate indicators to be compared against task, report, and architecture-specific alternatives.
    - Integration measures and workspace availability can covary in some paradigms, but their relationship is empirical rather than definitional and cannot by itself establish conscious-phase membership or an episode's experiential profile (Casali et al. 2013; Cogitate Consortium et al. 2025).
- **Divergence:**
    - UCT does not treat Φ (or any one information measure) as *defining* consciousness. It treats such measures as **indicators** of a certain structural regime in which collapse produces globally integrated contents.

- UCT does not endorse panpsychism or the claim that any system with nonzero Φ is conscious. WP04's current mind-phase hypothesis proposes a regime in which selected resolutions become available as content to the system itself. That is a defeasible phase-identification proposal, not a generic-kernel condition; very simple or weakly structured systems may fail it even if some information measure is nonzero.

IIT and perturbational-complexity approaches may supply candidate indicators of integration and differentiation, but WP04 does not reduce them to generic 'coherence pockets' or treat nonzero Φ as consciousness by fiat.

### 8.3 Quantum collapse theories: Orch-OR and beyond

Several proposals tie consciousness directly to **quantum collapse** at the microphysical level—for example, Orch-OR (orchestrated objective reduction), which suggests that conscious moments correspond to orchestrated collapses in microtubules (Hameroff & Penrose 1996; Hameroff & Penrose 2014).

UCT's notion of "collapse" is deliberately more general:

- It is a structural resolution schema, not necessarily a quantum transition:
  - UCT uses a declared transition kernel over a typed domain; it is not a synonym for every instance of selection, threshold crossing, or optimization.
  - The schema can be instantiated at macroscopic scales, including phase transitions and neural dynamics, as well as in proposed microscopic applications.
- UCT remains **agnostic** about whether microphysical collapse has a special role in consciousness beyond what standard decoherence and neural dynamics already provide. From the kernel's standpoint:
  - if quantum effects matter for consciousness, a mind-domain instantiation must show how they alter the declared neural state, constraints, traces, or update dynamics rather than merely note that quantum transitions occur;
  - but there is no need to **identify consciousness with quantum collapse itself** (Bassi et al. 2013).

Conscious Collapse does not deny that quantum transitions occur in brains. It says that the mind-phase proposal concerns how structured potential and constraints are organized at neural and cognitive scales, where global availability, FRLB, and CIM become the relevant constructs.

## 8.4 UCT's proposed contribution and its burden

UCT's proposed contribution is a typed family of domain kernels sharing a restricted schema and signature family. The claim is not that one concrete operator runs over unlike physical, biological, and cognitive state spaces.

- Mind, biology, and physics are modeled with domain-specific kernels that share the same architecture and are hypothesized to share testable restrictions and invariants.
- Conscious Collapse is proposed as an access-level regime in which selected resolutions become globally available and recursively usable, including the capacity to participate in later control, memory, and constraint update.
- FRLB supplies a provisional functional-role decomposition and a heuristic update ansatz; it is not yet a validated law of belief or identity change (Jones 2026g, How Minds Resolve).
- CIM supplies a bounded category for physically instantiated, record-bearing outputs shaped by conscious cognition that can non-trivially constrain later individual and collective dynamics without implying a collective phenomenal subject (Hutchins 1995; Clark & Chalmers 1998).

UCT also advances candidate law-level predictions that cut across theories:

- UCT hypothesizes that audited S-signatures ($S_1$, $S_2$, $S_3$) will survive appropriately controlled tests in perception, cognition, behavior, and intersubjective dynamics; qualitative resemblance alone does not qualify as a signature result.
- Candidate consciousness-threshold measures should track preregistered shifts in whether neural resolutions become globally available and recursively usable, including whether they become available to participate in later $K_{mind}$ update through U—not merely raw information content (Casali et al. 2013; Massimini et al. 2005; Dehaene 2014).
- Interventions intended to influence selected portions of $K_{mind}$ (therapy, reflection, education) should exhibit predictable patterns in how resolution trajectories and identity basins change over time (Beck et al. 1979; Hayes et al. 2011).

Beyond classical architectures, quantum probability models of cognition have also been proposed to capture certain decision anomalies (Pothos & Busemeyer 2013; Yearsley & Busemeyer 2016). UCT is agnostic about whether the correct microformalism for mind is classical or quantum; it focuses on the shared selection structure of collapse under constraint.

Where GWT/GNWT and predictive processing describe candidate functional architectures, IIT offers an information-centered theory, and quantum models explore microphysical proposals, UCT supplies a cross-domain research program. WP04 does not demonstrate the proposed shared restrictions; it specifies what would count as incremental support or failure.

### 8.5 The classical hard problem and the remaining explanatory burdens

The paper's operational target is access consciousness: content available for reasoning, report, flexible action, memory, and system-level control (Block 1995). The global-availability hypothesis and MIND-1 predictions live at this level.

On the classical production formulation targeted here, an ontologically exhaustive third-person description is assumed to omit phenomenal presence, which is then posed as an additional fact that the physical or functional process must produce (Chalmers 1995). UCT does not claim to solve that problem on those terms; it declines the primitive partition on which the production question depends. Its non-bifurcation thesis is that first-person presence and third-person neural organization are non-interchangeable modes of access to one conscious occurrence.

That proposal does not settle the empirical or phenomenological work. Four burdens remain: (1) phase identification—what distinguishes genuine conscious availability from sophisticated non-conscious processing; (2) episode structure—what establishes unity, boundary, duration, and cessation; (3) phenomenological differentiation—why different constraint architectures correspond to different experienced contents; and (4) correspondence—how first-person evidence, behavior, and third-person neural measures constrain one another without treating any one description as exhaustive. A completed account also owes structural completeness: if phenomenally different systems can remain identical on every declared mind-domain variable, the mapping lacks relevant structure.

The inferential directions must remain separate. Successful access-marker prediction would support the proposed phase mapping, not prove the non-bifurcation thesis; the philosophical appeal of the non-bifurcation thesis would not validate a neural marker. WP04 therefore neither announces that consciousness is fully explained nor leaves phenomenality as an unrelated add-on. It rejects the classical production problem while keeping the identification, differentiation, and correspondence problems open.

## §9 Predictions and Falsifiability (Mind)

Conscious Collapse makes structural claims about how conscious and non-conscious processes differ, how belief and identity update, and how intersubjective structures emerge. Those claims can be probed empirically using the S-signature audit protocols deposited as standalone methods (Jones 2026a, 2026b, 2026c).

For executable orientation to the three S-signatures, see the UCT Mind Starter Pack v1.0 (Jones 2026m). The pack is a runnable illustration and reproducibility aid, not evidence for the kernel or the cross-domain law claim. Neither local marker success nor cross-domain transfer would by itself establish the non-bifurcation thesis; those tests bear on the structural mapping.

This section sketches classes of tests, not completed protocols. Generic fit to C + U is not evidence because that architecture is extremely flexible. The burden falls on preregistered restrictions, effect directions, measurement mappings, transfer criteria, and comparisons against domain-native null models. This empirical discipline follows the Standards-layer requirement to predeclare restrictions and expose divergence claims to domain-native comparison (Jones 2026j).

## 9.1 Neural signatures of Conscious Collapse: MIND-1

If the workspace phase-identification model is useful, carefully matched conscious and non-conscious conditions should differ in preregistered patterns of global availability, recurrence, and complexity.

**Prediction 1 — Consciousness thresholds and integrated dynamics.**

Perturbational complexity (e.g., TMS–EEG PCI-like measures) should:

- be comparatively high in selected states where responses remain spatially differentiated and temporally integrated, subject to state- and method-specific exceptions;
- drop in states marked by substantially reduced effective complexity, differentiated recurrent coordination, or content-specific global availability, including deep anesthesia, selected non-REM conditions, and generalized seizures;
- recover when the system returns to a regime where global broadcasting and FRLB-style updating are possible (Massimini et al. 2005; Casali et al. 2013; Dehaene 2014).

A MIND-1 style test would:

1. Perturb the system (brief stimulation).
2. Measure the **spatiotemporal richness** and spread of the response.
3. Compare conditions (awake vs anesthetized vs minimally conscious).

The phase-identification hypothesis predicts systematic regime changes in global availability, recurrence, and complexity, but no single metric is expected to be necessary and sufficient across every state. Failure of preregistered multivariate classifiers and causal perturbation tests to distinguish carefully matched conscious and non-conscious conditions would undermine the current neural framing.

**MIND-1T — Sustained-resolution temporal discriminator.**

If the sustained-resolution hypothesis is useful, access-conscious contents should exhibit a preregistered interval of temporal persistence and recurrent causal influence beyond matched non-conscious processing. Tests must separately match or control attention, report preparation, ordinary evidence accumulation, and post-resolution maintenance, and must assess whether persistence and recurrence add held-out predictive value beyond established alternative accounts.

The hypothesis would be weakened if conscious access systematically occurs without the proposed interval, if the interval tracks the matched control processes rather than content availability, or if the added measures yield no incremental predictive gain. MIND-1T is a candidate temporal mechanism, not the definition of consciousness.

## 9.2 Perception and decision: S-signals in the mind-phase

Perception and decision-making are modeled as resolutions in $\Omega_{\text{percept}}$ and $\Omega_{\text{act}}$ under structured constraints. Their S-signatures are candidate parallels to those in physics and biology, not yet evidence of a common law.

**Prediction 2 — $S_2$ in ambiguity and conflict.**

In conditions of perceptual or decisional ambiguity:

- **reaction times** should be longer,
- measures of uncertainty or vacillation (e.g., confidence ratings, pupil dilation, EEG markers of conflict) should increase,
- and conscious content may **alternate between attractors** (bistable perception, indecisive "back and forth" in thought) (Leopold & Logothetis 1999; Sterzer et al. 2009; Devia et al. 2022).

This is a candidate mind-phase $S_2$ placement: neutrality → delayed resolution in $\Omega_{percept}$ or $\Omega_{act}$. A valid test must independently declare the neutrality or asymmetry variable, predict latency direction, and control relevant alternatives. Systematic violations under those conditions would undercut the proposed placement.

**Prediction 3 — $S_1$ in redundancy and consensus.**

When independently informative sensory channels or records are redundant and constraints are strong:

- repeated presentations under similar $K_{percept}$ should establish response reliability but count as amplification unless the presentations supply independently informative records;
- multimodal cues should converge on the same interpretation;
- across individuals with similar training and context, there should be strong **agreement** about what is seen or judged (Ernst & Banks 2002).

Here, an $S_1$ claim requires an independence audit: repeated or correlated copies of one source do not become independent evidence by multiplication. If independently informative channels converged no more reliably than matched single-channel or correlated-copy conditions, the proposed $S_1$ placement would fail.

## 9.3 FRLB and belief updating: task-bounded update tests

FRLB proposes structured, task-bounded differences in how premises, search, admission rules, and retained commitments relate to update. The claim remains provisional until the constructs can be coded reliably and add prediction beyond established models.

**Prediction 4 — Structured belief change under targeted interventions.**

In preregistered studies that explicitly define the unit of inquiry and the FRLB coding protocol, interventions such as structured education, reflective practice, or psychotherapy may test:

- identifiable shifts in $K_{mind}$ (reported beliefs, values, self-concepts) that follow **gradual update trajectories**, not unexplained jumps;
- tighter alignment between the task-bounded licensed target and retained belief in generative loops, as defined by preregistered task measures;

- direction- and context-sensitive effective update rates when focal starting premises or admission rules are selectively protected (Beck et al. 1979; Beck & Clark 1997; Hayes et al. 2011).

In practice, this means:

- tracking explicit belief ratings, behavior patterns, and self-report over the course of an intervention;
- modeling trajectories with update-like dynamics (e.g., exponential approaches to new baselines) rather than arbitrary stochastic changes.

FRLB would be weakened if its coding is unreliable, if generative and supplanted patterns do not predict held-out update trajectories, or if standard measures explain the same outcomes without incremental contribution.

Prediction 5 — Update-rigidity profiles (research hypothesis).

If FRLB captures a useful update dimension, preregistered rigidity profiles may covary with selected symptom or behavior trajectories after controlling for diagnosis-specific, developmental, biological, and contextual factors. WP04 does not predict a one-to-one FRLB signature for any disorder.

- candidate measures should separately assess:
    - stability or rigidity of task-defined starting premises,
    - asymmetry in candidate search or counterexample generation,
    - selective or shifting application of declared admission rules,
    - direction-specific effective update rates for matched evidence.

Where an intervention changes a targeted update process, FRLB measures should change in the preregistered direction and improve held-out prediction of relevant behavior or symptom trajectories. Null or non-incremental relationships would weaken the construct claim (Beck 1976; Clark & Beck 2010; Hayes et al. 2011).

## 9.4 CIM and intersubjective mind: collective predictions

At the intersubjective level, WP04 proposes that CIM-mediated systems may exhibit operational analogues of $S_1$–$S_3$. These are hypotheses about declared social variables, not proof that institutions instantiate the same concrete process as physical systems.

Prediction 6 — Audited convergence and unresolved competition in CIM-mediated systems ($S_1/S_2$).

In domains where:

- evidence is strong and multi-sourced, and
- incentives and $K_{\text{soc}}$ favor accuracy (e.g., mature natural sciences),

the preregistered hypothesis is stronger convergence ($S_1$) than in matched domains with less independent evidence or weaker accuracy incentives, after auditing channel independence and alternative social mechanisms (Kuhn 1962; Merton 1973).

Conversely, in domains where:

- evidence is sparse or equivocal, or
- incentives misalign (identity, status, power),

we expect **persistent neutral zones** ($S_2$): competing frameworks coexisting without resolution, prolonged debates, and sensitivity to small perturbations in context or authority (Lakatos 1978; Mahoney 2000).

If scientific and cultural histories showed no such pattern—if mature, high-evidence domains were as contentious as low-evidence ones, and vice versa—even after controlling for obvious social factors, the collapse-based reading of intersubjective dynamics would be weakened.

**Prediction 7 — Hysteresis in social/institutional change ($S_3$).**

Social systems and institutions are candidate $S_3$ placements and may show hysteresis-like path dependence:

- once a norm, platform, or paradigm is entrenched, simply reversing a driving parameter (e.g., removing a law, changing leadership, adjusting incentives) will often **not restore** the prior state;
- instead, durable records, infrastructures, habits, and narratives may keep the system in the new basin until stronger or qualitatively different interventions are applied (North 1990; Pierson 2000; Scheffer 2009).
- A preregistered $S_3$ test must declare the control parameter, forward and reverse sweeps, state or outcome coordinate, loop-area or threshold comparison, and sweep-rate audit; any record-state relation used to explain the loop must be specified independently.

If, contrary to this, social and institutional states routinely reverted smoothly and reversibly to prior configurations as soon as a single parameter (like leadership or policy) was rolled back, $S_3$ would fail as a general intersubjective signature.

## 9.5 Failure modes and falsification criteria

Several kinds of findings would force revision or rejection of the current Conscious Collapse framing:

- **No neural regime change:** If careful perturbation and integration/differentiation measures found **no systematic difference** between conscious and clearly non-conscious states, or if conscious content routinely appeared in regimes of trivial integration, the neural threshold story in Section 2 would be undermined (Casali et al. 2013; Massimini et al. 2005).
- **No S-signals in mind-phase domains:**

If perception, decision, belief change, and intersubjective dynamics showed none of the predicted $S_1/S_2/S_3$ patterns—even when carefully controlled for differences in $K$—the claim that the same audited restriction or signature family transfers across mind, life, and physics would be weakened.

- **FRLB patterns detached from behavior and well-being:** If reliably coded generative and supplanted FRLB patterns add no held-out prediction of task-defined updating or behavior beyond established measures, FRLB would be more metaphor than mechanism.
- Failure of restricted cross-domain transfer:

If each domain requires post hoc redefinition of redundancy, neutrality, sweep, outcome, or time scale—or if independently fit domain models outperform the shared restricted

model—UCT's cross-domain law claim is weakened. Mere ability to map a phenomenon into Ω, K, C, and U does not count as confirmation.

Positive evidence requires more than observing familiar effects. It requires preregistered mappings, successful out-of-domain prediction, and incremental performance relative to GWT/GNWT, predictive-processing, dynamical-systems, clinical, or institutional baselines as appropriate.

The mind-phase proposal is therefore accountable at two levels: local claims can fail within their domains, and the shared-schema claim can fail when restricted models do not transfer. Success at the first level does not guarantee success at the second.

## 9.6 The discriminating prediction: restricted transfer for each signature

Predictions 1–7 and the MIND-1T discriminator are individually compatible with established single-domain accounts. Integration signatures, ambiguity-related latency, multisensory convergence, structured belief change, and institutional path dependence are not uniquely UCT phenomena. Their replication confirms the relevant local effect, not the UCT law claim.

UCT's added prediction is quantitative transfer. For each signature $S_n$, a preregistered restricted model should preserve its declared functional relation and shared or hierarchically modeled parameters across typed domain mappings better than independently flexible models. Family-level support would depend on the combined results of separate $S_1$, $S_2$, and $S_3$ transfer tests. The claim concerns a typed kernel family sharing restrictions—not one operator and not the assertion that unlike phenomena are literally the same object.

For each $S_n$, a decisive design is leave-one-domain-out transfer: fit the restricted model in two domains, predict the third without redefining the signature, and compare predictive performance with domain-native baselines and an independently fit model. Failure to transfer, parameter instability beyond declared tolerances, or improved fit only after post hoc remapping counts against the law-level claim even if every local effect is real.

**Current status.** WP04 specifies this discriminator but does not yet supply the shared quantitative model, invariant parameter set, or completed cross-domain analysis. That is an open empirical obligation rather than a result claimed by this paper.

## §10 Limitations and Open Obligations

**Phenomenal stance and open burdens.** WP04 does not inherit the classical demand to derive a second phenomenal fact from an intrinsically non-experiential process; it adopts a non-bifurcation thesis. That thesis is not established by the generic kernel or by access-marker accuracy. Necessary and sufficient phase criteria, episode unity, content differentiation, first-person/third-person correspondence, and structural completeness remain open. Global availability is not sufficient by definition.

**Phase-identification evidence.** Broadcast, recurrence, integration, and differentiation remain competing or complementary indicators. The sustained-resolution interval lacks a validated measure and must outperform attention, report, evidence-accumulation, and post-resolution-maintenance alternatives.

Kernel fillings. No mind-domain instantiation yet supplies the full declared trace partition, domain trace-quality test, and six-part dependence specification. The open carrier remains uninterpreted. Candidate traces and example records in this paper are therefore provisional placements.

**FRLB measurement.** The functional roles are conceptually separable but not yet a validated scale. Inter-rater reliability, incremental validity, preregistered failure criteria, and comparison with established cognitive and clinical constructs remain owed.

**Cross-domain leverage.** $S_1$–$S_3$ are portable but not individually novel. The law-level claim requires quantitative transfer under fixed mappings and explicit null models; qualitative resemblance is insufficient.

**Normative and clinical boundaries.** The Structural Ethic adds value premises, and the rigidity/pathology discussion supplies research hypotheses rather than diagnoses or treatment guidance.

## §11 Conclusion: Conscious Collapse as a Mind-Phase Proposal

WP04 proposes a disciplined mind-phase mapping: neural and cognitive resolutions are provisionally identified as access-conscious when they enter a globally available, recursively usable regime; perception selects hypotheses under layered constraints; FRLB distinguishes starting trust, search, admission discipline, event records, and standing commitments; action and identity emerge from histories of constrained resolution and update.

The paper's strongest original territory remains the update architecture and the record layer. Generative and supplanted FRLB loops test corrigibility at the update seam while allowing failures to originate in protected constraints, candidate search, record access, admission rules, or U itself; the unit is a commitment-specific loop rather than intelligence or moral worth. CIM explains how cognition becomes durable outside the originating episode and returns as constraint without treating every artifact as CIM or every coupled system as a conscious collective.

The empirical claim is deliberately narrower than the ontology's reach. Local effects must survive domain-native tests, and the UCT contribution must then survive a harder transfer test: fixed S-signature definitions and restricted models should predict across physics, biology, and mind better than flexible independent accounts. That result has not yet been delivered.

WP04 therefore hands WP05 and the empirical program a bounded proposition. Interpretively, conscious occurrences require no second phenomenal output: first-person presence and third-person neural organization are non-interchangeable modes of access to one process. Scientifically, mind may instantiate the proposed constrained-resolution schema, but phase discriminators, certified mind-domain records, episode and content correspondence, and quantitative cross-domain invariants remain open.

## Appendix A: Kernel Interface Used in WP04 (Frozen 2026–07 Architecture)

**Status and scope.** This appendix states the interface used by WP04. It freezes the architecture, not the open domain fillings, and avoids restating the older universal postulates whose record, residue, time, and operator language has been superseded; reconciled successor clauses (LK-1–LK-4) are stated in the series capstone (WP05).

### A.1 Transition and certification

$$(x_t^*, q_t, S_t) \sim C_t^K(\omega_t, \cdot)$$

A draw emits the structured resolution, candidate trace, and open carrier shown above. The candidate trace is not yet a record, and WP04 assigns no entropy or discarded-interpretation meaning to the open carrier.

$$R_t := \mathrm{cert}_P^d(\mathrm{out}(x_t^*), q_t), \qquad \mathrm{Rec}_t := (R_t \neq \mathrm{None}), \qquad E_t := \mathrm{Rec}_t$$

Certification is Option-typed. It requires a declared trace partition, a domain trace-quality test, and a provenance/dependence test with six declarations: intervention, background, metric, threshold, causal model, and partition. WP04 supplies candidate mind-domain placements but not a complete instantiation.

Certification remains cell-wise over a declared partition. WP04 does not choose the mind-domain partition: a runnable instantiation must declare whether the partition cells are local carriers or declared carrier-groups treated as cells and must supply the corresponding recomposition rule.

## A.2 Update, derived state, and event levels

$$K_{t+1} := U(K_t, x_t^*, R_t, S_t), \qquad \Omega_{t+1} := \mathrm{Adm}(\Omega, K_{t+1}), \qquad \omega_{t+1} := \mathrm{post}(x_t^*)$$

$$\mathrm{Upd}_t := \begin{cases} \mathrm{Dep}_r^d(K_{t+1}), & R_t = \mathrm{Some}(r), \\ \mathrm{false}, & R_t = \mathrm{None} \end{cases}$$

$$\mathrm{Closure}_t := \mathrm{Rec}_t \wedge \mathrm{Upd}_t$$

The update map consumes the Option-typed certification result and runs whether certification returns Some(r) or None. Recordhood and uptake are distinct: a certified record survives failed uptake. The displayed uptake diagnostic tests immediate dependence on a supplied record; closure is the conjunction of record occurrence and uptake, while eventhood remains record occurrence.

## A.3 History and domain variation

History depth is derived over a provenance graph. A lineage edge requires a certified source record plus a demonstrated material path into constraints governing a later event; immediate Closure at the source is not required. History is not a sum of records, a transition counter, or an emitted variable.

The cross-domain claim concerns a typed family of domain-specific kernels sharing an architecture and testable restrictions. Physical, biological, neural, cognitive, and social kernels have different domains, codomains, record types, and update maps. One concrete operator does not run unchanged over all of them.

## Appendix B: Provisional Mind-Domain Mapping and Test Obligations

Appendix B lists provisional mind-domain placements and the obligations required to turn them into runnable tests. The examples are not certified simply because they can be named.

### B.1 Provisional mind-domain placements

**State spaces ($\Omega$)**

- $\Omega_{\text{neural}}$ — neural state space: possible patterns of spiking, synaptic strengths, and field dynamics in a given brain.
- $\Omega_{\text{local}}$ — local processing state space: activity/hypothesis spaces of local modules (e.g. sensory cortices, motor areas) under $K_{\text{local}}$.
- $\Omega_{\text{GW}}$ — global workspace state space: patterns of coordinated activity across multiple modules (workspace-like, globally broadcast states).
- $\Omega_{\text{percept}}$ — perceptual hypothesis space: possible interpretations of current sensory input (objects, causes, layouts).
- $\Omega_{\text{mind}}$ — higher-order hypothesis space: possible interpretations, explanations, plans, and self-models (conceptual and narrative-level states).
- $\Omega_{\text{act}}$ — action space: possible behaviors and policies, from immediate motor outputs to long-horizon plans.
- $\Omega_{\text{self}}$ — self-state space: possible configurations of identity, roles, self-narratives, and style.
- $\Omega_{\text{ext}}$ — possible externalization outcomes; CIM is the qualifying record-bearing output category, not the whole state space.

**Constraint sets ($K$)**

- $K_{\text{neural}}$ — neural constraints: anatomy, connectivity, neuromodulators, receptor distributions, metabolic limits.
- $K_{\text{local}}$ — local constraints for sub-modules: local connectivity, tuning, and inputs that shape non-conscious processing.
- $K_{\text{mind}}$ — mind-phase constraint architecture: current beliefs, values, priors, goals, self-models, inference rules, and emotional/attentional weights.
- $K_{\text{percept}}$ — perceptual constraints:

$$K_{\text{percept}} := \text{Compose}(K_{\text{prior}}, K_{\text{context}}, K_{\text{goal}}, K_{\text{neural}})$$

including priors, current context, goals/attention, and sensory architecture.

- $K_{\text{act}}$ — action constraints:

$$K_{\text{act}} := \text{Compose}(K_{\text{phys}}, K_{\text{bio}}, K_{\text{mind}}, K_{\text{soc}})$$

combining physical, biological, mind-level, and social constraints on behavior.

- $K_{\text{soc}}$ — social/institutional constraints: languages, media, norms, laws, infrastructures, economic and technological structures.

**Candidate trace and record placements**

- $q_{\text{neural},t}$ — candidate neural traces such as synaptic change, plasticity traces, and priming; record status requires certification.
- $q_{\text{mind},t}$ — candidate cognitive traces such as memory, narrative, explicit commitment, external notes, or feedback; $R_{\text{mind},t}$ is Option-typed after certification.
- $q_{\text{CIM},t}$ — candidate externalized traces; qualifying CIM records must satisfy the CIM category conditions and the kernel certification requirements.

**Proposed phase/regime placement**

- Conscious Collapse — the proposed mind-domain regime in which selected resolutions become globally available and recursively usable, including the capacity to participate in later control, memory, and constraint update. The phase boundary, experiential differentiation, and structure–report correspondence remain open; global availability is not sufficient by definition.

Conscious Collapse is a phase/regime placement, not a record type. A transition in this regime still emits a realized resolution, candidate trace, and open carrier. Certification may return Some(r) or None, and the mind-domain update runs in either case. If a record certifies, uptake remains a distinct empirical question; conscious-phase status guarantees neither certification nor uptake.

## B.2 Assumptions and scope (WP04)

WP04 operates at the same effective level as contemporary cognitive neuroscience and psychology, with the following assumptions:

- **Single-world program commitment.**

WP04 inherits UCT's single-world ontology. Each modeled transition realizes one outcome, and WP04 does not employ a many-worlds representation.

- **Mind-phase nested in life-phase.**

Human Conscious Collapse, as modeled in WP04, is a phase of Biological Collapse implemented in living nervous systems. The general conscious-phase category is realization-neutral in principle; no non-biological fulfillment is established here.

- Workspace-like regime as a candidate conscious-access architecture.

The present phase-identification hypothesis treats global availability, recurrence, integration, and differentiation as candidate third-person indicators of conscious access. It does not by itself define a sufficient conscious-phase criterion, settle first-person/third-person correspondence, or rule out alternative architectures.

- **Predictive processing and global broadcasting as implementations, not ontology.**

GWT/GNWT and predictive-processing or active-inference frameworks are candidate architecture-level implementations. WP04 is neutral among microcircuit models and requires incremental tests against recurrent and higher-order alternatives.

- Integration, recurrence, differentiation, and broadcast as indicators, not definitions.

Measures such as perturbational complexity are candidate proxies for aspects of the proposed regime. No one measure defines consciousness, and recent adversarial evidence limits strong GNWT- or IIT-specific identification (Cogitate Consortium et al. 2025).

- **Teleonomy, not teleology.**

Conscious systems are treated as teleonomic: their apparent purposefulness arises from structure shaped by evolution and learning, not from extra "goal forces" outside the kernel.

- **Human focus, graded continuum.**

WP04 focuses primarily on human consciousness as the best-characterized mind-phase, while acknowledging that non-human animals may instantiate similar regimes to varying degrees.

## B.3 Test classes and the law-level discriminator

Section 9 groups the local tests below and separates them from the law-level discriminator. No local effect is decisive for UCT by itself:

- **Neural thresholds and temporal persistence (MIND-1/MIND-1T; §9.1).**

Conscious vs non-conscious states should differ systematically in integration/differentiation patterns (e.g., PCI-like measures) and in global broadcasting behavior. Transitions between unconscious and conscious regimes should correspond to qualitative shifts in these patterns, not just to superficial changes in behavior.

- **S-signals in perception and decision (§9.2).**

Perception and decision-making modeled as typed resolutions under declared constraints should show preregistered S-signatures and beat domain-native null models:

  - $S_1$ (independent redundancy → consensus) in high-redundancy, well-constrained conditions after a declared independence audit;
  - $S_2$ (neutrality → delayed resolution) in ambiguous or conflicted conditions;

    - analogous signatures in reaction times, uncertainty measures, and bistable experiences.

- **FRLB patterns in belief and identity change (§9.3).**

If FRLB captures real structure in belief updating:

    - reliably coded generative FRLB episodes should predict task-bounded update trajectories beyond established measures,
    - supplanted episodes should show selective update insensitivity under preregistered evidence-direction tests,
    - no disorder-specific signature is assumed; any relation to symptoms or behavior must survive clinical, developmental, biological, and contextual controls.

- **CIM and intersubjective S-signals (§9.4).**

CIM-mediated systems may be tested for operational analogues of:

    - $S_1$: convergence and consensus in mature, high-evidence, accuracy-driven domains;
    - $S_2$: persistent controversies and neutral zones where evidence is underdetermining or incentives are misaligned;
    - $S_3$: audited hysteresis-like path dependence in institutional and normative shifts under declared control sweeps.

- Restricted cross-domain transfer (§§9.5–9.6).

Generic mappability into Ω, K, C, and U is not a decisive test. The law-level claim requires preregistered mappings, shared restrictions, leave-one-domain-out prediction, and performance against independently fit and domain-native alternatives.

Failure of local predictions would weaken the corresponding mind-phase components. Failure of restricted cross-domain transfer would weaken the law-level claim even if local neural, behavioral, and institutional effects remained real.

## Appendix C: On the Status of Consciousness in UCT

This appendix clarifies how the proposed Conscious Collapse regime fits within UCT and what is—and is not—being claimed about its relation to physics, biology, and value. It is interpretive context for the mind-phase proposition, not a completed formal law argument.

### C.1 Position of Conscious Collapse in the collapse stack

UCT proposes a typed family of domain kernels sharing the interface stated in Appendix A:

$$(x_t^*, q_t, S_t) \sim C_t^K(\omega_t, \cdot),$$
$$R_t := \mathrm{cert}_P^d(\mathrm{out}(x_t^*), q_t), \qquad K_{t+1} := U(K_t, x_t^*, R_t, S_t)$$

A transition emits a resolution, candidate trace, and open carrier; certification establishes whether a record exists; the update runs on Some(r) or None. Possibility regions and next state are derived, and history depth is reconstructed from certified record-mediated dependencies.

Within this stack:

- **First Collapse / physics (WP02):**

WP02 proposes that collapse under physical constraints $K_{\text{phys}}$ realizes matter, fields, and phases of matter as stabilized constraint architectures; the status of apparent laws and any law-level cross-domain claim remains a burden for WP05 and the wider empirical program.

- **First Biological Collapse / life-phase (WP03):**

A provisional biological-domain kernel under composed physical and chemical constraints models self-maintaining, self-propagating networks that regenerate parts of their own constraint architecture.

- **Conscious Collapse / mind-phase (WP04):**

A provisional mind-domain kernel under neural and cognitive constraints models a regime in which some resolutions become globally available and recursively usable and can influence later $K_{\text{mind}}$.

On the present proposal, conscious access is a late-arriving organization within the life-phase. This nesting is a structural implementation hypothesis; it does not by itself determine necessary or sufficient phase criteria, episode boundaries, or content correspondence.

## C.2 Life-phase precedence and value

Nothing in the descriptive kernel establishes where value originates. The following claims are philosophical interpretations continuous with the series, not formal consequences of C and U:

- **Biological Collapse** already constitutes a four-billion-year chain of coherence: cells, bodies, nervous systems, social groups, ecosystems, and proto-intent all maintaining themselves under gradients.
- **Conscious Collapse** names the regime in which a branch of that chain becomes able to *model itself*: to treat its own states, histories, and futures as content.

In this frame:

- Life-phase coherence is **structurally prior** to mind.
- Conscious Collapse is a **later, nested regime** that can recognize and act on that pre-existing coherence.

Put simply:

On this interpretation, consciousness does not retroactively grant value to life; it inherits a biologically prior organization and may recognize and act on it from within.

Conscious beings are one kind of node in a longer biological lineage. Any claim about the lineage's importance depends on declared evaluative premises.

## C.3 Metaphysical scope and implementation neutrality

UCT's neutrality is implementation neutrality, not neutrality about every ontology of mind.

UCT remains neutral about whether the relevant microphysical implementation is quantum, classical, or hybrid; about the exact microcircuit realization; and about whether later philosophy labels the non-bifurcation view process-monist, dual-aspect, neutral-monist, or something else.

It is not neutral about a primitive physical/phenomenal bifurcation that treats an otherwise complete neural process as producing an added mental substance or output. It also rejects the deletion of experience as unreal and an epiphenomenal duplicate disconnected from the occurrence's causal organization.

Instead, UCT proposes that, for a genuinely conscious occurrence, first-person presence and third-person neural organization are non-interchangeable modes of access to one process. Neither evidence stream is exhaustive, and the modes are not substitutes for one another.

This is a program-level interpretive commitment, not a theorem of C or U and not evidence that the proposed phase markers are correct. Phase identification, episode structure, phenomenological differentiation, and first-person/third-person correspondence remain active obligations.

### C.4 Neural realization rather than production of a second phenomenal object

In ordinary speech, 'the brain generates consciousness' is acceptable only as shorthand for neural constraint architectures realizing human conscious-phase organization. It must not imply that an otherwise complete neural process emits a second phenomenal object.

The brain is a physical constraint architecture $K_{neural}$: a network with particular anatomy, connectivity, dynamics, and plasticity rules. WP04 proposes that some neural and cognitive resolutions enter a globally available, recursively usable regime and treats broadcast, recurrence, integration, and differentiation as candidate third-person indicators of that organization.

Which organizations are necessary or sufficient remains empirical, architecture-sensitive, and open. Realization language marks implementation and dependence, not a completed account of episode unity or content differentiation.

This framing preserves the dependence of human consciousness on living neural organization while leaving open whether other architectures could realize related conscious-phase occurrences.

### C.5 Consciousness as lens, not throne

Because experience is epistemically primary (all other domains are known *through* it), there is a strong temptation to treat it as ontologically or morally primary. UCT resists this slippage.

In UCT:

- **Epistemically,** conscious experience is the medium through which we know physics and biology.
- Ontologically, UCT proposes physics and biology as structurally prior regimes in the series architecture.
- **Structurally,** consciousness is the regime where some of that coherence becomes *self-referential*: able to represent itself as content and act on those representations.

A concise way to say this is:

A bounded version of the claim is: conscious access is a regime in which selected representations of reality's structure and significance become available to a system that can act on them. 'Significance' here belongs to the paper's declared philosophical extension, not to the kernel alone.

This positioning allows WP04 to support later ethical and cultural work without importing extra metaphysical assumptions into the law stack.

The non-bifurcation thesis is the paper's interpretive stance; conscious-phase discrimination, episode and content correspondence, certified mind-domain record tests, and law-level invariants remain unsettled.

## References


Baars, B. J. (1988). *A Cognitive Theory of Consciousness*. Cambridge University Press.

Bassi, A., Lochan, K., Satin, S., Singh, T. P., & Ulbricht, H. (2013). Models of wave-function collapse, underlying theories, and experimental tests. *Reviews of Modern Physics*, 85(2), 471–527.

Beck, A. T. (1976). *Cognitive Therapy and the Emotional Disorders*. International Universities Press.

Beck, A. T., & Clark, D. A. (1997). An information processing model of anxiety: Automatic and strategic processes. *Behaviour Research and Therapy*, 35(1), 49–58.

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive Therapy of Depression*. Guilford Press.

Birch, J., Schnell, A. K., & Clayton, N. S. (2020). Dimensions of animal consciousness. *Trends in Cognitive Sciences*, 24(10), 789–801.

Block, N. (1995). On a confusion about a function of consciousness. Behavioral and Brain Sciences, 18(2), 227–287.

Bogacz, R. (2007). Optimal decision-making theories: Linking neurobiology with behaviour. *Trends in Cognitive Sciences*, 11(3), 118–125.

Burns, D. D. (1999). *Feeling Good: The New Mood Therapy*. Avon.

Casali, A. G., Gosseries, O., Rosanova, M., et al. (2013). A theoretically based index of consciousness independent of sensory processing and behavior. *Science Translational Medicine*, 5(198), 198ra105.

Chalmers, D. J. (1995). Facing up to the problem of consciousness. Journal of Consciousness Studies, 2(3), 200–219.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(3), 181–204.

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, 58(1), 7–19.

Clark, D. A., & Beck, A. T. (2010). *Cognitive Therapy of Anxiety Disorders: Science and Practice*. Guilford Press.

Cogitate Consortium, Ferrante, O., Gorska-Klimowska, U., et al. (2025). Adversarial testing of global neuronal workspace and integrated information theories of consciousness. Nature, 642, 133–142. https://doi.org/10.1038/s41586-025-08888-1

Conway, M. A., & Pleydell-Pearce, C. W. (2000). The construction of autobiographical memories in the self-memory system. *Psychological Review*, 107(2), 261–288.

Dayan, P., & Abbott, L. F. (2001). *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. MIT Press.

Dehaene, S. (2014). *Consciousness and the Brain: Deciphering How the Brain Codes Our Thoughts*. Viking.

Dehaene, S., & Changeux, J.-P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron*, 70(2), 200–227.

Dennett, D. C. (2003). *Freedom Evolves*. Viking.

Devia, C., Concha-Miranda, M., & Rodríguez, E. (2022). Bi-stable perception: Self-coordinating brain regions to make-up the mind. *Frontiers in Neuroscience*, 15, 805690.

Dretske, F. (1988). *Explaining Behavior: Reasons in a World of Causes*. MIT Press.

Edelman, D. B., & Seth, A. K. (2009). Animal consciousness: A synthetic approach. *Trends in Neurosciences*, 32(9), 476–484.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415(6870), 429–433.

Festinger, L. (1957). *A Theory of Cognitive Dissonance*. Stanford University Press.

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

Gallagher, S. (2000). Philosophical conceptions of the self: Implications for cognitive science. *Trends in Cognitive Sciences*, 4(1), 14–21.

Ghosh, V. E., & Gilboa, A. (2014). What is a memory schema? A historical perspective on current neuroscience literature. *Neuropsychologia*, 53, 104–114.

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience*, 30, 535–574.

Goldstein, E. B. (2014). *Sensation and Perception* (9th ed.). Cengage Learning.

Griffin, D. R., & Speck, G. B. (2004). New evidence of animal consciousness. *Animal Cognition*, 7(1), 5–18.

Habermas, J. (1984). *The Theory of Communicative Action, Volume 1: Reason and the Rationalization of Society*. Beacon Press.

Haggard, P. (2008). Human volition: Towards a neuroscience of will. *Nature Reviews Neuroscience*, 9(12), 934–946.

Hahn, U., & Oaksford, M. (2007). The rationality of informal argumentation: A Bayesian approach to reasoning fallacies. *Psychological Review*, 114(3), 704–732.

Hameroff, S., & Penrose, R. (1996). Orchestrated reduction of quantum coherence in brain microtubules: A model for consciousness. *Mathematics and Computers in Simulation*, 40(3–4), 453–480.

Hameroff, S., & Penrose, R. (2014). Consciousness in the universe: A review of the 'Orch OR' theory. *Physics of Life Reviews*, 11(1), 39–78.

Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (2011). *Acceptance and Commitment Therapy: The Process and Practice of Mindful Change* (2nd ed.). Guilford Press.

Hebb, D. O. (1949). *The Organization of Behavior: A Neuropsychological Theory*. Wiley.

Henrich, J. (2015). *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter*. Princeton University Press.

Hohwy, J. (2013). *The Predictive Mind*. Oxford University Press.

Hutchins, E. (1995). *Cognition in the Wild*. MIT Press.

Ingram, R. E., Miranda, J., & Segal, Z. V. (1998). *Cognitive Vulnerability to Depression*. Guilford Press.

James, W. (1950). *The Principles of Psychology* (Original work published 1890). Dover.

Jones, J. C. (2025a). Universal Collapse Theory—Collapse in Physics: Coherence as Law from Cosmology to Matter (WP02 v1.0). HoldingLight LLC. https://doi.org/10.17605/OSF.IO/3NYMP

Jones, J. C. (2025b). Universal Collapse Theory—Foundations of Collapse (WP01 v2.0). HoldingLight LLC. https://doi.org/10.17605/OSF.IO/VZ836

Jones, J. C. (2026a). Auditing Constraint Asymmetry in Latency-Based Resolution Tests ($S_2$), v1.0. HoldingLight LLC. https://doi.org/10.17605/OSF.IO/HRKWT

Jones, J. C. (2026b). Auditing Independence in Multi-Channel Measurement ($S_1$), v1.0. HoldingLight LLC. https://doi.org/10.17605/OSF.IO/7U8SK

Jones, J. C. (2026c). Auditing Record State in Constraint-Sweep Hysteresis Tests ($S_3$), v1.0. HoldingLight LLC. https://doi.org/10.17605/OSF.IO/CQGTD

Jones, J. C. (2026d). Biological Faith Systems: How Living Systems Commit Before Certainty (v1.0). HoldingLight LLC. https://doi.org/10.17605/OSF.IO/D37Q5

Jones, J. C. (2026e). Consciousness-Induced Material: A Structural Ontology of Externalized Cognition (v1.0). HoldingLight LLC. https://doi.org/10.17605/OSF.IO/7URMK

Jones, J. C. (2026f). Constraint-Dependent Perceptual Resolution: A Pre-Specified Reanalysis of the COGITATE iEEG Dataset (v1.0). HoldingLight LLC. https://doi.org/10.17605/OSF.IO/MXYU2

Jones, J. C. (2026g). How Minds Resolve: The Faith–Reason–Logic–Belief Cycle (v1.0). HoldingLight LLC. https://doi.org/10.17605/OSF.IO/92RQ5

Jones, J. C. (2026h). Human Interface Laws: A Philosophical Mechanics of Belief, Update, and Shared Reality (v1.0). HoldingLight LLC. https://doi.org/10.17605/OSF.IO/437HX

Jones, J. C. (2026i). Records Across Nature, Life, and Mind: The Persistence Layer of Collapse (v2.0). HoldingLight LLC. https://doi.org/10.17605/OSF.IO/7H6DY

Jones, J. C. (2026j). The Structuralization of Empiricism: Formalizing the Structural Conditions Under Which Empiricism Stabilizes Knowledge (v1.0). HoldingLight LLC. https://doi.org/10.17605/OSF.IO/J4GZ9

Jones, J. C. (2026k). The Tether: A Foundational Discipline for Universal Collapse Theory (v1.0). HoldingLight LLC. https://doi.org/10.17605/OSF.IO/CBR4J

Jones, J. C. (2026l). Universal Collapse Theory—Biological Collapse: Life as a Phase of Constraint-Guided Collapse (WP03 v1.0). HoldingLight LLC. https://doi.org/10.17605/OSF.IO/5SZ86

Jones, J. C. (2026m). Universal Collapse Theory—Mind Starter Pack (v1.0) [Data and software package]. HoldingLight LLC. https://doi.org/10.17605/OSF.IO/A25CV

Jones, J. C. (2026n). Update Integrity Standard (UIS v1.0): A Structural Ethic for Preserving Corrigibility in Record-Based Systems. HoldingLight LLC. https://doi.org/10.17605/OSF.IO/DWM29

Kabat-Zinn, J. (1990). *Full Catastrophe Living*. Dell.

Kandel, E. R. (2001). The molecular biology of memory storage: A dialogue between genes and synapses. *Science*, 294(5544), 1030–1038.

Kandel, E. R., Schwartz, J. H., Jessell, T. M., Siegelbaum, S. A., & Hudspeth, A. J. (2013). *Principles of Neural Science* (5th ed.). McGraw-Hill.

Knill, D. C., & Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neurosciences*, 27(12), 712–719.

Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*. University of Chicago Press.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108(3), 480–498.

Lakatos, I. (1978). *The Methodology of Scientific Research Programmes*. Cambridge University Press.

Leopold, D. A., & Logothetis, N. K. (1999). Multistable phenomena: Changing views in perception. *Trends in Cognitive Sciences*, 3(7), 254–264.

Mahoney, J. (2000). Path dependence in historical sociology. *Theory and Society*, 29(4), 507–548.

Mashour, G. A., Roelfsema, P., Changeux, J.-P., & Dehaene, S. (2020). Conscious processing and the global neuronal workspace hypothesis. *Neuron*, 105(5), 776–798.

Massimini, M., Ferrarelli, F., Huber, R., Esser, S. K., Singh, H., & Tononi, G. (2005). Breakdown of cortical effective connectivity during sleep. *Science*, 309(5744), 2228–2232.

McAdams, D. P. (2001). The psychology of life stories. *Review of General Psychology*, 5(2), 100–122.

Merton, R. K. (1973). *The Sociology of Science: Theoretical and Empirical Investigations*. University of Chicago Press.

Millikan, R. G. (1984). *Language, Thought, and Other Biological Categories*. MIT Press.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2(2), 175–220.

North, D. C. (1990). *Institutions, Institutional Change and Economic Performance*. Cambridge University Press.

Oaksford, M., & Chater, N. (2007). *Bayesian Rationality: The Probabilistic Approach to Human Reasoning*. Oxford University Press.

Ostrom, E. (1990). *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge University Press.

Ouellette, J. A., & Wood, W. (1998). Habit and intention in everyday life: The multiple processes by which past behavior predicts future behavior. *Psychological Bulletin*, 124(1), 54–74.

Pariser, E. (2011). *The Filter Bubble: What the Internet Is Hiding from You*. Penguin.

Pierson, P. (2000). Increasing returns, path dependence, and the study of politics. *American Political Science Review*, 94(2), 251–267.

Pothos, E. M., & Busemeyer, J. R. (2013). Can quantum probability provide a new direction for cognitive modeling? Behavioral and Brain Sciences, 36(3), 255–274.

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2(1), 79–87.

Redgrave, P., Prescott, T. J., & Gurney, K. (1999). The basal ganglia: A vertebrate solution to the selection problem? *Neuroscience*, 89(4), 1009–1023.

Rumelhart, D. E., & Norman, D. A. (1978). Accretion, tuning, and restructuring: Three modes of learning. In J. W. Cotton & R. L. Klatzky (Eds.), *Semantic Factors in Cognition* (pp. 37–53). Lawrence Erlbaum.

Scheffer, M. (2009). *Critical Transitions in Nature and Society*. Princeton University Press.

Squire, L. R., & Kandel, E. R. (2009). *Memory: From Mind to Molecules* (2nd ed.). Oxford University Press.

Sterzer, P., Kleinschmidt, A., & Rees, G. (2009). The neural bases of multistable perception. *Trends in Cognitive Sciences*, 13(7), 310–318.

Sunstein, C. R. (2017). *#Republic: Divided Democracy in the Age of Social Media*. Princeton University Press.

Synofzik, M., Vosgerau, G., & Newen, A. (2008). Beyond the comparator model: A multifactorial two-step account of agency. *Consciousness and Cognition*, 17(1), 219–239.

Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Harvard University Press.

Tononi, G. (2004). An information integration theory of consciousness. *BMC Neuroscience*, 5, 42.

Tononi, G. (2008). Consciousness as integrated information: A provisional manifesto. *Biological Bulletin*, 215(3), 216–242.

Tononi, G., Boly, M., Massimini, M., & Koch, C. (2016). Integrated information theory: From consciousness to its physical substrate. *Nature Reviews Neuroscience*, 17(7), 450–461.

West-Eberhard, M. J. (2003). *Developmental Plasticity and Evolution*. Oxford University Press.

Wood, W., & Neal, D. T. (2007). A new look at habits and the habit–goal interface. *Psychological Review*, 114(4), 843–863.

Yearsley, J. M., & Busemeyer, J. R. (2016). Quantum cognition and decision theories: A tutorial. Journal of Mathematical Psychology, 74, 99–116.

This paper is part of the Universal Collapse Theory library. For a reading guide and full architecture, visit universalcollapse.com/roadmap.

AI Disclosure: AI tools (Anthropic's Claude and OpenAI's GPT) were used to assist with manuscript preparation, drafting, organization, and editorial refinement. The underlying theory, structural decisions, analysis, and conclusions are the author's own.

Citation: Jones, J. C. (2026). Universal Collapse Theory—Conscious Collapse: Mind as a Phase of Constraint-Guided Collapse (WP04 v1.0). HoldingLight LLC. https://doi.org/10.17605/OSF.IO/T5H6W