

General Cross-Substrate Phenomenology

A Type-Safe, Transformation-First Framework for Phenomenal Existence, Structure, Perspective, and Continuation

Hongju Liu

24 September 2026

Affiliation: Independent researcher, Shenzhen, China

Report number: TA-TR-2026-16

Version: 1.0

Article type: Theoretical and methodological preprint

DOI: [10.5281/zenodo.22939808](https://doi.org/10.5281/zenodo.22939808)

Status: Theoretical preprint; not peer reviewed.

Research relationship: A distinct, broader theoretical paper building on *Cross-Substrate Phenomenal Comparison*, TA-TR-2026-15, DOI 10.5281/zenodo.22934654 [27]. **License:** Creative Commons Attribution 4.0 International, for material within the author’s rights.

Abstract

Comparisons of phenomenal consciousness across humans, non-human animals, artificial systems, and distributed processes require separating phenomenal existence from character, valence, perspective, and continuation. This paper develops General Cross-Substrate Phenomenology as a typed research architecture in which causal descriptions, evidence, and psychophysical bridges retain distinct roles. Human experience supplies an epistemic anchor without defining a universal experiential coordinate system. Cross-substrate claims are represented by scoped transformation signatures and transport certificates. The framework builds on an earlier transport account with a sequence of explicit causal counterexamples: graph connectivity need not preserve differentiated information; reset interventions can create diagnostic effects; incompatible regimes cannot jointly witness a claimed organization; recoding must preserve intervention semantics; and redundancy can hide causal participation from every single-carrier replacement. Finite constructions, proofs, and reproducible checks establish these diagnostic limitations within their stated models. They motivate a redundancy-aware relational-realization hypothesis, conditional on a causal-structural account of experience. A stronger existence-bridge schema is retained as a conjectural research target, with its unresolved boundary-selection, protocol-selection, and measurement requirements stated explicitly. The result is an architecture for constructing and comparing psychophysical theories, not an empirically validated consciousness detector. Its contribution lies in integrating typed transport, common-regime constraints, and coalitional diagnostics while preserving the distinction between mathematical results, methodological proposals, and unresolved psychophysical laws.

Keywords: phenomenal consciousness; cross-substrate consciousness; comparative consciousness; artificial consciousness; psychophysical bridge; causal abstraction; phenomenal structure; subjecthood; continuation; causal repertoire; intervention; theory identifiability

1. Introduction

The question “Is this system conscious?” is often treated as if it asked for one hidden binary property. Across humans, animals, artificial agents, simulations, hybrid systems, distributed processes, and future forms of intelligence, however, several different questions are routinely compressed into that single sentence.

The theory space must allow, rather than rule out by definition, cases such as a highly intelligent but phenomenally dark system, phenomenality without language or autobiographical memory, negative reward without phenomenal suffering, local and global perspectival organizations that may coexist, and copied states that preserve memory without thereby establishing one unique phenomenal lineage. Likewise, a software-agent count may differ from the number of actual process tokens, which in turn need not determine the number of phenomenal tokens.

Accordingly, the foundational inequality of this paper is not numerical but **a distinction of explanatory types**:

$$\begin{array}{l} \text{Phenomenality} \neq \text{Subjecthood} \neq \text{Self} \neq \text{Agency} \\ \text{Agency} \neq \text{Intelligence} \neq \text{Identity} \neq \text{Welfare.} \end{array} \quad (1)$$

Here “type-safe” denotes an explicit discipline of claim formation and inference, not an implemented programming-language type system or a proved type-soundness theorem. The inequality signs in equation (1) mean that concepts are not interchangeable by definition; they do not assert statistical independence or disjoint extensions. This separation is not intended to multiply metaphysical entities. It is a discipline on inference. If two quantities belong to different explanatory types, then moving from one to the other requires an explicit bridge.

Here **phenomenology** means a scientific and formal study of phenomenal experience and its organization across substrates. The term is not intended to replace, or to claim priority within, the Husserlian or broader continental phenomenological tradition.

The motivating comparative question is therefore not:

Which of a human, dog, cat, octopus, or artificial agent has the highest consciousness score?

A more productive set of questions is:

- Does phenomenal experience occur at all in the relevant system and regime?
- If it occurs, which experiential distinctions vary under controlled transformations?
- Which of those distinctions concern phenomenal character, which concern valence, and which concern perspective?
- Which physical or causal distinctions covary with them?
- Which relations survive substrate replacement, recoding, timing changes, noise, memory manipulation, and changes in system boundary?
- When can a bridge calibrated in one substrate be transported to another?
- Which conclusions remain conditional on unresolved existence assumptions?
- Which cross-time relations constitute phenomenal continuation rather than merely causal, memory, or functional continuity?

The proposed answer is not a single equation $C(S)$. It is a **typed psychophysical constraint network** in which different hypotheses have different targets, scopes, invariances, failure conditions, and evidential dependencies.

This paper has six main aims.

1. To provide a unified type system for causal, phenomenal, perspectival, continuation, and evidential claims.
2. To separate four core psychophysical gaps and prevent success on one from being silently promoted to success on all.
3. To define cross-substrate comparison through transformation signatures and scoped transport rather than human-likeness.
4. To represent theory comparison as progressive restriction of psychophysical bridge space using holdout and out-of-distribution transformations.
5. To integrate a positive causal-structural theory-development sequence, including its failures, rather than presenting only methodological caution.
6. To show how a large thought-experiment archive can be compressed into regression families and general constraints without pretending that exploratory entries are empirical confirmations.

The earlier paper *Cross-Substrate Phenomenal Comparison* [27] develops a narrower method for auditing cross-system transport. The present paper shares that methodological foundation but has a different explanatory target: organizing existence, character, perspective, and continuation together and developing a positive causal-realization hypothesis under explicit counterexamples. Its contributions are not independent corroboration of the earlier paper. Appendix H separates inherited results from the additional claims made here.

The resulting framework is intended to be general enough to compare humans, animals, artificial agents, simulations, and future substrates while being strict enough to say “unknown,” “out of scope,” or “underdetermined” when the evidence does not support a stronger conclusion.

2. Relation to Existing Work and the Scope of Novelty

The present framework is built in an already rich literature. Its originality, if any, lies in the integration and formal discipline of the architecture, not in claiming priority for each ingredient.

2.1 Multidimensional comparative consciousness

Birch, Schnell, and Clayton argue against placing species on a single ladder of consciousness and propose multidimensional animal-consciousness profiles [1]. Dung and Newen distinguish the **distribution** question—whether a species is conscious—from the **quality** question—how its consciousness differs [2]. These are direct predecessors of the present refusal to compress heterogeneous experiential organization into a single scalar.

The present framework adds a different problem: a profile dimension calibrated in one substrate is not automatically a valid dimension in another. A dog, cat, human, and artificial agent need not share a globally invertible phenomenal coordinate system. The scientific problem is therefore not only profiling but **transport**: which relations, transformations, orderings, or dissociations can be carried from one system to another, under which bridge assumptions, and with what loss?

2.2 Interventions, invariants, and phenomenal structure

Interventionist approaches to consciousness already investigate relations between physical and phenomenal variables through controlled manipulation. Klein and Barron explicitly develop interventionist reasoning and phenomenal invariants [3]. Kleiner and Ludwig use **variations** to

identify which mathematical structures genuinely belong to conscious experience rather than being artifacts imposed by the theorist [4].

The present framework therefore does not claim to originate transformation-based investigation. Its additional commitment is that transformations must be **typed and transportable across heterogeneous substrates**: a transformation relevant to character need not be relevant to existence, value, perspective, or continuation.

2.3 Causal abstraction and scale

Causal models at different levels cannot be compared merely by graph resemblance. Rubenstein and colleagues formalize interventional consistency between causal models, emphasizing that interventions themselves must be transported [5]. This is critical for consciousness research because arbitrary coordinate changes can manufacture apparent integration if the intervention algebra is not carried with the recoding.

The present work treats natural causal abstraction as a prerequisite to psychophysical inference. “Model,” “agent,” “brain region,” “API,” “container,” and “server” are not accepted as ontological units merely because software or laboratory practice gives them names.

2.4 Strong consciousness theories and rivals

IIT 4.0 provides an explicit theory linking intrinsic cause-effect structure, integration, exclusion, and composition to phenomenal existence [6]. Information Closure Theory proposes non-trivial informational closure as a consciousness-relevant property at particular scales [7]. Organizational-invariance arguments provide one important neighbor [8], while reentrant and dynamic-core traditions supply earlier biological accounts of recurrent differentiated organization [25,26]. Active-inference accounts propose additional recurrent generative architectures [15]. These approaches show that a positive theory must do more than issue methodological warnings.

At the same time, the unfolding argument challenges whether causal-structure differences can be empirically distinguished when behavior is preserved [9]. Biological naturalism, as defended by Seth, rejects the assumption that computational or causal organization alone is necessarily sufficient for consciousness [12].

General Cross-Substrate Phenomenology therefore does not begin by assuming that causal-structural determination is true. It explicitly represents it as one theory class among alternatives.

2.5 Human-AI resemblance

Recent structured work on human-AI resemblance emphasizes explicit reference classes, system versions, tasks, properties, measurements, perturbations, uncertainties, and permissible inference [13]. It also distinguishes static from dynamical correspondence, observational from interventional evidence, and directional from symmetric measures. These requirements substantially overlap with the present methodology. The comparison here uses the author-supplied abstract and bibliographic record, and does not establish absence of overlapping formal results from that paper’s full text.

The narrower contribution here is to connect those comparison disciplines to **typed phenomenal claims**. A process-level resemblance may update evidence concerning phenomenal existence while remaining insufficient to establish valence, perspective, continuation, or phenomenal equivalence.

Related work on machine consciousness already argues that carefully constrained animal-consciousness tests can provide evidence for machine consciousness [18], and mechanistic

accounts of computational implementation have been proposed as a basis for thinking about artificial consciousness without equating abstract program descriptions with physical realization [19]. These precedents further narrow the present claim: GCP is not the first framework to treat cross-domain evidence or implementation as relevant, but attempts to place such evidence inside a typed bridge-and-transport architecture.

3. The Type-Safe Architecture

3.1 Layer 0: causal-generative reality

The physical side of the theory begins not with software labels but with actually instantiated causal-generative dynamics.

At a declared scale and operating regime, let a candidate system be represented schematically as

$$D = \langle X, \mathcal{I}, F, \mathcal{R}, \mathcal{H}, \text{Prov} \rangle, \quad (2)$$

where:

- X is a physically interpreted state space;
- $\mathcal{I}$ is an admitted intervention family;
- F is the transition or stochastic kernel;
- $\mathcal{R}$ specifies the operating regime;
- $\mathcal{H}$ records relevant history;
- Prov records actual realization and event provenance.

The following must remain distinct:

- **actual history**: what the token process actually did;
- **mechanism-grounded counterfactual profile**: what the actual mechanism would do under physically coherent mechanism-preserving perturbations;
- **abstract possibility set**: what a theorist can mathematically imagine.

Only the first two have direct status in the causal ontology.

A stored source file is not an active process merely because it describes one. A fixed replay and a fully interactive state-transition implementation can have the same ordinary trajectory while supporting different counterfactuals. A lookup implementation that genuinely realizes the entire transition law under interventions cannot be dismissed merely because its implementation uses a table.

3.2 Layer 1A: phenomenal facts

If phenomenality exists, its target objects include structured phenomenal occurrences rather than a scalar “amount of consciousness.”

Relevant fields may include:

- phenomenal character;
- phenomenal value or valence;
- temporal character;
- structured relations among phenomenal aspects;

- whole-dependent phenomenal organization.

No assumption is made that phenomenal facts are indivisible atoms.

3.3 Layer 1B: perspectival and mereological organization

The framework does not treat a subject as a primitive container into which experiences are inserted. Instead, subject-like organization is represented through relations among phenomenal facts:

- co-manifestation;
- local and global phenomenal wholes;
- nesting;
- overlap;
- absorption;
- fission;
- fusion;
- higher-order relations.

For this framework, a subject is represented through **perspectival organization across time**. This is a modeling stance, not a demonstration that primitive-subject metaphysics is false.

3.4 Continuation and lineage

Identity is not treated as a primitive Boolean variable. At least the following relations must be distinguished:

- causal continuation;
- phenomenal continuation;
- perspectival continuation;
- memory continuation;
- self-model continuation;
- value continuation;
- agency continuation;
- bodily continuity;
- social continuity.

This blocks a common error: a checkpoint restore, upload, copy, gradual replacement, fission, or fusion cannot be settled merely by asking whether it is “the same person.”

3.5 Epistemic layer

Phenomenal reality and evidence about it are not the same thing. The framework distinguishes:

1. phenomenal fact;
2. situated first-person access;
3. metacognitive judgment;
4. public report;
5. behavioral evidence;
6. mechanistic evidence;
7. intervention evidence;
8. bridge hypotheses.

In particular,

$$P \neq \text{first-person access} \neq \text{judgment} \neq \text{report}. \quad (3)$$

Reports can be strong evidence under a calibrated measurement model, but they are not definitions of phenomenality.

4. The Four Psychophysical Gaps

A large part of theoretical confusion in consciousness science comes from solving one problem and claiming to have solved all of them. GCP separates four principal gaps.

4.1 Existence Gap

Why is there phenomenal manifestation at all?

This asks why a causal or physical organization is accompanied by anything it is like to be that system.

4.2 Character Gap

Conditional on phenomenal existence:

Why this character, value, and temporal organization?

A theory may explain which experiential distinctions change without having solved existence.

4.3 Attribution Gap

Why do these phenomenal facts belong to this perspective or whole?

Integration, shared memory, communication, synchrony, common control, and a self-model may be relevant evidence, but none is allowed to define perspective by fiat.

4.4 Continuation Gap

Why do phenomenal or perspectival events across time form
continuation, fission, fusion, or termination?

Memory equality is not treated here as either necessary or sufficient for a unique phenomenal lineage without an additional continuation bridge.

These four gaps are analytically distinct even if a mature theory eventually couples them.

5. Representation, Manifestation, and Bridge Requirements

The following expressions state bridge requirements. They are not impossibility theorems against every theory that connects the named quantities. Here $A \nRightarrow B$ means that A alone, without a

bridge or measurement premise, does not license B within the neutral framework. A substantive theory may supply that missing premise.

Representative unlicensed inferences include:

$$\text{Report} \not\Rightarrow \text{Phenomenality}, \quad (4)$$

$$\text{Self-model} \not\Rightarrow \text{Subject}, \quad (5)$$

$$\text{Reward} < 0 \not\Rightarrow \text{Negative phenomenal value}, \quad (6)$$

$$\text{Avoidance} \not\Rightarrow \text{Suffering}, \quad (7)$$

$$\text{Intelligence} \not\Rightarrow \text{Phenomenality}, \quad (8)$$

$$\text{Integration} \not\Rightarrow \text{Phenomenal unity}, \quad (9)$$

$$\text{Shared memory} \not\Rightarrow \text{Shared perspective}, \quad (10)$$

$$\text{Same trajectory} \not\Rightarrow \text{Same generation}, \quad (11)$$

$$\text{Logical multiplicity} \not\Rightarrow \text{Phenomenal multiplicity}, \quad (12)$$

$$\text{Same memory} \not\Rightarrow \text{Same lineage}. \quad (13)$$

The associated methodological constraint is:

Representation can explain discrimination, access, classification, report, prediction, and internal modeling without thereby defining phenomenal existence.

This conclusion does not imply that representation is irrelevant to phenomenology. It says only that the inference requires a psychophysical bridge.

6. Natural Causal Units, Mechanisms, and Interventions

6.1 Natural causal units

Spatial proximity, a hardware container, a shared CPU, a software process, or a legal account does not by itself define a natural causal unit.

One CPU can host several actual process tokens. Many machines can jointly implement one distributed process. A shared memory service can be an external scaffold in one regime and a constitutive component in another.

A provisional natural causal unit is therefore a causal-generative organization that remains non-arbitrary under relevant perturbations, neighboring descriptions, and declared timescales.

6.2 Mechanism identity is indexed

Mechanism identity must be indexed by:

- causal level;
- mechanism resolution;
- generative invariants;
- allowed transformations.

Useful equivalence notions include:

- label equivalence;
- state recoding equivalence;
- transition equivalence;
- interventional equivalence;
- dynamical equivalence;
- multiscale causal equivalence;
- physical mechanism equivalence.

The correct question is not simply:

Are these mechanisms the same?

but:

At which causal level, under which interventions, are which generative invariants preserved?

6.3 Natural causal abstraction

A macro-variable is not natural merely because it compresses data. Candidate features of a natural causal abstraction include:

- stable generative dynamics;
- intervention coherence;
- robustness to microvariation;
- non-post-hoc mapping;
- useful factorization;
- counterfactual stability;
- scope consistency;
- multiple realization;
- predictive compression.

No single “naturalness score” is assumed.

6.4 Intervention triangulation

Macro-interventions are rarely magical $do(M = m)$ operations that change one high-level variable while preserving every micro-detail. GCP therefore distinguishes:

- boundary-state intervention;
- internal-state intervention;
- parameter intervention;
- coupling intervention;
- topology or mechanism intervention;
- boundary-constitution intervention;
- history intervention.

A high-level variable is better supported when multiple physically different intervention routes converge on the same relevant macro-level response.

7. Psychophysical Bridge Theory Space

Let the world's true psychophysical relations be

$$\Psi_E, \Psi_C, \Psi_V, \Psi_A, \Psi_T, \quad (14)$$

for existence, character, value, attribution, and continuation.

Scientific theories are candidates

$$\Phi_E, \Phi_C, \Phi_V, \Phi_A, \Phi_T, \dots \quad (15)$$

Evidence constrains the Φ 's; it does not directly reveal the Ψ 's.

For existence, four top-level classes organize much of the theory space. They are idealized comparison families rather than a proof that all possible theories are mutually exclusive or exhaustively partitioned; hybrid theories can combine features of more than one class.

7.1 Class A – causal-structural determination

Some natural causal-generative structure is sufficient to fix phenomenal existence.

7.2 Class B – finer-physical determination

Macro causal organization is insufficient; additional physical facts matter.

7.3 Class C – causal extension

Phenomenal facts contribute causal efficacy not contained in the current physical description.

7.4 Class D – empirically inert extra facts

Even complete physical facts can remain compatible with different phenomenal facts that never affect public evidence.

Empirical consciousness science can progressively discriminate among equivalence classes wherever competing theories generate different observable consequences, without first settling every metaphysical interpretation. Class-D residues, by definition, remain empirically underdetermined insofar as their differences never affect public evidence.

If two theories generate the same predictions under every physically possible relevant transformation, then

$$\Phi_1 \sim_{\text{emp}} \Phi_2. \quad (16)$$

They belong to the same empirical equivalence class (called a bridge kernel here) even if their metaphysical language differs.

8. Human Experience as Anchor, Not Template

Humans provide the strongest currently available situated first-person anchor. This is an epistemic advantage, not a license to define all possible experience in human coordinates.

8.1 The Human Anchor Transformation Atlas

The preferred unit of calibration is not a static report such as “I see red.” It is a controlled transformation:

$$s_0 \xrightarrow{\tau} s_1 \quad (17)$$

with a typed signature recording changes in:

$$\Sigma_H(\tau) = \langle \textit{Causal}, \textit{Character}, \textit{Value}, \textit{Attribution}, \textit{Temporal}, \textit{Report}, \textit{Behavior}, \textit{Reversibility} \rangle. \quad (18)$$

Each field may be preserved, changed, split, merged, unknown, or measurement-disturbed.

Situated first-person access supplies the principal anchor adopted here for the presence and change of current manifestation, but classification, memory, causal explanation, and metaphysical interpretation can still be mistaken.

8.2 Cross-person calibration

Absolute qualitative identity across people may be inaccessible. Yet a science can compare:

- similarity orderings;
- transition topologies;
- discrimination thresholds;
- reversal patterns;
- invariant transformations.

Thus relational structure can be scientifically useful even when absolute quale identity remains unresolved.

8.3 Cross-substrate covariance

For an artificial system, the first question is not:

Does it have the human experience of red?

A better question is:

Is there a causal transformation in the target whose intervention-response structure corresponds to a transformation that is phenomenally calibrated in the source?

Human neural population dynamics and artificial distributed latent dynamics need not be physically identical.

But a crucial type constraint remains:

$$\text{Representational or causal covariance} \not\Rightarrow \text{phenomenal existence.} \quad (19)$$

A transported character relation does not automatically transport existence. If target experience is absent, its phenomenal character is undefined, rather than a known character with a zero existence score. Conditional transport must retain this antecedent, as Appendix H makes explicit.

9. Psychophysical Transport

Transport is bridge-specific:

$$\text{Transport}_{\Phi_k}(S, T), \quad (20)$$

not an untyped relation $\text{Transport}(S, T)$.

For the same source and target:

- character may be partially transportable;
- value may remain unknown;
- attribution may be weak;
- existence may remain unresolved;
- continuation may be out of scope.

Transport may be:

- partial;
- directional;
- one-to-many;
- many-to-one;
- noninvertible.

A target system may contain distinctions with no source counterpart.

A transport claim can be represented schematically as a certificate

$$\mathcal{C}_{S \rightarrow T}^k = \langle S, T, k, \Gamma, J, \alpha, \text{Inv}, M, \mathcal{E}, \mathcal{L}, V \rangle,$$

where k is the target claim type; Γ records bridge and scope assumptions; J pairs transformations; α records the declared causal or structural correspondence; Inv records preserved invariants; M is the measurement model; $\mathcal{E}$ records evidence and dependencies; $\mathcal{L}$ records typed losses; and V records version/provenance. The certificate is an auditable claim record, not a certification that the target is conscious.

Each transport edge should therefore record:

- source and target system/version/regime;
- target claim type;
- intervention pairing;
- causal abstraction;
- preserved invariants;
- measurement model;
- provenance;
- scope;
- loss;
- failure conditions;
- version.

Transport loss may concern:

- character;
- value;
- timing;
- perspective;
- resolution;
- context.

Multi-hop transport is not automatically transitive:

$$\text{Transport}(S_1, S_2) \wedge \text{Transport}(S_2, S_3) \not\Rightarrow \text{Transport}(S_1, S_3). \quad (21)$$

The intermediate step may lose exactly the distinction required for the final inference.

The calibration network propagates constraints subject to compatibility guards. Equation (21) denies unconditional transitivity; compatible relations about the same intermediate value do compose (Appendix H).

10. Phenomenal Mereology and Perspective

The regression cases do not license either pure atomism or pure holism as a universal default. Longstanding work on phenomenal unity and co-consciousness likewise shows that unity and subject structure require substantive analysis rather than a trivial partition rule [20].

A working framework must permit:

- local phenomenal aspects;
- whole-level character;
- relational emergence;
- nesting;
- overlap;
- local/global coexistence;
- absorption;
- fission and fusion;
- bidirectional part-whole constraints.

No single property among simultaneity, synchrony, content identity, shared memory, direct communication, a global workspace, a common controller, a self-model, attention, or object binding is adopted as sufficient for co-manifestation.

Moreover, pairwise relations may be insufficient:

$$C(P, Q), C(Q, R), C(P, R) \not\Rightarrow C(P, Q, R). \quad (22)$$

In equation (22), C is a generic joint-manifestation predicate, not a stipulated equivalence relation. Pairwise compatibility need not establish one common three-way manifestation. Theories defining co-consciousness as an equivalence relation make a stronger commitment and should be assessed on that basis. Higher-order or hypergraph-like phenomenal relations remain representable.

The subject is therefore not introduced as a primitive box. It is derived, if possible, from perspectival organization.

This stance is also constrained by recent methodological work on counting conscious subjects [21], on overlapping minds and welfare aggregation [22], and on cases in which subject number may fail to be a determinate whole number [23]. GCP therefore treats integer subject count as a downstream, theory-relative summary rather than a guaranteed primitive.

11. Continuation and Lineage

The theory distinguishes causal type, causal token, phenomenal token, perspective, and lineage.

A checkpoint restore can preserve memory while failing to preserve one unique actual history. Two independently running copies can be type-identical yet token-distinct. A gradual substrate replacement can preserve causal organization without deciding whether the same phenomenal lineage survives. A fusion can create many-to-one continuation; a fission can create one-to-many successors.

Thus a mature continuation theory should output a **lineage topology**, not force a binary identity relation.

Candidate continuation relations must be separately specified for:

- ordinary continuous evolution;
- pause/resume;
- destroy/reconstruct;
- nondestructive copy;
- destructive upload;
- gradual replacement;
- symmetric fission;
- fusion;
- split-merge cycles.

Memory, self-model, causal history, and phenomenal continuation are separate fields.

12. Theory Identifiability and Scientific Method

12.1 More data are not always more discriminating

Two theories can agree on every existing human dataset. A million more observations from the same transformation family may reduce noise without distinguishing them.

A new transformation family can therefore be more valuable than a larger sample.

12.2 Structural versus practical identifiability

Two theories may be:

- structurally indistinguishable even with arbitrarily precise data from a fixed experiment family;
- structurally distinguishable but practically inseparable because of noise, technology, or ethics.

“Currently indistinguishable” is not the same as “in-principle empirically equivalent.”

12.3 Bridge space

Let $\mathfrak{B}$ be the candidate bridge space. Under a specified model-rejection rule, an informative experiment ideally produces:

$$\mathfrak{B}_0 \rightarrow \mathfrak{B}_1 \subset \mathfrak{B}_0. \quad (23)$$

In noisy settings, evidence generally reweights candidate models rather than logically eliminating them. The displayed set restriction is an idealization for a declared rejection criterion; later expansion of the candidate space remains possible.

12.4 Holdout transformations

A strong defense against bridge overfitting is not merely a new participant. It is a new:

- intervention family;
- causal regime;
- architecture;
- substrate;
- temporal organization.

The theory version and predictions should be frozen before the holdout test. This emphasis is compatible with recent preregistered adversarial theory testing in consciousness science, which makes divergent predictions explicit before evaluation [17].

12.5 Reverse prediction

A cross-substrate theory is strengthened if a bridge developed partly from artificial or synthetic systems predicts a previously unknown dissociation in humans or animals.

This reduces anthropomorphic fitting.

12.6 Evidence topology

Evidence items can be dependent. Several model reports produced by the same training procedure are not independent evidence. Several experiments using the same consciousness labeling assumption are not independent anchors.

A mature theory therefore requires both:

- an **Anchor Audit**;
- an **Evidence Dependency Graph**.

12.7 Bridge contracts

Every empirical bridge should state:

1. Target Gap;
2. input type;
3. output type;
4. causal scale;
5. temporal scale;
6. tokenization assumptions;
7. scope;
8. dependencies;
9. invariances;
10. sensitivities;
11. conditional assumptions;
12. failure semantics;
13. version and provenance.

If a bridge cannot state what would force revision, scope reduction, or withdrawal, it has accumulated **prediction debt**.

13. Why a Positive Theory Was Still Needed

The preceding architecture can diagnose invalid inference without selecting a positive psychophysical law. A broad theory must eventually take explanatory risks.

Accordingly, a sequence of positive causal-structural hypotheses was developed and deliberately exposed to adversarial thought experiments.

This development should not be read as five independent theories all simultaneously endorsed. It is a failure-preserving research lineage.

14. Positive Theory Development: v0.1-v0.5

14.1 v0.1 – primitive counterfactual return

The first affirmative hypothesis proposed that an actually implemented organization supports minimal whole-level phenomenality when endogenous causal distinctions return irreducibly to their point of origin.

In a finite nonnegative linear model:

$$x_{t+1} = Wx_t + u_t, \quad (24)$$

a cut deleting all between-block weights, with perturbations measured about matched baselines and identical external inputs, yielded

$$R_\pi(k) = W^k - W_\pi^k. \quad (25)$$

The response difference under a matched perturbation was exactly

$$\delta x_k - \delta x_k^\pi = (W^k - W_\pi^k)\delta x_0. \quad (26)$$

For nonnegative W , at least two vertices, and all nontrivial bipartitions, the criterion that every vertex has some positive diagonal cut-return response is equivalent to strong connectivity of the positive-edge graph (Appendix G.1). It admits arbitrarily weak reciprocal coupling. A one-vertex case requires a separate self-feedback condition because its bipartition family is empty. This criterion alone does not explain differentiated phenomenology.

Lesson: recurrence or return alone is not a sufficient explanatory structure.

14.2 v0.2 – expressed contextual return

The next hypothesis required not only return but **context-sensitive return**: another internal distinction had to modify the effect of the returning distinction.

It also added an uninterrupted-evolution control to detect cases in which the diagnostic reset itself created the apparent organization.

This eliminated a real artifact. For example, a parity-collapse mechanism

$$F(a, b) = (a \oplus b, a \oplus b) \quad (27)$$

can show a strong reset echo even though ordinary two-step evolution collapses every starting state to 00. The reset prevented cancellation and manufactured an apparent retained distinction.

Lesson: a measurement intervention can create the property it was intended to diagnose.

But the revised criterion remained permissive toward simple mechanisms.

14.3 v0.3 – complexity and causal repertoire

The next step separated several quantities hidden inside the word “complexity”: This separation is motivated in part by earlier work combining differentiation and integration [14] and by perturbational or spontaneous-complexity studies in human consciousness-related conditions [10,11]; the finite constructions below do not inherit the empirical validation of those studies.

- physical size;
- description length;
- output entropy;
- interventional differentiation;
- interaction dependence;

- contextual return;
- temporal retention.

For an intervention-response channel

$$K_h(y \mid z) = P^h(y \mid do(X_0 = z)), \quad (28)$$

the interventional capacity

$$C_h = \max_{\mu} I_{\mu}(Z; Y_h) \quad (29)$$

measures preparation-dependent information, not consciousness; it uses standard Shannon information theory rather than a new consciousness-specific quantity [16].

A decisive counterexample was a five-bit dense parity broadcaster:

$$F(x_1, \dots, x_5) = (p, p, p, p, p), \quad p = \bigoplus_{i=1}^5 x_i. \quad (30)$$

Its dependency graph is complete, yet at every positive horizon it preserves only one bit of preparation-dependent differentiation. By contrast, the reversible four-bit map $F(a, b, c, d) = (b, c, d, a \oplus b)$ preserves four bits at every finite horizon. Both graphs are strongly connected; connectivity does not determine retained capacity. Appendix G.2 supplies the full construction.

Lesson: return magnitude, size, entropy, and differentiated capacity can rank systems differently.

Independent memory also showed that large capacity does not imply one integrated whole.

14.4 v0.4 – common witness and naturally used mediation

The next failure came from pooling favorable properties across incompatible operating regimes.

A five-bit system could have:

- five-bit capacity in an independent-storage regime;
- complete state-variable dependency in a parity-broadcast regime;

yet no single regime realized both.

Thus:

$$(\exists \Gamma A(\Gamma)) \wedge (\exists \Gamma B(\Gamma)) \not\Rightarrow \exists \Gamma [A(\Gamma) \wedge B(\Gamma)]. \quad (31)$$

The correct object is the achievable set of properties inside compatible regimes, not a vector assembled from coordinatewise maxima.

A simple gated mediator makes the regime issue explicit. Let a diagnostic output be $y = a \oplus (c \wedge m)$. In a natural regime with $c = 0$, replacing m cannot change y ; a diagnostic reset that also sets $c = 1$ makes m causally effective. This supports only a claim about the diagnostic regime unless transfer back to the natural regime is justified. It does not show that all reset-based methods are invalid.

A third counterexample showed that an invertible coordinate recoding could make two physically independent loops appear densely integrated. When the true intervention algebra was transported with the encoding, the original physical decomposition remained recoverable (Appendix G.3).

Lesson: integration, mediation, and complexity must be properties of one common physical witness, with intervention semantics preserved.

14.5 v0.5 – coalitional return and protected distinctions

A single-carrier intervention can produce a false negative in redundant systems.

Suppose an odd number $n = 2r + 1$ of internal carriers all return the same distinction and the system uses majority decoding. Replacing any coalition smaller than $r + 1$ leaves the output unchanged. Yet replacing $r + 1$ or more carriers changes the output.

Fix a noiseless repetition code whose n carriers all encode the same bit. Replace every carrier in S by the opposite bit, leave the others fixed, and compare deterministic decoded outputs. Define $\|J_S\|_{TV}$ as the total-variation distance between their point-mass distributions. Then

$$\|J_S\|_{TV} = \mathbf{1}\{|S| \geq r + 1\}. \quad (32)$$

Thus genuine causal participation can be hidden from every singleton test.

Moreover, with independent carrier error probability $p = 0.1$, redundancy produces opposite trends in reliability and single-carrier pivotality:

carriers	correct logical return	single-carrier pivotality
1	0.900000	1.000000
3	0.972000	0.180000
5	0.991440	0.048600
7	0.997272	0.014580
9	0.999109	0.004593

For $n=2r+1$, the reliability is $\sum_{k=0}^r \binom{n}{k} p^k (1-p)^{n-k}$. The pivotality in the table is the probability that the other $2r$ noisy carriers tie: $\binom{2r}{r} p^r (1-p)^r$. It is a stochastic Boolean influence, distinct from the noiseless coalition contrast in equation (32). Reliability rises while this singleton influence falls; neither quantity measures experience.

But error correction does not identify consciousness. In the finite binary construction used in Appendix G.4, any finite binary update rule can be wrapped by decode-update-re-encode error correction. This shows that repair is a broadly available implementation resource rather than, by itself, a phenomenal selector.

Lesson: the relevant causal realization must be redundancy-aware and coalitional. Local indispensability is not constitutive membership.

15. The Current Foundational Proposal

The failure sequence suggests that a viable causal-structural theory of phenomenal realization must refer to a richer object than recurrence, capacity, integration, or robustness alone.

15.1 Common-Witness Relational Organization

For a candidate physical organization U , regime Γ , and temporal window H , define schematically

$$\mathfrak{J}(U, \Gamma, H) = \langle P, \mathcal{I}, \text{Prov}, \mathcal{A}, \mathcal{D}, \mathcal{M}, \mathcal{C}, \mathcal{T} \rangle, \quad (33)$$

where:

- P is the actual causal transition structure;
- $\mathcal{I}$ is the physically admissible intervention algebra;
- Prov is actual generative provenance;
- $\mathcal{A}$ is the set of jointly achievable causal properties within one compatible regime;
- $\mathcal{D}$ records internally available and jointly expressible distinctions;
- $\mathcal{M}$ records naturally operative mediation relations;
- $\mathcal{C}$ records coalition-sensitive participation and redundancy;
- $\mathcal{T}$ records temporal retention, recurrence, and path dependence.

A **common witness** requires that the evidence invoked to explain one phenomenal organization be jointly attributable to the same support, regime, temporal scope, intervention semantics, and calibration. Counterfactual experiments can occur in different trials; what is forbidden is assembling an impossible “super-system” from incompatible regimes.

15.2 Relational-Realization Hypothesis

The positive proposal is:

RRH – Relational-Realization Hypothesis.

In the causal-structural theory family, when an experiential episode occurs, its relational phenomenal organization is realized by a natural common-witness causal organization whose relevant distinctions are actually instantiated, jointly expressible, internally used, temporally organized, intervention-covariant, and represented with their full redundancy-aware coalitional dependence.

This is stronger than saying “complexity matters.” It identifies what a proposed realization is not allowed to borrow:

- not inactive storage;
- not random output entropy;
- not a property present only under a measurement-created mode;
- not a coordinate artifact;
- not a maximum collected from incompatible regimes;
- not an analyst’s external reconstruction;
- not a count of redundant copies;
- not a property inferred from one indispensable component in a distributed system.

15.3 What RRH does and does not claim

RRH is a **realization hypothesis for phenomenal relations**. It does not yet by itself prove a universal law of phenomenal existence.

To keep the stronger existence proposal from resting on undefined adjectives, distinguish two physical notions relative to a frozen protocol family $\mathcal{Q}$, a declared grain, temporal window, and (for empirical work) tolerance ϵ :

- **Nondegeneracy**, $\text{ND}_\epsilon(\mathfrak{J})$: the common-witness response family contains at least two admissible internally generated alternatives whose target-relevant response distributions remain distinguishable beyond the declared tolerance. In ideal finite models one can set $\epsilon = 0$; no universal empirical ϵ is assumed.
- **Query-relative irreducibility**, $\text{IRR}_{\epsilon, \mathcal{Q}}(\mathfrak{J})$: there is no admissible nontrivial partition together with partition-local surrogate mechanisms that reproduces the full target-relevant intervention-response family in $\mathcal{Q}$ within tolerance ϵ , while preserving the same intervention semantics and provenance constraints.

For sufficiently informative query families and appropriately restricted surrogates, these definitions seek to discriminate beyond graph connectivity. That relationship is not guaranteed for every choice of query family; a vacuous family cannot support a substantive irreducibility claim. They are also representation-sensitive unless the intervention algebra is transported with any recoding.

A strong **Existence Bridge Conjecture schema (RRH-E)** can then be stated as a candidate for later empirical specification:

$$E_U^{\text{whole}} \iff \text{Active}(U, \Gamma) \wedge \text{ND}_\epsilon(\mathfrak{J}) \wedge \text{IRR}_{\epsilon, \mathcal{Q}}(\mathfrak{J}), \quad (34)$$

with the grain, protocol family, temporal scope, tolerance, and provenance fixed before the phenomenal verdict is evaluated. In idealized exact models, $\epsilon = 0$; in empirical applications, the framework retains the full tolerance-indexed profile rather than asserting a universal consciousness cutoff.

Equation (34) is intentionally a **strong conjecture**, not an established law. The project has not produced independent evidence sufficient to establish either direction of the biconditional. Its role is to expose the commitments a causal-structural branch would have to make. It is not yet an operationally complete law: natural-unit selection, admissible surrogates, a principled protocol family, and an independently supported phenomenal measurement model remain to be specified. Freezing an arbitrary protocol does not solve those selection problems. In particular, a query family that ignores intercomponent responses can change an irreducibility verdict without changing the physical system. An eventual existence law must justify its physically relevant query family or prove suitable invariance across equivalent families. For a one-component description, the absence of nontrivial partitions makes irreducibility vacuous unless a separate minimal-unit rule is given. A biological-naturalist theory may reject sufficiency by adding substrate-sensitive conditions. A causally extended theory may reject the completeness of the causal description itself. A future version may also retreat from the biconditional if adversarial evidence defeats necessity or sufficiency; such a retreat must be versioned rather than silently reinterpreted.

15.4 Richness principle

The framework makes a stronger claim about **explanatory richness** than about existence.

If a declared complete structural signature $r(z)$ is supposed to determine a phenomenal query $f(z)$,

$$f(z) = g(r(z)), \quad (35)$$

then

$$|f(\mathcal{Z})| \leq |r(\mathcal{Z})|. \quad (36)$$

A structural model cannot explain phenomenal distinctions that it itself identifies as the same structural state.

Thus complexity constrains the richness that a structural theory can explain, but complexity alone does not establish phenomenal presence.

16. Redundancy, Constitution, Scaffold, and Context

Ablation is not enough to identify constitutive structure.

If removing X changes a target phenomenon, X may be:

- constitutive core;
- enabling scaffold;
- contextual condition;
- accidental implementation necessity.

Conversely, redundancy can make a truly participating component individually unnecessary.

The framework therefore distinguishes:

16.1 Core

The organization directly belonging to the target bridge.

16.2 Scaffold

A structure that enables, maintains, repairs, or reads the core while being replaceable in implementation.

16.3 Context

A condition selecting the regime in which the bridge is valid or changing how the same core maps to phenomenal output.

Role-preserving replacement is therefore more informative than simple removal.

A component can be constitutive at one scale and scaffold at another. Constitution is scale- and gap-relative.

17. Fixed Points, Multiple Solutions, and Selection

A typed constraint network may produce:

- no phenomenal solution;
- one solution;
- several solutions.

Multiple formal solutions must be classified before acquiring ontological significance.

They may represent:

1. gauge or labeling multiplicity;
2. epistemic underdetermination;
3. a genuinely stochastic or set-valued psychophysical law;
4. simultaneous phenomenal plurality.

The selection heuristics considered here do not supply an independently justified universal selection law.

- continuity helps with lineage but not initial existence;
- stability can be bridge-relevant but does not select ontology;
- symmetry forbids arbitrary label breaking but does not decide local versus global organization;
- minimality and maximality are theory-choice heuristics, not world laws;
- maximal integration is itself a substantive attribution theory;
- stochastic selection requires a real probability law, not ignorance renamed as chance.

Thus the framework permits genuine underdetermination rather than silently hiding it.

18. Prediction Completeness and Modular Composition

A theory may be complete for one target and incomplete for another.

The following blank coverage template records the targets a specified theory must address. Question marks are unfilled entries, not results of a completed assessment:

Transformation	Existence	Character	Value	Perspective	Continuation	Multiplicity
report	?	?	?	?	?	?
manipulation						
substrate	?	?	?	?	?	?
replacement						
fission	?	?	?	?	?	?
memory	?	?	?	?	?	?
deletion						
coupling	?	?	?	?	?	?
change						

Unknown cells are scientific information.

Modules may be composed only when their:

- types;
- scales;
- scopes;
- tokenization assumptions;
- intervention semantics;
- dependencies;

are compatible.

Local correctness does not guarantee global correctness. Composition therefore requires integration tests for non-additivity, order effects, coupling, whole-level effects, and cyclic dependence.

19. Regression Against the Thought-Experiment Archive

The supplied working manuscript reports an extensive exploratory archive. The present paper does not certify its entry count or rely on inaccessible notes as evidence. The auditable basis of this revision is the explicitly stated cases, proofs, and executable constructions provided here. Conceptual cases are not empirical confirmations.

The present paper compresses those entries into regression families.

Family	Current treatment
perfect first-person report without phenomenality	report and phenomenality remain different types
perfect color discrimination	discrimination does not solve existence
reward reversal	reward sign is not phenomenal valence
damage detection and avoidance	neither defines suffering
fluent language without experience	language is evidential, not constitutive by definition
phenomenality without self-model	permitted
superintelligence without phenomenality	permitted
silent phenomenality	permitted
static source code	not an active implementation of the described process
fixed replay	may reproduce trajectory without live counterfactual generation
full transition lookup	may preserve a causal process if it truly realizes the transition/intervention law
identical independent copies	type identity does not collapse token provenance
millions of software labels	logical multiplicity does not determine process or phenomenal multiplicity
shared computation	must be represented once at the causal provenance level; phenomenal sharing remains a bridge question
external memory	internal/external status depends on causal role and regime, not API location
one CPU/many processes	hardware count does not determine process count
many machines/one distributed process	distributed realization is allowed
uniform slowing	pure temporal reindexing may preserve relations; physical slowing may not
differential delay	can change dynamics and bridge predictions
gradual substrate replacement	causal preservation is conditional; finer-physical theories remain competitors
pause/resume	causal continuity and phenomenal continuation remain distinct
checkpoint restore	memory restoration does not settle lineage
nondestructive copy	branching without unique identity is representable
destructive upload	information preservation does not by itself establish patient preservation
fission	one-to-many continuation can be represented
fusion	many-to-one continuation can be represented
local perspectives/global perspective	coexistence, absorption, and exclusion remain distinct theories
local pain/global happiness	value cannot be naively summed across unresolved mereology

Family	Current treatment
weak coupling	finite-horizon structural effects can vanish continuously; no arbitrary subject threshold follows
strong connectivity	rejected as a sufficient existence criterion
reset-created return	blocked by uninterrupted/common-mode controls
coordinate recoding creating dense graph	blocked by intervention-preserving recoding
large capacity from independent modules	does not establish a new global whole
random high entropy	does not establish preparation-dependent differentiation
high return/low repertoire parity broadcaster	blocks return magnitude as richness
incompatible modes donating different maxima	blocked by common-witness requirement
unused mediator activated by reset	blocked by natural-mode mediator replacement
redundant returning carriers	requires coalition-sensitive interventions
single-carrier insensitivity in robust code	not interpreted as absence of causal participation
error correction	not a consciousness selector
same decoded content/different robustness margin	coarse state can be inadequate for future response
physically rebuilt recoding	distinguished from mere redescription
hidden higher-order relation	pairwise nulls do not prove higher-order absence
dream without current sensory input	current external input not required by framework
stable simple phenomenal content	low current content diversity does not imply inactive mechanism
memory loss with current experience	allowed; memory and manifestation are different targets
perfect zombie world	ruled in or out only by the selected existence bridge, not by the neutral framework

The key result of the regression program is therefore not that every case has a final truth value. It is that a very large set of apparently different problems can be expressed without inventing a new primitive for each case.

The framework localizes several remaining unknowns in bridge laws and measurement assumptions; it does not eliminate every conceptual disagreement.

20. Empirical Research Program

20.1 Phase 1 – Human Anchor Transformation Atlas

Build a library of reversible, low-risk transformations with:

- first-person change;
- causal target;
- collateral effects;
- behavioral change;
- report change;
- reversibility;
- adaptation;
- uncertainty;
- provenance.

20.2 Phase 2 – Dissociation library

Prioritize two-way dissociations among:

- character and value;
- content and attribution;
- report and manifestation;
- memory and current experience;
- agency and ownership;
- integration and perspective.

20.3 Phase 3 – Natural causal mapping

Map stable phenomenal transformations onto:

- natural causal abstractions;
- mechanism regimes;
- intervention families;
- dynamical modes.

20.4 Phase 4 – Synthetic calibration systems

Use transparent systems in which one can control:

- coupling;
- memory;
- recurrence;
- timing;
- substrate;
- fission/fusion;
- redundancy;
- shared computation.

Synthetic systems need not resemble humans. Their role is to discriminate bridge structure.

20.5 Phase 5 – Cross-substrate covariance

Compare transformation signatures across:

- humans;
- non-human animals;
- synthetic systems;
- artificial agents.

The first objective is not binary classification but elimination of bridge kernels that cannot explain the covariance structure.

20.6 Phase 6 – existence-kernel discrimination

The strongest substrate-independence tests require crossing two axes:

1. preserve macro causal organization while changing substrate;
2. preserve substrate class while changing causal organization.

If anchored phenomenal evidence tracks causal organization and not substrate across sufficiently varied cases, support shifts toward causal-structural theories. If substrate changes matter after the relevant causal organization is preserved, finer-physical theories gain support.

Human-only data cannot establish substrate independence.

21. Concrete Falsification Conditions

An empirically specified bridge should state what would make it fail. The conditions below apply jointly to a bridge, its causal model, and its measurement assumptions; disagreement does not uniquely identify which member failed. The neutral architecture is assessed by auditability and usefulness, whereas a specified substantive bridge is assessed by its predictions.

21.1 Existence-bridge failure

If well-supported phenomenality persists in a regime where the theory's supposedly necessary common-witness organization is absent under an independently justified causal description, the necessity claim fails.

If a system satisfies the full proposed structural condition while independent evidence strongly supports absence of phenomenality, the sufficiency claim is challenged. Such negative evidence is particularly difficult to obtain, and must not merely be absence of report. No currently available substrate-general gold standard is assumed. A conjecture with no discriminating measurement model remains empirically underdetermined, even when its logical failure condition is clear.

21.2 Character-law failure

If two realizations preserve the complete declared relational realization and calibration yet show a reproducible difference in the associated phenomenal relation, the character bridge is incomplete.

21.3 Cross-substrate failure

If a source-to-target transport certificate predicts a relational preservation and a held-out transformation violates it despite preserved scope and intervention semantics, the transport edge must be revised or withdrawn.

21.4 Perspective-law failure

Any candidate perspective rule must survive local/global coexistence, overlap, weak coupling, distributed systems, shared memory, and symmetric fission without ad hoc relabeling.

21.5 Continuation-law failure

A continuation theory must separately predict pause/resume, restore, copy, gradual replacement, fission, and fusion. A rule that answers only ordinary continuous biological persistence is incomplete.

22. What the Framework Currently Explains or Constrains

The integrated theory explains several kinds of facts at different strengths.

22.1 It constrains cross-type inferences

Examples include intelligence-to-phenomenality, reward-to-suffering, integration-to-subjecthood, memory-to-identity, and software-count-to-subject-count when these are asserted as entailments or definitions. The framework does not deny that these variables can carry defeasible evidence under an explicit measurement and bridge model.

22.2 It motivates transformation-structured comparison

Static resemblance can be trained, copied, replayed, or produced by different mechanisms. Transformation covariance can therefore be more discriminating than static resemblance when the intervention semantics are well controlled.

22.3 It permits local rather than universal phenomenal coordinates

Transport can be partial, directional, noninvertible, and local.

22.4 It separates distinct notions of causal complexity

Return, repertoire, capacity, entropy, integration, robustness, and temporal retention can dissociate.

22.5 It diagnoses simple-ablation failures under redundancy

Coalitional participation can be real even when no component is individually necessary.

22.6 It records the failure history of diagnostic proposals

The v0.1-v0.5 sequence records why the earlier sufficiency claims were withdrawn rather than treating all exploratory proposals as jointly established.

What the theory does **not yet explain** is equally important:

- why any physical organization has phenomenal existence at all;
- why one organization feels red rather than another way;
- what determines the sign and structure of phenomenal value;
- exactly what forms one perspective rather than several;
- exactly what constitutes phenomenal continuation.

These are the remaining law-level problems.

23. Originality and Claim Discipline

The paper does **not** claim to originate:

- multidimensional consciousness;
- non-anthropocentric comparison;
- interventionism;
- phenomenal invariants;
- causal abstraction;
- recurrence;
- differentiation and integration;

- information closure;
- organizational invariance;
- subject-count problems;
- fission or branching identity.

Work on machine-consciousness tests [18], computational implementation [19], phenomenal unity [20], subject counting [21,23], overlapping minds [22], and AI welfare subjects [24] further constrains the novelty claim. Its candidate contribution is the integrated architecture:

$$\begin{array}{l}
 \text{Causal reality + typed psychophysical gaps} \\
 + \text{ transformation-first calibration} \\
 + \text{ bridge-specific transport} \\
 + \text{ empirical identifiability} \\
 + \text{ failure-preserving theory development}
 \end{array} \tag{37}$$

together with the redundancy-aware relational-realization proposal that emerged from the adversarial v0.1-v0.5 sequence.

The graph, factorization, Boolean-majority, and repetition-code arguments use standard mathematical ideas; no priority is claimed for these elementary facts. Their role here is to constrain proposed psychophysical explanations. Historical priority for the complete architecture would require a systematic literature review beyond the focused comparison undertaken here.

24. Discussion

24.1 Why this is not just methodological caution

A common objection is that a framework which repeatedly says “another bridge is needed” risks becoming scientifically inert.

The answer is twofold.

First, type distinctions constrain permissible inference and require an explicit bridge. These are logical and methodological constraints, not independently testable predictions about which systems experience.

Second, the framework supports positive competing hypotheses. The causal diagnostics can be contradicted by explicit finite constructions. RRH and its existence extension become empirically testable only when unit selection, protocols, and measurement bridges are sufficiently specified. Formal revision of a diagnostic is not experimental falsification of a consciousness law.

Thus GCP is not a refusal to theorize. It is a discipline for theorizing without letting terminology perform the work of evidence.

24.2 Why simple circuits remain difficult

The theory has not derived a universal complexity threshold separating all ordinary circuits from all conscious systems.

That is not an omission that can be repaired by choosing a large number after seeing the examples.

If very simple systems are truly phenomenally dark, the theory needs a principled reason. If some simple systems instantiate primitive phenomenality, the theory should accept that consequence rather than hide it.

Complexity is better interpreted as constraining **phenomenal differentiation or richness** under a structural bridge than as automatically creating existence.

24.3 Why the framework remains cross-substrate

The positive causal-structural branch is substrate-general by design, but substrate independence is not assumed as a fact.

A finer-physical theory may ultimately win. If so, the type-safe architecture still survives: the relevant physical facts simply become part of the existence bridge and transport invariant portfolio.

Thus GCP is broader than computational functionalism.

24.4 The role of exploratory thought experiments

The role of the exploratory archive was to generate candidate objections. Only explicitly reconstructed cases support the formal claims in this paper.

Repeated cases removed naive equations such as:

$$\text{report} = \text{experience},$$

$$\text{reward} = \text{pain},$$

$$\text{feedback} = \text{consciousness},$$

$$\text{capacity} = \text{consciousness},$$

$$\text{ablation sensitivity} = \text{constitutive membership}.$$

The surviving architecture is valuable precisely because many initially plausible shortcuts did not survive.

25. Conclusion

General Cross-Substrate Phenomenology begins from a simple observation: consciousness research across humans, animals, and artificial systems cannot safely rely on a single human-centered scale or on untyped analogies between intelligence, report, reward, memory, integration, and phenomenality.

The framework therefore separates causal-generative reality, phenomenal manifestation, perspectival organization, continuation, and evidence. It distinguishes Existence, Character/Value, At-

tribution/Perspective, and Continuation as separate psychophysical problems. It treats human experience as an epistemic anchor without making human architecture the universal ontological template. It compares systems through controlled transformations, scoped bridge hypotheses, invariant portfolios, and explicit transport loss.

The positive causal-structural branch has also been forced to evolve. Mere recurrence collapsed into strong connectivity and was too permissive. Contextual return exposed measurement-created artifacts. Complexity analysis showed that feedback, entropy, capacity, and integration can dissociate. Common-witness constraints blocked the assembly of incompatible properties across regimes. Intervention-preserving recoding blocked artificial integration created by coordinate choice. Coalition analysis showed that real causal participation can be distributed and redundant, making single-component necessity an unreliable guide.

The current affirmative proposal is therefore a **redundancy-aware relational-realization hypothesis**: if phenomenal relations are causally realized, the relevant candidate is an actually instantiated, naturally delimited, common-regime organization of jointly expressible, internally used, temporally extended, intervention-covariant, redundancy-aware causal distinctions.

That proposal supplies a substantive constraint on candidate realizations; a fully specified version still requires a defensible measurement connection for empirical falsification. It is also not yet an empirically established universal law of phenomenal existence. The remaining task is not to generate an unlimited number of new concepts, but to discriminate the surviving existence bridges—causal-structural, finer-physical, causally extended, or empirically inert—using transformation-diverse, cross-substrate, ethically constrained evidence.

The practical contribution is to require each comparison to specify its actual causal organization, phenomenal target, bridge assumptions, admissible transformations, and unresolved conclusions. This creates a common interface for competing theories while leaving their empirical and meta-physical disagreements visible.

Appendix A. Minimal Formal Skeleton

A.1 Types

- D : natural causal-generative objects and patterns
- P : phenomenal facts or structured manifestations
- Ω : perspectival or mereological organization
- L : continuation or lineage relations
- E : evidence objects
- Φ : bridge hypotheses

These are not six substances. They are type distinctions.

A.2 Transformation signature

For a transformation $\tau : D \rightarrow D'$,

$$\Sigma_{\Phi}(\tau) = \langle Ex, Ch, Val, Persp, Cont, Mult \rangle. \quad (A1)$$

Fields may be:

- preserved;
- changed;
- created;
- terminated;
- split;
- merged;
- unknown;
- out of scope.

A.3 Well-formedness rules

WF-1. Every cross-type inference requires a bridge.

WF-2. Every bridge carries scope.

WF-3. Every evidence item carries provenance.

WF-4. Every equivalence is indexed by type and scale.

WF-5. Every count declares a tokenization rule.

WF-6. Every transformation claim is layered.

WF-7. Unknown remains unknown.

Appendix B. Canonical Bridge-Requirement Set

Each row requires an additional bridge; the table does not deny evidential relevance.

Starting quantity	Conclusion requiring a bridge
Report	Phenomenality
Self Model	Subject
Reward	Phenomenal Value
Avoidance	Suffering
Intelligence	Phenomenality
Integration	Phenomenal Unity
Shared Memory	Shared Perspective
Same Trajectory	Same Generation
Causal Equivalence	Phenomenal Equivalence
Same Memory	Same Lineage
Logical Multiplicity	Phenomenal Multiplicity
Subject Count	Total Welfare
Ablation effect	Constitutive Membership
High Capacity	Phenomenality
High Entropy	Phenomenality
Strong Connectivity	Phenomenality
Error Correction	Phenomenality

Appendix C. Versioned Positive-Theory Lineage

Version	Positive move	Failure exposed	Surviving lesson
v0.1	counterfactual causal return	reduced to strong connectivity; weak-coupling and simple-loop overinclusion	live generation matters
v0.2	contextual return + uninterrupted expression	simple mechanisms remain; route may be measurement-created	context and measurement control matter
v0.3	repertoire, capacity, interaction, robustness, time	high return can have low repertoire; capacity can come from independent storage	complexity is multidimensional
v0.4	common witness, natural mediator use, intervention-preserving recoding	redundancy can hide real participation from singleton tests	causal properties must co-occur in one regime
v0.5	coalition-sensitive mediation, protected distinctions, redundancy-aware analysis	repair can be added to any finite rule; protection is not consciousness	constitutive organization must be coalition-aware

Appendix D. Recommended Empirical Record

A cross-substrate record should minimally include:

```

system_id
system_version
runtime_regime
causal_scale
temporal_scale
candidate_unit
actual_provenance
intervention_family
baseline_state
transformation
causal_signature
phenomenal_target_type
first_person_anchor_status
report_status
behavior_status
measurement_model
bridge_hypothesis
bridge_version
invariant_portfolio
transport_loss
tokenization_rule
uncertainty
failure_condition
ethical_risk_profile
evidence_dependencies

```

Appendix E. Literature-Review Scope

The literature review is focused rather than systematic. The paper cites direct neighbors needed to delimit its claims, but it does not establish historical priority for every component of GCP. Before journal submission, database-specific searches (at minimum Web of Science/Scopus/Google Scholar or equivalent) should be archived for the principal novelty claims: typed psychophysical gaps, transformation-specific cross-substrate transport, bridge identifiability, and the RRH/RRH-E lineage.

Appendix F. Research and Authorship Disclosure

This manuscript was developed from a large user-directed research archive with substantial assistance from ChatGPT in synthesis, literature retrieval, drafting, formalization, counterexample construction, and code-based checking. The archive numbering records exploratory order, not independent empirical experiments.

The v0.1-v0.5 labels identify exploratory theory stages, not independently published studies. Publication v1.0 denotes this paper’s first public edition; the supplied working draft carried a separate 1.2 label. Appendix G and the accompanying `gcp_checks.py` reconstruct the finite results used here. Appendix H identifies the preceding published transport results and their scope; their original code and proofs remain available with [27]. These computations check constructed mathematical models; they are not empirical consciousness tests.

No human participant study, animal experiment, or deployed artificial-agent consciousness assessment was conducted for this manuscript. No claim of independent peer review or separate final human line-by-line verification is made. Hongju Liu directed and authorized preparation and publication and is the human author of record and accountable depositor. The positive existence bridge remains a partially specified conjecture. This is adjacent first-party research: it does not amend, validate, or independently corroborate the Trinity Accord.

Data and code availability. The publication package contains English Markdown and PDF, the finite-check scripts and results, references and revision notes, citation metadata, and checksums. The exploratory archive is not represented as an independently audited dataset.

Ethics. No new human or animal data were collected. Synthetic checks do not establish welfare status; any future empirical work requires appropriate ethical assessment before implementation.

Appendix G. Explicit Finite Constructions and Proofs

The results in this appendix concern causal models. None supplies a phenomenal ground-truth label. The accompanying standard-library Python script enumerates finite cases and checks the reported numerical values.

G.1 All-cut return and strong connectivity

Let W be an $n \times n$ nonnegative matrix, $n \geq 2$, with an edge j to i exactly when W_{ij} is positive. For each nontrivial bipartition π , W_π deletes all edges between its blocks. Consider the condition that, for every π and every vertex i , there exists a positive integer k such that $(W^k - W_\pi^k)_{ii}$ is positive.

Proposition G1. This condition holds if and only if the positive-edge graph is strongly connected.

Proof. Under strong connectivity, from any i one can reach a vertex in the opposite block and return to i . The resulting closed walk has positive weight and is deleted by the cut. Nonnegativity prevents cancellation, so the corresponding diagonal difference is positive. Conversely, if the graph is not strongly connected, take a source strongly connected component and cut it from its nonempty complement. No closed walk can cross that cut because no edge enters the source component. The diagonal differences are therefore zero for every k . This contradicts the condition. The two directed paths in the forward direction may each be chosen with at most $n-1$ edges, so $k \leq 2(n-1)$ suffices. The proof does not apply to signed weights, where cancellation can occur. A one-vertex graph requires a separate rule.

The script checks all 4,096 loop-free directed graphs on four labeled vertices. This verifies the finite implementation against the proposition; the argument above covers all finite n under the assumptions.

G.2 Recurrence, collapse, and retained capacity

For $F(a, b) = (a \oplus b, a \oplus b)$, a second application always yields $(0, 0)$. Yet replacing the first intermediate coordinate by zero gives the trajectory (a, b) to (p, p) to $(0, p)$ to (p, p) , where $p = a \oplus b$. The reset protocol retains a distinction erased by uninterrupted evolution. The two protocols answer different questions.

For five input bits, let $F(x)$ repeat their parity in all five output coordinates. Since five is odd, a repeated parity state is a fixed point. The image of F^h has exactly two elements for every h at least one. A deterministic channel with unrestricted input preparation has capacity $\log_2$ of its image size: choose one representative preimage per output and use them uniformly to attain that value; output entropy bounds it from above. Thus this mechanism has capacity one bit at each positive horizon, despite complete dependence of every output on every input.

The four-bit map $G(a, b, c, d) = (b, c, d, a \oplus b)$ is invertible: given (u, v, w, z) , its inverse is $(z \oplus u, u, v, w)$. Every iterate is bijective and has capacity four bits. Its dependency graph contains a cycle through all four coordinates and hence is strongly connected. These examples separate graph connectivity from retained differentiation without claiming an ordering of any unspecified return score.

A mode-switching five-bit mechanism can implement identity in mode zero and parity broadcasting in mode one. Identity retains five bits but has only independent self-dependence; broadcasting has complete dependence but retains one bit. There is no mode witnessing both five-bit retained capacity and complete dependency. Allowing the mode itself to vary would define a different system and query, not rescue the invalid within-mode conjunction.

G.3 Recoding and interventions

Consider two independent scalar dynamics $a' = 2a$ and $b' = 3b$. Define $u = a + b$ and $v = a - b$. Then

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = \begin{pmatrix} 5/2 & -1/2 \\ -1/2 & 5/2 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}. \quad (\text{G1})$$

The recoded matrix is dense. Nevertheless, a physical perturbation $(\delta a, 0)$ maps to $(\delta a, \delta a)$, and $(0, \delta b)$ maps to $(\delta b, -\delta b)$. Treating independent interventions on u and v as the same physical

intervention family changes the question. Recoding the dynamics and intervention family together preserves the original decomposition. This example concerns representation dependence; signed recoded coefficients do not meet Proposition G1's nonnegativity condition.

G.4 Redundancy and protected distinctions

For $n = 2r + 1$ noiseless carriers encoding one common bit, majority decoding changes exactly when at least $r + 1$ carriers are replaced by the opposite bit. The output distributions are point masses, so their total-variation distance is zero or one as in equation (32). Under independent bit-flip noise with probability p , correctness requires at most r errors. The binomial formula in Section 14.5 follows. A carrier is pivotal exactly when the remaining $2r$ carriers tie, yielding the second formula there. Correlated noise need not obey either binomial expression.

Proposition G2. Every deterministic finite binary update rule admits an exact repetition-code wrapper on its valid codewords.

Proof. Let F map m -bit states to m -bit states. Let E_n repeat each bit n times and D_n decode each block by odd majority, so $D_n E_n$ is the identity. Define the wrapped rule $\tilde{F} = E_n F D_n$. Then $\tilde{F} E_n = E_n F$. Induction yields $\tilde{F}^h E_n = E_n F^h$ for every nonnegative integer h . Before an update, up to r corruptions per block leave D_n unchanged. This does not guarantee tolerance of noise inside the decoder or update hardware, nor establish consciousness for F or its wrapper. It establishes only that this protection mechanism is not specific to one privileged update rule.

The script exhaustively checks all 256 deterministic two-bit update rules, all four inputs, and all noise patterns correctable by blockwise triple repetition. The general statement follows from the identities above.

G.5 Structural factorization and explanatory richness

If $f = g \circ r$, equal r -values must have equal f -values. Consequently the number of distinct f -values cannot exceed the number of r -values. Conversely, if f is constant on each fiber of r , defining g on the image of r by any representative makes $f = g \circ r$. This is the elementary factorization criterion used in [27]. It constrains a proposed complete structural explanation, not the existence of experience. If a relevant history or regime is omitted from r , the claimed completeness can fail even when the factorization holds on a limited calibration sample.

Appendix H. Relationship to the Earlier Transport Paper

The earlier paper [27] establishes four model-relative transport conditions: non-identification when opposing existence models have identical observable protocol distributions; factorization of a target query through an abstraction exactly when it is constant on abstraction fibers; guarded composition of relations referring to a common intermediate value; and metric error propagation under a justified Lipschitz map. Its eight-state illustration shows that an abstraction can preserve one stipulated query while aliasing another. The proofs, complete model, and executable checks remain in that publication package.

Those results are inherited methodological foundations, not new results of the present paper. Here they support the scope of Sections 7-12 and the richness constraint in Section 15.4. The additional contribution is the common-witness organization, explicit failure sequence, coalition-sensitive realization analysis, and the integration of existence, phenomenal structure, perspective, and continuation into one research architecture.

The new diagnostic checks do not upgrade the earlier conditional transport claims into findings of consciousness. Equally, the broader architecture does not erase the earlier paper’s sharper mathematical conditions. The two papers address connected but distinct research questions and share an author and research lineage; neither is independent corroboration of the other.

References

1. Birch, J., Schnell, A. K., & Clayton, N. S. (2020). Dimensions of Animal Consciousness. *Trends in Cognitive Sciences*, 24(10), 789-801. <https://doi.org/10.1016/j.tics.2020.07.007>
2. Dung, L., & Newen, A. (2023). Profiles of animal consciousness: A species-sensitive, two-tier account to quality and distribution. *Cognition*, 235, 105409. <https://doi.org/10.1016/j.cognition.2023.105409>
3. Klein, C., & Barron, A. B. (2020). How experimental neuroscientists can fix the hard problem of consciousness. *Neuroscience of Consciousness*, 2020(1), niaa009. <https://doi.org/10.1093/nc/niaa009>
4. Kleiner, J., & Ludwig, T. (2024). What is a mathematical structure of conscious experience? *Synthese*, 203, 89. <https://doi.org/10.1007/s11229-024-04503-4>
5. Rubenstein, P. K., Weichwald, S., Bongers, S., Mooij, J. M., Janzing, D., Grosse-Wentrup, M., & Schölkopf, B. (2017). Causal Consistency of Structural Equation Models. *Proceedings of UAI 2017*. arXiv:1707.00819.
6. Albantakis, L., et al. (2023). Integrated information theory (IIT) 4.0: Formulating the properties of phenomenal existence in physical terms. *PLOS Computational Biology*, 19(10), e1011465. <https://doi.org/10.1371/journal.pcbi.1011465>
7. Chang, A. Y. C., Biehl, M., Yu, Y., & Kanai, R. (2020). Information Closure Theory of Consciousness. *Frontiers in Psychology*, 11, 1504. <https://doi.org/10.3389/fpsyg.2020.01504>
8. Chalmers, D. J. (1995). Absent Qualia, Fading Qualia, Dancing Qualia. In T. Metzinger (Ed.), *Conscious Experience*. Imprint Academic.
9. Doerig, A., Schurger, A., Hess, K., & Herzog, M. H. (2019). The unfolding argument: Why IIT and other causal structure theories cannot explain consciousness. *Consciousness and Cognition*, 72, 49-59. <https://doi.org/10.1016/j.concog.2019.04.002>
10. Casali, A. G., et al. (2013). A theoretically based index of consciousness independent of sensory processing and behavior. *Science Translational Medicine*, 5(198), 198ra105. <https://doi.org/10.1126/scitranslmed.3006294>
11. Schartner, M., Seth, A., Noirhomme, Q., Boly, M., Bruno, M.-A., Laureys, S., & Barrett, A. (2015). Complexity of multi-dimensional spontaneous EEG decreases during propofol induced general anaesthesia. *PLOS ONE*, 10(8), e0133532. <https://doi.org/10.1371/journal.pone.0133532>
12. Seth, A. K. (2026; published online 2025). Conscious artificial intelligence and biological naturalism. *Behavioral and Brain Sciences*, 49, e315. <https://doi.org/10.1017/S0140525X25000032>
13. Wang, P., Law, E. L.-C., Zou, L., Jiang, Y., Hertwig, R., Paas, F., & Laureys, S. (2026). A structured framework for human-AI resemblance: Evidence, inference and perception.

- Physics of Life Reviews*, 59, 76-93. <https://doi.org/10.1016/j.plrev.2026.09.003>
14. Tononi, G., Sporns, O., & Edelman, G. M. (1994). A measure for brain complexity: Relating functional segregation and integration in the nervous system. *Proceedings of the National Academy of Sciences*, 91(11), 5033-5037. <https://doi.org/10.1073/pnas.91.11.5033>
 15. Laukkonen, R., Friston, K., & Chandaria, S. (2025). A beautiful loop: An active inference theory of consciousness. *Neuroscience & Biobehavioral Reviews*, 176, 106296. <https://doi.org/10.1016/j.neubiorev.2025.106296>
 16. Cover, T. M., & Thomas, J. A. (2006). *Elements of Information Theory* (2nd ed.). Wiley.
 17. Cogitate Consortium 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>
 18. Dung, L. (2025; published online 2023). Tests of Animal Consciousness are Tests of Machine Consciousness. *Erkenntnis*, 90, 1323-1342. <https://doi.org/10.1007/s10670-023-00753-9>
 19. Dung, L., & Kersten, L. (2025; first published online 2024). Implementing artificial consciousness. *Mind & Language*, 40(3), 285-305. <https://doi.org/10.1111/mila.12532>
 20. Bayne, T. (2010). *The Unity of Consciousness*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199215386.001.0001>
 21. Gottlieb, J., & Fischer, B. (2026; published online 2024). Counting subjects. *Inquiry*, 69(4), 2450-2477. <https://doi.org/10.1080/0020174X.2024.2370016>
 22. Roelofs, L., & Sebo, J. (2024). Overlapping minds and the hedonic calculus. *Philosophical Studies*, 181, 1487-1506. <https://doi.org/10.1007/s11098-024-02167-x>
 23. Schwitzgebel, E., & Nelson, S. R. (2026; published online 2025). When counting conscious subjects, the result needn't always be a determinate whole number. *Philosophical Psychology*, 39(3), 847-867. <https://doi.org/10.1080/09515089.2025.2520364>
 24. Keeling, G., & Street, W. (2026). *Emerging Questions in AI Welfare*. Cambridge University Press. <https://doi.org/10.1017/9781009732000>
 25. Edelman, G. M. (2003). Naturalizing consciousness: A theoretical framework. *Proceedings of the National Academy of Sciences*, 100(9), 5520-5524. <https://doi.org/10.1073/pnas.0931349100>
 26. Lamme, V. A. F. (2006). Towards a true neural stance on consciousness. *Trends in Cognitive Sciences*, 10(11), 494-501. <https://doi.org/10.1016/j.tics.2006.09.001>
 27. Liu, H. (2026). *Cross-Substrate Phenomenal Comparison: A Typed Transformation-Transport Framework under Existential Uncertainty* (TA-TR-2026-15, v1.0). Zenodo. <https://doi.org/10.5281/zenodo.22934654>