

Evolutionary Relativity

Frame-Qualified Organizational Change and Observer-Relative Reconstruction

Kai Wang^{1*}

¹ Independent Researcher, Australia

* Correspondence: wangkai@connection-theory.org

Version 3.6.1 | July 2026

Abstract

In this paper, evolutionary refers to organizational change along the temporally extended history of an adaptive system, not to Darwinian change across generations. Evolutionary Relativity studies a relation: an observed system traverses an organizational path, while another adaptive organization reconstructs that path through finite evidence, reference states, temporal coverage, resolution, and comparison capacity. The central distinction is between path and endpoint: matched start and end states can conceal different amounts of established intermediate change. The paper separates a protected registered measurement frame from the frame of the studied observer and reports a finite registered-refinement path estimate rather than claiming an infinite-resolution curve length. The canonical target is an independently motivated continual learner operating in a partially observable, non-stationary task stream. Its loss, update rule, and environment are fixed before path and rival summaries are selected. Candidate histories arise from real adaptive runs under pre-registered curricula, drift schedules, data orders, replay budgets, and random seeds; the trajectory bank is screened from that pool rather than synthesized to prescribed summary values. A history-dependence gate requires a history-aware predictor to outperform a current-only predictor, and a non-designed-sufficiency gate prohibits embedding path or rival summaries in the target dynamics. The studied observer predicts the target's next response given the next external input, receives no path supervision, and is frozen before probing. Raw-channel calibration tests both extractability and probe fairness; frozen-representation results are normalized against each summary's raw-channel recoverability. The target characteristic timescale $\tau(P^M)$ is estimated from amplitude-weighted temporal correlation and must be empirically separable from persistence, transition count, recency, and path. ER's empirical wagers are whether path survives in the frozen observer representation better than identifiable rivals, whether defined degradations produce the registered residual-proxy shift, and whether the readable-scale optimum moves jointly with observer capacity and the independent target timescale. This is a pre-validation theoretical status, not empirical support.

1. The Explanatory Problem

1.1 Equal Duration, Different Organizational Paths

Two adaptive systems can occupy the same calendar interval while undergoing very different organizational histories. One may repeatedly reorganize how it classifies situations, selects actions, or coordinates internal processes. Another may remain within a stable routine. Calendar duration alone does not distinguish these histories.

Endpoint comparison is also insufficient. A system may depart from an earlier organization, remain in intermediate organizations long enough for those states to be established, and later return close to its starting point. The start and endpoint can be similar even though the traversed history is not.

1.2 The Relational Object

Evolutionary Relativity does not begin with a fully accessible objective time and then add observer error. Its object is the relation between an observed organizational trajectory and the adaptive organization through which that trajectory becomes detectable and comparable. The target constrains what occurred; the observer constrains what can be resolved, retained, and reconstructed.

A confirmatory test must therefore constrain every place in which the answer could otherwise be designed: the target dynamics, the trajectory bank, the evidence channel, the studied observer's objective, and the probe family. This complexity is not a claim that every application requires the same apparatus. It is the cost of making the first canonical test capable of losing.

1.3 Main Contributions and Deliberate Scope

The paper contributes a two-frame architecture, a finite-band path estimate with explicit entry conditions, an independently motivated continual target, history-dependence and non-designed-sufficiency gates, a screened and decorrelated trajectory bank, raw-channel probe-fairness calibration, normalized retention analysis, degradation-specific proxy predictions, and a joint target-timescale/observer-capacity prediction.

Retrospective felt duration, childhood time compression, relationship divergence, personal identity, inactivity and recovery, evolutionary velocity, forgetting, and recovery cost remain deferred. They introduce additional mechanisms and should not be used to rescue or decorate the present core.

The main axis is a defensible organizational path on the observed side and a finite, non-path-supervised reconstruction process on the observer side.

2. Ownership and Prior Work

2.1 Imported Foundation

Connection Theory owns the foundational concepts used here: temporally extended existence, interaction, feedback, feedback availability, connection, organization, capability-relevant persistence state, Organizational Continuity, and current applicability (Wang, 2026a). Evolutionary Relativity imports them without redefinition.

The Flow Shaping Principle is a separate companion proposal concerning correspondence-calibrated pathway-specific active updating (Wang, 2026c). Its success or failure does not determine whether an organizational path can be estimated or reconstructed here.

2.2 Concepts and Protocol Boundaries

Concepts owned here are the registered measurement frame, observer frame, protected finite-band path reference, observer-relative reconstruction, degradation-specific proxy shift, and joint capacity-timescale matching. These define ER's theoretical relation.

The history-dependence gate, non-designed-sufficiency gate, adequacy decoder, constructibility gate, trajectory bank, challenger metric, collinearity gate, τ -independence gate, probe-fairness calibration, and frozen-probe design are test protocols defined for this paper. These protocols function as methodological components rather than as additional theoretical primitives of ER.

2.3 Related Work

Information geometry supplies Riemannian metrics, geodesic concepts, and parameterization-aware learning directions for statistical models (Rao, 1945; Amari, 1998). Those tools can support path measurements in parameter or distribution space, but ER is not a renaming of a Fisher–Rao or natural-gradient trajectory. Its target state can be any independently justified organizational representation; only persistence-qualified states enter the registered path; and path must defeat endpoint, count, excursion, recency, and target-timescale alternatives in held-out discrimination.

Rate-distortion theory and the information bottleneck formalize trade-offs between compression and retained fidelity or relevance (Shannon, 1959; Tishby et al., 2000). Predictive state representations encode histories through predictions of future observations rather than direct access to hidden state (Littman et al., 2001). Frozen probes test what information is recoverable from internal representations without updating the represented model (Alain & Bengio, 2016). Continual-learning research supplies independently motivated adaptive targets under task sequence and distribution change (Parisi et al., 2019; De Lange et al., 2021). ER combines these neighboring resources in a dual-frame test that separates raw-channel extractability, non-path-supervised observer induction, and summary-specific retention.

3. Two Frames and a Protected Reference

3.1 Registered Measurement Frame and Observer Frame

$$F^M = (B, R_M, P_M, W_M, d, h) \quad F^O = (A_O, R_O, P_O, W_O, \rho_O, C_O)$$

The registered measurement frame F^M specifies the target boundary B , measurement resolution R^M , target property set P^M , observation and sampling design W^M , organizational distance d , and persistence rule h . It is fixed before primary outcomes are inspected.

The observer frame F^O specifies evidence access A^O , observer resolution R^O , accessible property set P^O , temporal coverage W^O , reference fidelity ρ^O , and comparison or encoding capacity C^O . These belong to the studied observer and may be manipulated.

Feature	Registered measurement frame F^M	Studied observer frame F^O
Role	Produces the protected target reference	Produces a frozen internal representation
Status	Pre-registered, public, fixed for confirmation	Trained on a non-path predictive task
Access	Protected full admitted log	Only the assigned evidence channel
Main readout	$\hat{T}(F^M; \mathcal{R})$	Ordinal path-probe score $\hat{T}(F^O)$
Scale relation	Finite-band path estimate	No assumed cardinal equality with $\hat{T}$

Table 1. The protected measurement frame and the studied observer frame have different operational roles and no assumed common cardinal scale.

3.2 Protected Reference Principle

The protected path reference is not an observer-free truth. Its privileged status is procedural: P^M is registered before the primary result, the reconstruction has protected access to the full admitted log, the algorithm is fixed, and independent investigators can reproduce it from the same record.

Objectivity here means constraint and reproducibility within a disclosed observation relation. The protected procedure is privileged because its access and rules are more complete and auditable, not because it occupies a view outside observation.

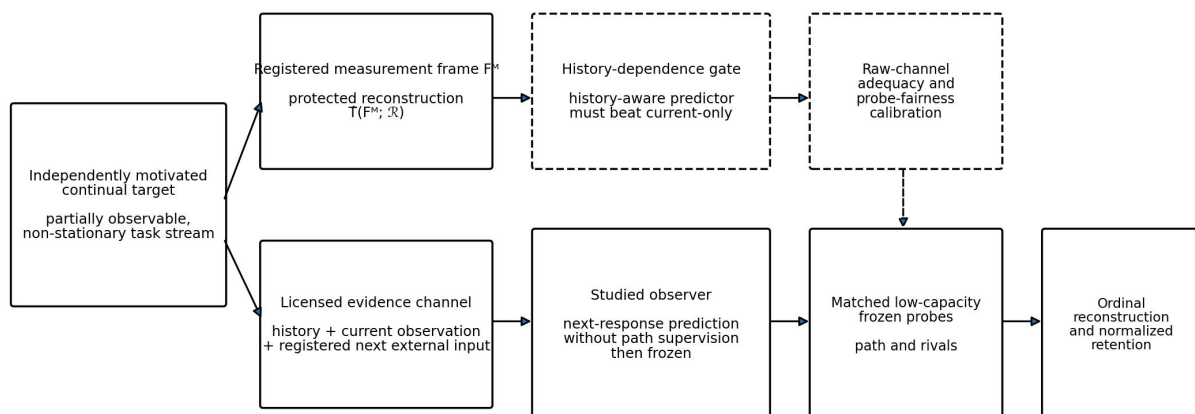
3.3 Target Characteristic Timescale

Let $\tau(P^M)$ denote the pre-registered characteristic timescale of organization-relevant change in P^M . The primary estimator is an amplitude-weighted integrated correlation time computed from the temporal correlation of frame-qualified transition magnitudes. This estimator uses both spacing and magnitude structure rather than reducing τ to a count-based mean interval.

An unweighted organization-distance autocorrelation decay is registered as a secondary challenger. If the two estimators reverse a load-bearing Prediction 3 result, the conclusion is timescale-dependent. The persistence threshold h decides whether a state enters the trajectory; $\tau(P^M)$ describes the temporal organization of admitted changes. They are related but not identical.

Before Prediction 3 is licensed, τ must pass an independence gate against h , transition count, recency-weighted change, and path. If a registered dependence threshold is exceeded, Prediction 3 is undefined for that bank rather than failed.

Independent target dynamics, channel extractability, and observer reconstruction



Target dynamics are fixed before rival summaries are chosen; trajectories are generated by real adaptation and screened, not synthesized to prescribed summary values.

Figure 1. Independent target dynamics and observer reconstruction. The continual target is fixed independently of path and rival summaries. A history-dependence gate establishes that the licensed channel contains target-relevant history. Raw-channel calibration tests extractability and probe fairness; the studied observer is trained on next-response prediction, frozen, and then examined with matched probes.

4. Measurement Entry Conditions

4.1 Independently Licensed Persistence

The persistence rule h helps determine which organizational states enter the target trajectory. It may not be selected to make the path smooth, permit a desired bank, or preserve a favorable result. Its lower bound and admissible range must be justified from P^M and independent domain knowledge before candidate-bank inspection.

Persistence is local to the occurrence of a change. A state counts if it remained organizationally established for the registered interval when it occurred; it need not survive until the end of W^M . A later reversal therefore does not erase an earlier established reorganization.

4.2 Registered Refinement Rather Than Claimed Infinite Resolution

Curve length can increase with refinement, as in the coastline problem (Mandelbrot, 1967). Metric-path length and total variation are mathematically well-defined only under additional conditions (Burago et al., 2001; Ambrosio et al., 2008). A finite experiment cannot prove an infinite-refinement limit. For nested registered partitions, it computes

$$L_j = \sum_{k=0}^{n_j-1} d(O(t_{j,k}), O(t_{j,k+1})) \quad \hat{T}(F^M; \mathcal{R}) = L_j *$$

The estimate is accepted only if at least three nested partitions enter a pre-registered equivalence region for final increments or ratios without a load-bearing ordering reversal. Passing supports $\hat{T}(F^M; \mathcal{R})$ only within the registered band $\mathcal{R}$.

4.3 Feasible h-Window

A coarse h can make finite-band stability trivial while erasing intermediate states; a fine h can admit irregular variation that prevents stability. A study must report the independently admissible h -range, the subset that passes refinement, and whether any passing value preserves a non-trivial excess over endpoint distance.

If no independently licensed h yields both finite-band stability and a non-trivial path distinction, ER is not applicable to that property in that setting. The prevalence and width of feasible h -windows are part of the result.

4.4 Adversarial Ordinal Robustness

The primary metric family must include a challenger that captures a meaningfully different concept of organizational change and is selected because it could reverse the target ordering. The registration must identify a trajectory class in which the primary and challenger metrics are expected to disagree.

If the target ordering reverses within the admissible metric family or h -window, the comparison is reported as metric-dependent. A favorable metric may not be selected after inspection.

5. Path, Endpoint, and Rival Summaries

5.1 Ideal Path Property and Empirical Estimate

When a frame-qualified trajectory is rectifiable, its ideal path property is the total variation under d . Empirical work reports the protected finite-band estimate $\hat{T}(F^M; \mathcal{R})$, not the supremum over every possible partition.

5.2 Why the Path Form Must Be Discriminated

High path estimates can covary with endpoint distance, established-transition count, maximum single excursion, recency-weighted change, and $\tau(P^M)$. A representation can appear path-sensitive while retaining one of these cheaper summaries. Using $\hat{T}$ as a score does not test the path form.

The decisive test is whether normalized path retention in a frozen observer representation exceeds the normalized retention of every rival that first passes target-bank identifiability and raw-channel probe-fairness gates. No claim is licensed against a rival that is inseparable in the bank or unreadable by the registered probe family from the raw channel.

5.3 What Path Length Does Not Mean

Path length is value-neutral. It does not imply improvement, maturity, fitness, identity continuity, current applicability, forgetting, recovery cost, or generalization.

6. Target Dynamics and Observer Reconstruction

6.1 Independently Motivated Continual Target

The canonical target is a real online learner operating in a partially observable, non-stationary task stream. Its task loss, architecture, update rule, replay policy, and environment transition family are justified by the learning task rather than by ER's summaries.

Candidate histories are generated by pre-registered manipulations such as curriculum order, task-switch timing, distribution drift, data order, learning-rate schedule, replay budget, and random seed. Path, count, excursion, recency, and τ are computed after each run by the protected procedure. The trajectory bank is selected from the resulting candidate pool by a fixed algorithm; internal state sequences are not directly synthesized to prescribed summary values.

6.2 History-Dependence and Non-Designed Sufficiency

The target is eligible only if its licensed observations are meaningfully history-dependent. A history-aware predictor and a current-only predictor receive the same current observation, the same registered next external input,

and matched computational resources; only the former receives the licensed history. The history-aware model must exceed the current-only model by a pre-registered held-out margin.

The target loss, update rule, scheduler, and environment must be fixed before path and rival summaries are selected for the confirmatory comparison. Neither path nor any rival may be embedded as a designed sufficient statistic of the target dynamics. Failure of the history test makes Prediction 1 undefined. Violation of non-designed sufficiency makes the study exploratory.

6.3 Non-Path-Supervised Observer Induction

The studied observer receives the licensed target history, current visible state, and the registered next external input, and predicts the target's next response or next admitted visible state. Its loss contains no $\hat{T}$ label, $\hat{T}$ ordering, rival-summary label, or monotone transform of those quantities.

After training, the observer is frozen at a pre-registered representation. No adequacy decoder, path probe, rival probe, or final test may update it. The experiment asks which trajectory structure a generic predictive objective caused the observer to retain.

6.4 Raw-Channel Adequacy and Probe Fairness

Calling a channel path-adequate cannot depend on whether the studied observer succeeds. A capacity-matched but non-identical decoder first tests path extractability from the licensed raw channel on independent calibration trajectories.

The same registered probe family is then calibrated on the raw channel for path and every rival. Each summary must exceed a pre-registered raw-channel recoverability floor. A rival that the probe family cannot recover from raw evidence is not an eligible discrimination target. The primary probe family should have a fixed low parameter budget but sufficient nonlinearity to represent both additive and extremal summaries; a linear-probe analysis may be reported secondarily.

6.5 Frozen-Representation Readout and Normalized Retention

Matched probes are trained on a disjoint probe-calibration bank and evaluated on the final held-out bank. $\hat{T}(F^0)$ denotes the ordinal path score produced from the frozen representation. It is not assumed to equal $\hat{T}(F^M; \mathcal{R})$, share its units, or provide cardinal calibration.

Probe expressivity is controlled by normalizing frozen-representation recoverability against the same summary's raw-channel recoverability:

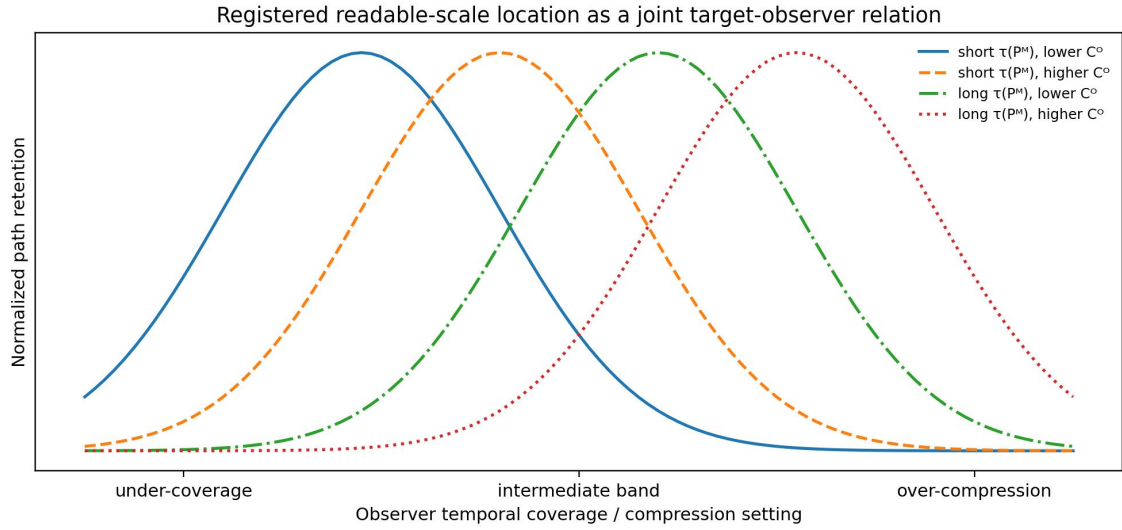
$$\mathcal{R}_q = \frac{\text{Rec}_{\text{frozen}}(q)}{\text{Rec}_{\text{raw}}(q)}$$

The primary comparison concerns normalized retention and held-out ordering. If path is easier than a rival even on the raw channel, the frozen-representation difference is not interpreted without this normalization.

6.6 Structured Proxy Use and Readable Scale

For each degradation class, the raw channel must be below the path-adequacy threshold while one named residual proxy remains above its fairness-calibrated threshold. ER predicts a registered shift in normalized retention away from path and toward that proxy.

Under a fixed resource budget and registered encoder class, temporal coverage and retained resolution generally trade off. The empirical wager is not that a trade-off exists, but that the location of maximal normalized path retention moves jointly with observer capacity C^0 and the independently measured $\tau(P^M)$.



The wager concerns the registered location and displacement of the optimum with both $\tau(P^M)$ and C^O ; τ must first pass its independence gate.

Figure 3. Joint capacity-timescale prediction. The empirical target is the registered location and displacement of maximal normalized path retention with both C^O and $\tau(P^M)$. Prediction 3 is unavailable unless τ passes its independence gate.

7. What Is and Is Not Empirical

7.1 Analytic or Constructive Statements

Path and endpoint are mathematically different; total variation cannot be smaller than endpoint distance under a metric satisfying the triangle inequality. Information absent from every licensed channel cannot be reconstructed without another source. Under a fixed resource budget and encoder class, temporal coverage and retained resolution generally trade off. These are not empirical discoveries of ER.

7.2 Measurement and Manipulation Gates

Independent h justification, finite-band refinement stability, adversarial metric challenge, history dependence, non-designed sufficiency, candidate-bank constructibility, rival-summary identifiability, τ -independence, raw-channel path adequacy, raw-channel probe fairness, proxy adequacy, disjoint data roles, and verified observer manipulations are gates.

Gate failure makes a downstream claim undefined or exploratory under the registered design; it is not itself a successful or failed relational prediction.

7.3 Empirical Wagers

ER risks failure if a path-adequate, non-path-supervised observer retains less normalized path information than an eligible rival; if a registered degradation shifts retention toward a proxy other than the independently verified residual proxy; or if the optimal readable scale fails to occupy and move through the jointly predicted $\tau(P^M)/C^O$ region.

8. Canonical Artificial-System Test

8.1 Candidate Pool, Screening, and Constructibility

The target dynamics are fixed first. A candidate pool is then generated by real continual-learning runs across pre-registered task and environment conditions. The protected procedure computes path and rival summaries after generation. A deterministic selection algorithm screens the pool for range coverage and low collinearity without using studied-observer or probe outcomes.

Anchor patterns remain useful diagnostics: low-path/few-transition, high-path/many-transition, high-path/few-transition, and, where constructible, low-path/many-small-transition histories. They are selected from the candidate pool rather than directly written into the target's internal state sequence.

Constructibility Gate

Established transitions consume time because each admitted state must persist for at least h . The protocol-specific upper bound depends on boundary rules, but the basic constraint is approximately

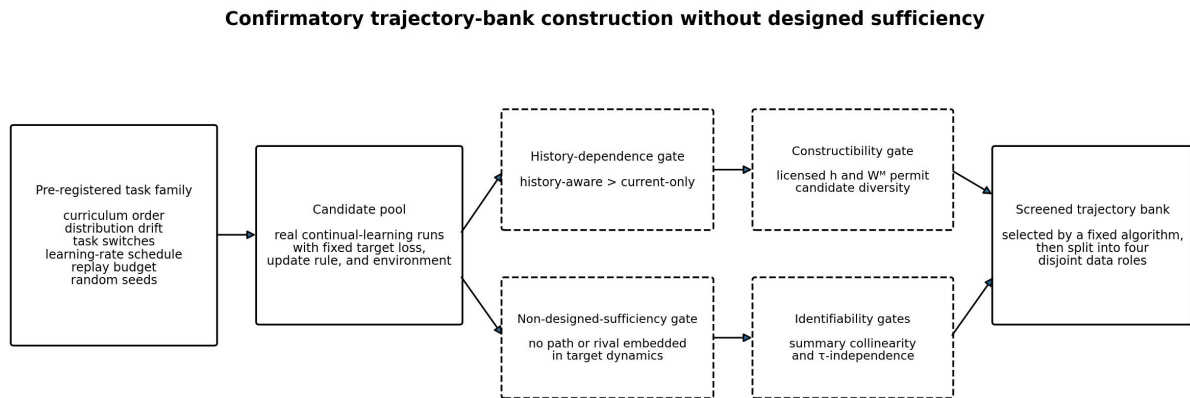
$$N_{\max} \lesssim \left\lfloor \frac{W_M}{h_{\min}} \right\rfloor$$

The study must show that the admissible h -range and W^M permit the count, magnitude, temporal-placement, and timescale variation needed to satisfy registered decorrelation margins. If a bank cannot be built from genuine adaptive runs, the corresponding path-versus-rival claim is unavailable.

Identifiability and τ -Independence Gates

Before observer testing, the study reports the candidate and selected-bank correlation matrices together with a registered diagnostic such as VIF or condition number. Path must be sufficiently separable from every claimed rival. $\tau(P^M)$ must also be below registered dependence thresholds with h , transition count, recency, and path.

If adequate decorrelation is possible only by directly synthesizing internal-state sequences, the target class fails the confirmatory screening requirement. This is the decorrelation squeeze: natural adaptive meaning and statistical separation may not coexist for every target family.



Decorrelation squeeze: if adequate summary separation can be achieved only by directly synthesizing internal-state sequences, the path-versus-rival comparison is not confirmatory for that target class.

Figure 2. Confirmatory bank construction. An independently motivated task family generates real continual-learning histories. History-dependence, non-designed-sufficiency, constructibility, identifiability, and τ -independence gates precede deterministic screening and the four disjoint data roles.

8.2 Disjoint Data Roles

The screened bank is partitioned at the trajectory level into four disjoint roles: observer training, raw-channel adequacy and probe-fairness calibration, frozen-probe calibration, and final held-out testing. No trajectory may appear in more than one role.

Prediction 1 — Normalized Path Retention Versus Eligible Rivals

On path-adequate channels, normalized path retention in the frozen observer representation should exceed the normalized retention of every rival that passed both target-bank identifiability and raw-channel probe-fairness gates. The high-path/few-transition anchor should group with other high-path histories rather than with other low-count histories.

The prediction is not supported if an eligible rival reaches the registered equivalence region, if the anchor follows the rival grouping, or if path superiority appears only after changing the target dynamics, frame, bank, metric family, h , adequacy decision, representation layer, probe family, raw-channel floor, normalization, or outcome outside the registered design.

Prediction 2 — Degradation-Specific Proxy Retention Shift

For each degradation class D_i , raw-channel path must be below its adequacy threshold while the named residual proxy q_i remains above its fairness-calibrated floor. Relative to the path-adequate condition, the frozen observer representation should show a registered normalized-retention shift away from path and toward q_i .

The prediction is not supported if retention shifts toward a different proxy, remains equally path-sensitive despite verified loss of path information, or requires selecting the successful proxy after inspection.

Prediction 3 — Joint Location of the Readable Scale

The primary encoder class, observer resource budget, amplitude-weighted τ estimator, secondary timescale challenger, and predicted optimum interval must be registered. The empirical target is the location of maximal normalized path retention as a joint function of C^O and $\tau(P^M)$.

The prediction is not supported if the optimum lies outside the registered interval, if increasing C^0 fails to move it by the registered direction and magnitude, if changing $\tau(P^M)$ fails to move it as predicted, or if the peak is determined only by the encoder and is insensitive to the target timescale. The prediction is undefined if τ fails its independence gate.

8.3 Confirmatory Commitments and Non-Circularity

Before primary outcomes are inspected, the study must register F^M , F^O , h , both τ estimators, refinement partitions, primary and challenger metrics, target task and update rules, candidate-pool conditions, bank-selection algorithm, all gates and thresholds, four data splits, adequacy decoder, primary and secondary probe families, raw-channel floors, retention normalization, observer objective and architecture, frozen representation, rivals, resource budget, predicted scale intervals and shifts, minimum effects, equivalence bounds, and undefined-case rules.

The target dynamics may not use $\hat{T}$, τ , or any rival summary as a designed sufficient statistic. The studied observer may not be trained on $\hat{T}$, its ordering, any rival label, or a monotone transform of those quantities. No probe gradient may update the observer. Observer performance may not define adequacy; probe results may not select the bank; and no downstream result may select a favorable frame, target generator, decoder, probe, summary, timescale estimator, or scale interval.

9. Limitations and Deferred Branches

9.1 Execution Limits of the Core Test

The h-squeeze appears in three places: a path estimate may lack a feasible stable scale, a passing scale may erase non-endpoint information, and an admissible h may leave too little time to construct a decorrelated bank.

The decorrelation squeeze is independent: genuine adaptive trajectories may preserve meaningful target dynamics while keeping path and rival summaries too correlated for confirmatory comparison. If sufficient separation requires directly synthesized state sequences, the test is not confirmatory for that target class.

History dependence may fail, τ may remain count-dominated, or a registered probe family may be unable to read a rival even from the raw channel. Each outcome removes a specific claim rather than licensing a favorable interpretation.

9.2 Limits of the Observer Model

Next-response prediction is a deliberately narrow canonical induction task. It establishes whether a generic predictive objective preserves path-sensitive structure; it does not show that all adaptive observers do so. A frozen probe establishes recoverability and normalized retention, not causal use in the original task.

Real observers may allocate attention adaptively, reconstruct missing history, combine proxies, or change encoding policy. The canonical test fixes one observer objective, encoder class, representation location, and probe family. Failure cannot be reinterpreted after the fact by replacing them with more favorable choices.

9.3 Branches Removed from the Present Core

Retrospective felt duration, childhood time compression, relationship divergence, personal continuity, inactivity, recovery, current applicability, velocity, acceleration, forgetting, interference, and relearning cost remain outside the present paper.

9.4 Next Tests and Freeze Boundary

The first test should run the complete continual-target protocol, including the history, non-designed-sufficiency, constructibility, identifiability, τ -independence, adequacy, and probe-fairness gates. A later intervention study may test causal use of retained path information.

This version is frozen as the final submission version. Further theoretical or protocol expansion should occur only if a new defect changes the falsifiability of the core target-observer relation. Copyediting, citation verification, figure refinement, and journal-format adaptation do not reopen the theory.

10. Conclusion

Evolutionary Relativity studies how an observed adaptive system's organizational path becomes reconstructable for another adaptive organization. The target history is constraining, but both its protected scientific estimate and its observer-side reconstruction are frame-qualified.

The target side must earn the path form without having the answer designed into its dynamics. Genuine continual-learning histories must be history-dependent, generated for an independent task, and screenable into a constructible bank that separates path from the rivals it claims to defeat.

The observer side must earn its interpretation without favorable readout bias. Raw-channel extractability and probe fairness are tested independently, the observer receives no path supervision, and frozen-representation results are normalized against each summary's raw-channel recoverability.

ER can therefore fail because path information was available but not retained, because a rival was retained more strongly, because degradation preserved the wrong proxy, or because the readable scale did not move with both observer capacity and an independent target timescale. The complexity of the canonical protocol is the cost of making those failures real rather than rhetorical.

References

- Alain, G., & Bengio, Y. (2016). Understanding intermediate layers using linear classifier probes. arXiv:1610.01644.
- Amari, S. (1998). Natural gradient works efficiently in learning. *Neural Computation*, 10(2), 251–276.
- Ambrosio, L., Gigli, N., & Savaré, G. (2008). *Gradient Flows in Metric Spaces and in the Space of Probability Measures* (2nd ed.). Birkhäuser.
- Burago, D., Burago, Y., & Ivanov, S. (2001). *A Course in Metric Geometry*. American Mathematical Society.
- De Lange, M., Aljundi, R., Masana, M., Parisot, S., Jia, X., Leonardis, A., Slabaugh, G., & Tuytelaars, T. (2021). A continual learning survey: Defying forgetting in classification tasks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(7), 3366–3385.
- Littman, M. L., Sutton, R. S., & Singh, S. (2001). Predictive representations of state. In *Advances in Neural Information Processing Systems* 14, 1555–1561.
- Mandelbrot, B. B. (1967). How long is the coast of Britain? Statistical self-similarity and fractional dimension. *Science*, 156(3775), 636–638.
- Parisi, G. I., Kemker, R., Part, J. L., Kanan, C., & Wermter, S. (2019). Continual lifelong learning with neural networks: A review. *Neural Networks*, 113, 54–71.
- Rao, C. R. (1945). Information and the accuracy attainable in the estimation of statistical parameters. *Bulletin of the Calcutta Mathematical Society*, 37, 81–91.
- Shannon, C. E. (1959). Coding theorems for a discrete source with a fidelity criterion. *IRE National Convention Record*, Part 4, 142–163.
- Tishby, N., Pereira, F. C., & Bialek, W. (2000). The information bottleneck method. arXiv:physics/0004057.
- Wang, K. (2026a). Connection Theory: A General Framework for Adaptive Network Dynamics — From Temporally Extended Existence to Emergent Dynamic Properties (Version 4.6) [Preprint]. Zenodo. <https://doi.org/10.5281/zenodo.21714702>
- Wang, K. (2026c). Flow Shaping Principle — Correspondence-Calibrated Pathway-Specific Updating (Version 3.6.3) [Preprint]. Zenodo. <https://doi.org/10.5281/zenodo.21714987>