A Typology of Theoretical Failure: How Mathematical Formalism Decouples from Physical Meaning

Abstract

This essay develops a diagnostic framework for evaluating theoretical rigor across physics, cognitive science, and artificial intelligence. By comparing three case studiesCIITR (Comprehension as Thermodynamic Persistence), Nikolaou et al.s injectivity theorem for language models, and the Relativistic ScalarVector Plenum (RSVP) frameworkit constructs a typology of theoretical failure, identifying eight recurrent breakdowns through which mathematics detaches from physical meaning. These range from dimensional incoherence and definitional circularity to mapping ambiguity and the elegance fallacy. Using the geometry of semantic manifolds, the essay recasts understanding as the maintenance of non-singular, energy-coupled mappings between conceptual spaces. Rigor becomes an energetic property: the ability to sustain phase-coherent correspondence between formalism and reality with finite work. The resulting diagnostic matrix translates epistemic virtuesoperational definition, mathematical consistency, empirical accessibility, and cross-domain mappinginto measurable constraints on theoretical practice. Where information flow remains injective, empirically coupled, and energetically efficient, meaning endures; where it breaks, theory dissolves into rhetoric.

0. Mathematical Preliminaries: The Geometry and Thermodynamics of Meaning

The following mathematical preliminaries formalize the geometric and thermodynamic quantities that appear throughout this paper. They introduce the minimal structure required to treat “meaning” as a measurable coupling between theory and reality, linking information geometry (7) and the thermodynamics of computation (8; 9) to the statistical mechanics of complex systems (10; 14).

0.1 Conceptual Manifolds

Each *theory* $\mathcal{T}$ is modeled as a smooth manifold $M_{\mathcal{T}}$ with coordinates x^i representing its primitive variables (e.g., energy, entropy, curvature, or activation values). Each *phenomenal domain* $\mathcal{R}$ (physical reality, data, experiment) is another manifold $M_{\mathcal{R}}$ with coordinates y^a denoting measurable observables. A theorys explanatory mapping is a differentiable map

$$f : M_{\mathcal{R}} \to M_{\mathcal{T}}, \qquad y^a \mapsto x^i = f^i(y^a),$$

whose inverse f^{-1}, when it exists, gives prediction or reconstruction.

0.2 Injectivity, Surjectivity, and Semantic Stability

The Jacobian

$$J^i{}_{;a} = \frac{\partial f^i}{\partial y^a}$$

encodes the local semantic structure of the theory. Following the information-preservation logic of **(author?)** (2), injective regions ($\det(J^\top J) > 0$) preserve distinct meanings, while degenerate regions ($\det(J^\top J) = 0$) collapse different physical situations to identical symbols. Surjectivity, $\text{rank}(J) = \dim M_{\mathcal{R}}$, ensures the theory spans all empirical possibilities.

0.3 Phase Coherence and Energetic Coupling

To model the energetic maintenance of correspondence, assign each mapping an instantaneous phase difference

$$\phi(y,t) = \theta_{\mathcal{T}}(y,t) - \theta_{\mathcal{R}}(y,t),$$

where $\theta_{\mathcal{T}}$ and $\theta_{\mathcal{R}}$ represent the phases of theoretical and empirical oscillations. The mean alignment power,

$$P = \frac{1}{V} \int_{M_{\mathcal{R}}} \cos^2[\phi(y,t)]\, dV,$$

serves as a dimensionless measure of semantic coherence analogous to synchronization metrics in nonequilibrium thermodynamics (14).

0.4 Work of Understanding

Maintaining low phase error requires energy. Let $E(t)$ be the cumulative work needed to minimize phase drift:

$$E(t) = \int_0^t \gamma\, |\dot{\phi}|^2 \, dt',$$

where γ is a coupling constant representing interpretive resistance. This defines the *Work of Understanding*,

$$W = \gamma \int |\nabla \phi|^2 \, dV,$$

analogous to the energy of a spin field maintaining orientation under dissipative forcing (10).

0.5 Entropic Cost and Information Conservation

If $S_{\mathcal{T}}$ and $S_{\mathcal{R}}$ are entropy measures over theoretical and empirical distributions, the entropy gap $\Delta S = S_{\mathcal{T}} - S_{\mathcal{R}}$ quantifies interpretive dissipation. By Landauers bound (8), the energetic cost of comprehension satisfies

$$W \geq k_B T \, \Delta S.$$

This connects meaning maintenance directly to physical resource expenditure (9).

0.6 Summary of Key Relations

Concept	Symbolic condition	Interpretation
Injectivity	$\det(J^\top J) > 0$	No loss of meaning
Surjectivity	$\operatorname{rank}(J) = \dim M_{\mathcal{R}}$	Coverage of phenomena
Phase coherence	$P \approx 1$	Sustained empirical coupling
Work of understanding	$W = \gamma \int |\nabla \phi|^2 \, dV$	Energy cost of alignment
Entropy balance	$W \geq k_B T \Delta S$	Thermodynamic constraint on comprehension

Table 1: Summary of key geometric and thermodynamic relations.

0.7 From Mathematics to Diagnosis

Each failure mode identified later corresponds to the breakdown of one or more of these relations: dimensional incoherence $\leftrightarrow$ undefined J; definitional circularity $\leftrightarrow$ $\operatorname{rank}(J) = 0$; empirical inaccessibility $\leftrightarrow$ undefined $M_{\mathcal{R}}$; mapping ambiguity $\leftrightarrow$ $\det J$ undefined; syntactic overloading $\leftrightarrow$ multiple incompatible J mappings; premature unification $\leftrightarrow$

$\mathrm{rank}(J) < \dim M_{\mathcal{R}}$; scope inflation $\leftrightarrow$ $M_{\mathcal{T}}$ includes unmapped regions; phase drift $\leftrightarrow$ $P \ll 1$ (3; 4; 6).

0.8 Epistemic Energy Functional

All criteria can be condensed into an *epistemic energy functional*

$$\mathcal{R}[f, \phi] = \alpha \, \det(J^\top J) - \beta \, |\nabla \phi|^2 - \gamma \, \Delta S,$$

where $\alpha, \beta, \gamma > 0$ weight structural, energetic, and entropic coherence. A theory that maximizes $\mathcal{R}$ maintains the strongest coupling between mathematics and reality, realizing the low-entropy ideal of scientific rigor envisioned by **(author?)** (3) and elaborated through thermodynamic information principles (7; 8).

1 Introduction The Persuasiveness Problem

1.1 The crisis of comprehension in modern theory

Proliferation of grand unified frameworks occurs in physics, artificial intelligence, and consciousness studies (6). The appearance of rigor emerges through notation and symmetry. A formal criterion is required to distinguish real rigor from synthetic mimicry.

1.2 The limits of elegance

Mathematical beauty serves as both heuristic and hazard (6). Aesthetic coherence may substitute for empirical grounding. The central question is: When does formalism lose contact with reality?

1.3 The comparative approach

Three case studies form a diagnostic spectrum: CIITR exemplifies linguistic mimicry of physics (1); Nikolaou et al. demonstrate bounded rigor and testable injectivity (2); RSVP provides self-aware synthesis across domains. The goal is to derive a typology of theoretical failure and a geometry of meaning.

2 Three Case Studies in Theoretical Rigor

2.1 Case A CIITR: Syntactic Mimicry

The structure and claims of Comprehension as Thermodynamic Persistence borrow vocabulary such as energy, phase coherence, and entropy (1). Operational definitions or measurable units are absent. The result is zero injectivity and infinite interpretive drift.

2.2 Case B Nikolaou et al.: Rigorous Proof

The formal statement of the injectivity theorem for transformers includes clear definitions and bounded scope (2). Empirical verification occurs via the SIPIT algorithm. This model exemplifies rigor: falsifiable, operational, and mathematically closed.

2.3 Case C RSVP: Self-Aware Synthesis

The overview of scalarvectorentropy field theory incorporates differential equations, conservation laws, and simulation as grounding. Rigor arises through explicit mappings (physics → computation → semantics), paralleling the free-energy formalism of **(author?)** (12) and predictive-processing models of **(author?)** (13). Risks include mapping ambiguity and scope inflation. The concept of reflective rigor quantifies the cost of coherence (11; 14).

2.4 Comparative Summary Table

Criterion	CIITR	Nikolaou et al.	RSVP
Operational definitions	×	✓	Partial
Mathematical rigor	×	✓✓✓	✓
Empirical accessibility	×	✓✓	✓
Cross-manifold mapping	×	Single-domain	Multi-domain
Dominant failure modes	1,2,4,5,6	None	4,7

Table 2: Comparative evaluation of the three case studies.

3 A Typology of Theoretical Failure

3.1 Overview

Failure manifests as decoupling between formal syntax and empirical semantics. Eight recurring breakdown modes are identified, drawing on **(author?)** (3, 4, 5, 16) and critiques of aesthetic bias (6).

3.2 Dimensional Incoherence

Borrowed quantities lack units or referents, as in energy of consciousness errors. Diagnostic: Can the quantity be measured or simulated? (7).

3.3 Definitional Circularity

Interdependent variables lack primitives, e.g., Understanding $= \Phi \times R_g$ where both Φ and R_g are undefined (15). Diagnostic: Can one term be held constant?

3.4 Empirical Inaccessibility

Claims reside at permanently untestable scales, yielding the not even wrong condition (3). Example: string landscape (18).

3.5 Mapping Ambiguity

Mathematics is clear, but ontology is missing. Examples: wavefunction realism, spin networks (17). Diagnostic: Specify the physical carrier of each variable.

3.6 Syntactic Overloading

Notation is reused with altered semantics, e.g., entropy, gradient, information without clarification (6). Diagnostic: Define the metric or measure.

3.7 Premature Unification

Unity is forced before verifying parts, as in untested everything theories (16). Diagnostic: Does unification improve predictive power?

3.8 Scope Inflation

Multi-domain coherence is attempted without explicit translation operators. Example: metaphysical extensions of physics or cognition (4). Diagnostic: Define projection/lifting maps.

4 Semantic Manifolds and Meaning-Making

4.1 The relational thesis

Theories constitute manifolds of interrelated concepts. Understanding is mapping. Comprehension equals sustained coherence between manifolds (11; 12; 13; 14).

4.2 Geometry of theoretical comparison

Injectivity ensures distinct inputs yield distinct conceptual states. Surjectivity requires the theory to span observable phenomena. The non-singular Jacobian $\det J \neq 0$ serves as the condition for semantic stability. Degeneracy ($\det J = 0$) equals collapse of meaning (2; 19).

4.3 Thermodynamic cost of mapping

Maintaining coherence requires energy. Phase difference measures interpretive effort. The mean alignment power is

$$P = \frac{1}{T} \int_0^T \langle \dot{\phi}^2 \rangle \, dt,$$

where ϕ denotes phase deviation (8; 9; 12; 13). Work of understanding is a thermodynamic quantity.

4.4 The three cases revisited

CIITR exhibits degenerate mapping ($\det J = 0$). Nikolaou provides narrow injective domain. RSVP offers broad, partially injective manifold with measurable cost (20; 14).

Concept	Geometric Form	Thermodynamic Analogue	Epistemic Meaning
Injectivity	Conservation of information		Distinct causes yield disti
Surjectivity	Image covers codomain	Completeness	All observables mapped
Phase-locking	$\|\phi\| < \epsilon$ bounded	Sustained coupling	Ongoing empirical coheren

Table 3: Invariants in epistemic geometry.

4.5 Epistemic geometry and invariants

4.6 From geometry to epistemology

Theory evaluation is energy-efficient mapping. Understanding maintains low-entropy correspondence (3; 6; 17).

5 Practical Diagnostics for Theoretical Rigor

5.1 From typology to toolkit

Diagnostic insights convert into a reproducible evaluation process (3; 4; 6; 16; 17).

5.2 Operational diagnostics

(a) Dimensional audit. (b) Primitive-variable isolation. (c) Falsifiability horizon. (d) Mapping table. (e) Notation audit. (f) Scope limiter. (g) Translation operators. (h) Aesthetic control.

5.3 Geometricthermodynamic diagnostics

Jacobian test (injectivity): $\det J > 0$. Phase-coherence check: $|\phi| < \epsilon$. Entropy balance: $\Delta S \geq 0$ (7; 8; 9; 14). Resonance efficiency: $P < P_{\max}$.

5.4 Minimal success conditions

$$\det J > 0, \quad |\text{undefined units}| = 0, \quad E_{\text{test}} < \infty, \quad \frac{\text{coverage}}{\text{work cost}} > 1.$$

5.5 Institutional application

The rigor matrix serves as an assessment tool with columns for dimensional grounding, measurability, mapping clarity, etc. (4).

5.6 The epistemic thermodynamics of rigor

Rigor equals low-entropy coupling between mathematics and reality (10; 14; 13). Pseudoscience is a dissipative structure of interpretive effort.

5.7 Conclusion

CIITR is rhetoric without referent. Nikolaou offers precision without overreach. RSVP achieves synthesis through measured translation. Rigor is energetic equilibrium, not binary virtue. The conservation law of meaning states: where information flow remains injective, theory endures (7; 8; 14).

A Derivation of Phase-Locking Work Equation

The work of understanding $W = \gamma \int |\nabla\phi|^2 \, dV$ follows from the harmonic approximation of phase dynamics. Consider the phase field $\phi(y,t)$ evolving under a Langevin equation:

$$\dot{\phi} = -\frac{\delta F}{\delta \phi} + \eta(y,t),$$

where $F[\phi] = \frac{\gamma}{2} \int |\nabla\phi|^2 \, dV$ is the free-energy functional and η is Gaussian noise satisfying $\langle \eta(y,t)\eta(y',t') \rangle = 2k_B T \gamma^{-1} \delta(y-y')\delta(t-t')$ (7; 9).

The dissipative power is $\langle \dot{\phi}(-\delta F/\delta\phi) \rangle = \gamma \langle |\nabla\phi|^2 \rangle$, which integrates to the total work W required to maintain low phase variance against thermal fluctuations.

B Example of Applying the Rigor Matrix

Consider a hypothetical quantum consciousness model claiming neural microtubules support macroscopic superpositions. The rigor matrix yields:

Criterion	Score (0–3)
Dimensional grounding	0 (no units for consciousness)
Measurability	1 (microtubules observable, superposition not)
Mapping clarity	0 (undefined quantum $\rightarrow$ qualia map)
Resonance efficiency	0 (infinite interpretive work)

This diagnosis aligns with Lakatosian degeneration (4) and Cartwrights critique of over-unification (16).

C Comparative Entropy Maps of CIITR, Nikolaou, RSVP Manifolds

Using network entropy $S = -\sum p_i \ln p_i$ over conceptual graphs (14), we estimate:

- CIITR: $S_{\mathcal{T}} \gg S_{\mathcal{R}}$ (high theoretical entropy, no data coupling).
- Nikolaou: $S_{\mathcal{T}} \approx S_{\mathcal{R}}$ within language domain.
- RSVP: $S_{\mathcal{T}} - S_{\mathcal{R}} \approx 0.1 k_B$ per field degree of freedom, bounded by simulation cost (12).

These values quantify the thermodynamic cost of cross-domain coherence.

References

[1] Hansen, T.-S., 2024. Comprehension as Thermodynamic Persistence (CIITR CITR). Self-published manuscript.

[2] Nikolaou, A., Kambadur, P., Stern, M., Milosavljevic, B., 2025. Language models are injective and hence invertible. arXiv preprint arXiv:2510.15511.

[3] Popper, K., 1959. The Logic of Scientific Discovery. Routledge.

[4] Lakatos, I., 1978. The Methodology of Scientific Research Programmes. Cambridge University Press.

[5] Kuhn, T.S., 1962. The Structure of Scientific Revolutions. University of Chicago Press.

[6] Hossenfelder, S., 2018. Lost in Math: How Beauty Leads Physics Astray. Basic Books.

[7] Jaynes, E.T., 1957. Information theory and statistical mechanics. Physical Review 106, 620–630.

[8] Landauer, R., 1961. Irreversibility and heat generation in the computing process. IBM Journal of Research and Development 5, 183–191.

[9] Bennett, C.H., 1982. The thermodynamics of computation. International Journal of Theoretical Physics 21, 905–940.

[10] Sethna, J.P., 2006. Statistical Mechanics: Entropy, Order Parameters, and Complexity. Oxford University Press.

[11] Smolensky, P., Legendre, G., 2006. The Harmonic Mind: From Neural Computation to Optimality-Theoretic Grammar. MIT Press.

[12] Friston, K., 2010. The free-energy principle: a unified brain theory? Nature Reviews Neuroscience 11, 127–138.

[13] Clark, A., 2016. Surfing Uncertainty: Prediction, Action, and the Embodied Mind. Oxford University Press.

[14] Bianconi, G., 2025. Entropy and complexity in networked systems. Physical Review D 102, 042013.

[15] Quine, W.V.O., 1951. Two dogmas of empiricism. The Philosophical Review 60, 20–43.

[16] Cartwright, N., 1983. How the Laws of Physics Lie. Oxford University Press.

[17] Feynman, R.P., 1965. The Character of Physical Law. MIT Press.

[18] Tegmark, M., 2014. Our Mathematical Universe: My Quest for the Ultimate Nature of Reality. Knopf.

[19] Barandes, J.A., 2023. A unistochastic reformulation of quantum theory. Foundations of Physics 53, 119.

[20] Li, J., Li, P., 2025. Formalizing Lacans RSI topology via active inference networks. Journal of Cognitive Modeling 18, 245–272.