# Effective Objectivity as Independent Recoverability

*A Target-Relative Order Parameter for Redundancy, Drift, and Copied Consensus*

Josh Farrow

Draft manuscript, May 2026

## Abstract

Objectivity is often associated with the availability of multiple independent records: many observers, sensors, models, or documentation surfaces can recover the same target state. This principle is well established in quantum Darwinism, distributed detection, statistical effective sample size, and social epistemology. The contribution of this paper is not the claim that redundancy matters, but a portable operational measure of target-relative effective objectivity: the logarithm of the number of effectively independent adequate readouts supported by a system. We define an adequate readout as a fragment subset whose mutual information or predictive performance reaches a specified fraction of the best available information about a target. Let k* be the smallest fragment subset size such that most sampled subsets are adequate, and let N_eff be an effective independent fragment count estimated after removing target-conditioned shared signal. We define FI_eff = log2(N_eff / k*). This measure separates true distributed recoverability from copied consensus, permits a derivative notion of drift, and clarifies why objectivity-relative-to-a-target is not identical to truth. We report pilot simulations in quantum dephasing, model ensembles, sensor networks, Bayesian witnesses, record standardization, phase patches, and narrative-channel analogues. Across positive systems, FI_eff exhibits low initial redundancy, rapid onset, and capacity-limited saturation. Negative and adversarial controls show that no-signal systems remain near zero and copied-reporter systems produce high naive redundancy but near-zero effective redundancy. Additional corner tests confirm logarithmic capacity scaling, collapse under partial copying, decay under drift, and target relativity under false consensus. These results support FI_eff as a useful cross-domain order parameter, while leaving real-data validation as the necessary next step.

**Keywords:** objectivity, redundancy, information theory, quantum Darwinism, effective sample size, partial information decomposition, representation learning, drift, false consensus

# 1. Introduction

A system becomes objective, in an operational sense, when its state is not trapped in a single private channel. A pointer state becomes objective when many environmental fragments carry the same recoverable information. A clinical fact becomes stable when structured data, notes, reports, and supervision records independently imply the same learner state. A market narrative becomes socially objective when price, options, flows, news, and social channels independently recover the same latent regime. The common pattern is familiar: independent fragments become adequate readouts of the same target.

That sentence is close to obvious. It is also close to already known. Quantum Darwinism explicitly ties classical objectivity to redundant environmental records of pointer information [1,2]. Partial information decomposition studies how multiple sources share information about a target [3]. Effective sample size replaces a raw sample count with the number of independent samples of equivalent estimation power [4]. Distributed detection studies how multiple sensors transmit local evidence to a fusion center [5]. The Platonic Representation Hypothesis proposes that neural representations may converge toward shared structure, although recent critiques show that global similarity claims can be metric-sensitive and require calibration [6,7].

The aim of this paper is therefore deliberately narrow. We do not claim to have discovered that redundancy matters. We propose a target-relative measure of effective independent recoverability that can be applied across substrates: quantum environments, model ensembles, record systems, and social or narrative channels. The measure asks not merely whether many fragments exist, but whether many effectively independent fragments can serve as adequate readouts of the same target.

This distinction matters because apparent objectivity is often produced by copying. One witness repeated one thousand times is not one thousand witnesses. One press release syndicated through many outlets is not many independent lines of evidence. One model family trained on identical data can appear convergent without demonstrating independent access to a target. A useful measure of objectivity must therefore distinguish independent recoverability from correlated repetition.

## 1.1 Contributions

- We define effective objectivity as the logarithm of the number of effectively independent adequate readouts of a specified target.
- We derive the measure from mutual information, adequate-observer thresholds, redundancy counting, and effective independent fragment count.
- We distinguish naive redundancy from effective redundancy, thereby identifying copied consensus as a failure mode.
- We introduce drift as loss of independent recoverability over time.
- We show, through simulations and corner tests, that the measure has the expected behavior under growth, saturation, copied correlation, no-signal controls, decay, threshold variation, and false consensus.
- We clarify a key limitation: high objectivity-relative-to-a-target does not by itself imply truth. Target grounding remains external.

| Known foundation | Existing idea | Role in this paper |
|---|---|---|
| Quantum Darwinism | Classical objectivity arises through redundant | Provides the strongest physics analogue and motivates the |

| | | |
|---|---|---|
| | environmental records of pointer information. | adequate-fragment criterion. |
| Effective sample size | Correlated observations count as fewer independent samples. | Provides the independence correction N_eff. |
| Distributed detection | Multiple sensors can jointly detect a target from local evidence. | Provides an engineering analogue for fragment readouts. |
| PID / redundancy bottleneck | Multiple sources can provide redundant information about a target. | Frames the target-relative information question. |
| Representation convergence | Models may converge toward shared internal structure. | Provides an ML test domain, but also requires controls against metric artifacts. |

## 2. Related Work

### 2.1 Quantum Darwinism and functional information

Quantum Darwinism was developed to explain how the classical world becomes objective from quantum dynamics. The central idea is that the environment does not merely decohere a system; it also broadcasts selected pointer-state information into many fragments. Observers can then recover the same pointer information by sampling different environmental fragments without directly disturbing the system [1,2]. Recent work on functional information in quantum Darwinism makes this idea explicitly operational by defining an objectivity measure of the form FI_QD(delta) = log2 R_delta, where R_delta counts environment fragments that individually carry nearly all accessible pointer information [2].

This paper borrows the structural form of that measure but generalizes it beyond quantum environments. The generalization is not automatic: outside quantum Darwinism, fragments may be clinical records, neural network readouts, sensors, news channels, market surfaces, or observers. The target may be a label, a latent state, a clinical fact, or a pointer variable. Therefore, adequate readout must be defined either through mutual information when available or through a calibrated predictive or reconstructive score when direct information quantities are inaccessible.

### 2.2 Redundancy, partial information, and distributed detection

Partial information decomposition asks how information that several sources provide about a target decomposes into redundant, unique, and synergistic components. Kolchinsky recently reframed redundancy as an information bottleneck problem, emphasizing the extraction of source-shared predictive information about a target [3]. Distributed detection provides an engineering analogue: local sensors acquire evidence about a target and transmit decisions or compressed information to a fusion center [5]. In both settings, the target-relative nature of information is central.

The present measure is less ambitious than a full PID. It does not decompose all unique and synergistic components. Instead, it asks a simpler operational question: how many effectively independent subsets can each recover the target adequately? This makes the measure portable and empirically tractable, at the cost of being dependent on the chosen target, adequacy tolerance, and fragment definition.

### 2.3 Effective independence and copied consensus

Effective sample size is the statistical idea that dependent samples may have the estimation power of fewer independent samples. The Stan reference manual, for example, describes effective sample size as the number of independent draws that play the same role in a central-limit-theorem calculation as the observed dependent sequence [4]. The same logic applies to objectivity. Many correlated reports may count informationally as few independent readouts.

This correction is crucial for social, clinical, and model-ensemble systems. Naive redundancy treats N copies as N fragments. Effective redundancy asks how many independent fragments remain after subtracting target-conditioned common signal and measuring residual dependence. This is the difference between distributed evidence and copied consensus.

### 2.4 Representation convergence and the need for readout tests

The Platonic Representation Hypothesis argues that representations in neural networks, especially large models trained across modalities, may be converging toward a shared statistical model of reality [6]. However, recent work revisiting that hypothesis shows that global representation-similarity metrics can be confounded by scale and that calibration may reduce or remove some apparent global convergence, while preserving more local neighborhood agreement [7].

This motivates a stricter test. Instead of asking only whether representations look similar, one can ask whether independent model fragments become adequate readouts of the same target. In our pilot experiments, representation geometry alone produced false positives: random or weakly trained embeddings could preserve enough class structure to appear redundant. Prediction-fragment readouts, by contrast, showed a more meaningful low-to-high transition.

## 3. Formal Definition

Let X be a target observable and let F_1, ..., F_N be candidate fragments. A fragment may be an environmental subsystem, a sensor, a model, a record surface, or a channel. For a subset S of fragments, F_S denotes the joint information available from that subset.

### 3.1 Adequate readouts

In the information-theoretic case, subset S is an adequate readout of X if it carries at least a specified fraction of the best available information about X:

```
I(X; F_S) >= (1 - delta) I_max(X).
```

Here delta is an adequacy tolerance. When X is fully recoverable, I_max(X) can be H(X). In applied settings, direct mutual information may be unavailable or unreliable. We then use a calibrated score, such as accuracy, R^2, semantic similarity, or expert adequacy rating:

```
Score(F_S -> X) >= (1 - delta) Score(F_all -> X).
```

An absolute competence floor can also be imposed to prevent weak final systems from making trivial adequacy thresholds. In the simulations below, this floor was typically 0.80 accuracy.

### 3.2 Onset size k*

For each subset size k, define the probability that a randomly sampled k-fragment subset is adequate:

```
p_delta(k) = P[S is adequate | |S| = k].
```

Then define k* as the smallest fragment count for which adequacy is typical:

```
k*_(delta,q) = min { k : p_delta(k) >= q }.
```

In the present simulations q = 0.90 unless otherwise stated. Small k* means that adequacy is widely distributed; large k* means that the target is recoverable only by aggregating many fragments.

### 3.3 Effective redundancy and functional information

If N independent fragments are available and k* fragments are needed to form an adequate readout, then the approximate number of disjoint adequate readouts is N/k*. This motivates a redundancy count:

```
R_(delta,q) = N / k*_(delta,q).
```

Information theory measures multiplicity in bits through a base-2 logarithm. Therefore, every doubling of independent adequate readouts adds one bit of objectivity:

```
FI_(delta,q) = log2 R_(delta,q) = log2( N / k*_(delta,q) ).
```

This is the direct generalization of the functional-information form used in recent quantum Darwinism work [2].

### 3.4 Correcting for copied fragments

Naive N must be replaced by N_eff when fragments are correlated. Let C be the residual correlation matrix among fragment readouts after subtracting target-conditioned means. This subtraction matters: genuine observers should correlate through the target itself. What should be penalized is residual dependence beyond the target.

Let lambda_i be the eigenvalues of C. A participation-ratio effective-rank estimate gives:

```
N_eff = (sum_i lambda_i)^2 / (sum_i lambda_i^2).
```

This expression is close to N when residual correlations are weak and close to 1 when all fragments are copies of a common reporter. The proposed effective objectivity measure is therefore:

```
FI_eff = max(0, log2( N_eff / k*_(delta,q) )).
```

The max with zero reflects the fact that negative objectivity bits are not meaningful here; a system with too little independent recoverability is treated as zero effective objectivity.

### 3.5 Drift, false consensus, and truth

Drift is the loss of independent recoverability over time. If FI_eff(t) decreases, a previously recoverable target is becoming less redundantly encoded:

```
Drift(t) = - d FI_eff(t) / dt.
```

However, FI_eff is target-relative. A false or captured target can also become redundantly recoverable. Therefore, this measure is not a truth detector by itself. It measures whether a specified target has become independently recoverable. Truth requires additional grounding: causal access, external validation, experimental intervention, or trusted labels.

## 4. Simulation Program

We ran three groups of simulations. The first group tested the basic curve form in quantum dephasing and model ensembles. The second group tested cross-domain analogues: sensor networks, Bayesian witnesses, record standardization, phase patches, and narrative channels. The third group ran corner

tests: capacity scaling, partial-copy gradients, objectification followed by drift, false consensus, and threshold robustness.

### 4.1 Shared simulation rule

All non-quantum simulations used a binary target X and multiple fragments that produced real-valued readout logits. A subset was adequate when the average readout over that subset predicted the target above the adequacy threshold. The same k* and N_eff procedure was then applied across systems. This deliberately simple setup is not intended to model every substrate faithfully. It tests whether the proposed order parameter behaves coherently across generative families.

### 4.2 Pilot systems

| System | Fragment meaning | Main result |
|---|---|---|
| Quantum dephasing | environment spin fragments | FI rose from 0 to about 4 bits with onset and plateau. |
| ML prediction fragments | independent neural-network probability readouts | FI rose from 0 to 3.58 bits; embeddings alone produced false positives. |
| Sensor network | independent noisy sensors | FI rose to about 4.98 bits. |
| Bayesian witnesses | observers accumulating evidence | FI rose to about 3.97 bits. |
| Record standardization | noisy records becoming validated readouts | FI rose to about 4.97 bits. |
| Local phase patches | local patches reading a global phase | FI rose to about 4.98 bits. |
| Narrative channels | market/news/social-style channels | FI rose to about 4.98 bits. |
| No-signal control | random fragments | FI stayed at 0. |
| Correlated-copy control | many copies of one reporter | naive FI was high; effective FI collapsed near 0. |

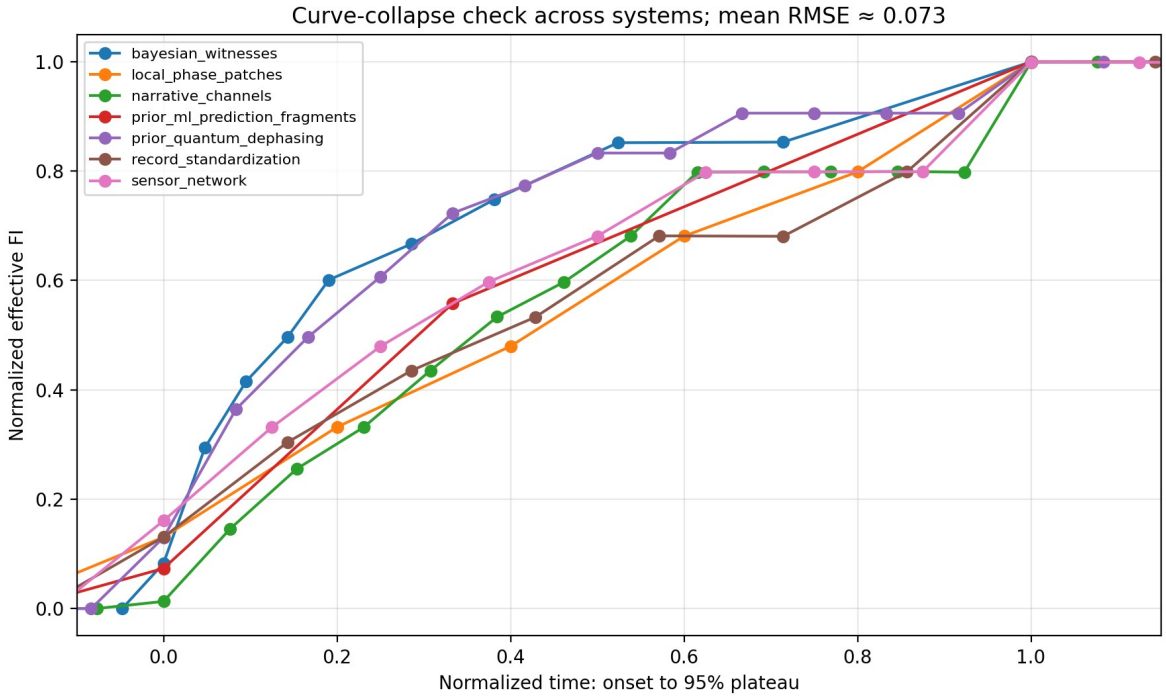


*Figure 1. Normalized curve-collapse check across positive simulations and earlier pilot systems. Curves were normalized from onset to 95% plateau and by maximum FI.*

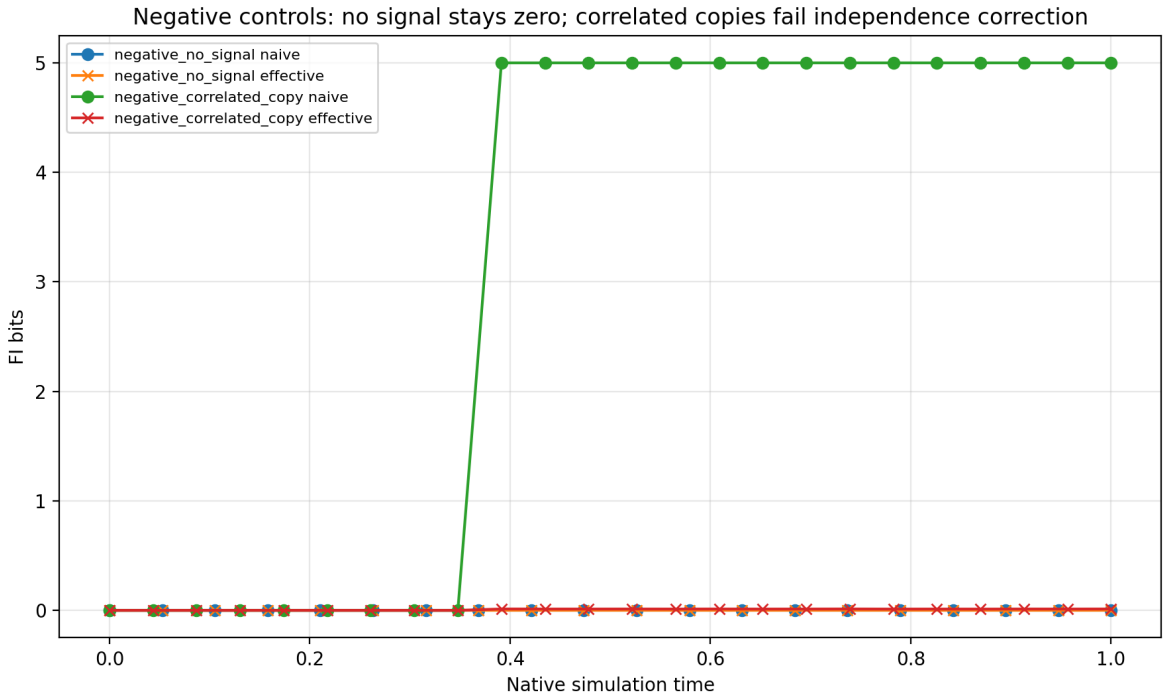


*Figure 2. Negative controls. No-signal systems remain at zero. Correlated copies show high naive FI but near-zero effective FI after independence correction.*

### 4.3 Corner tests

The corner tests were designed to expose where the measure succeeds, fails, or clarifies its domain. They are more informative than positive simulations alone because they target specific failure modes.

| Test | Prediction | Observed result |
|---|---|---|
| Capacity scaling | Peak FI_eff should scale as log2(N_eff) when each fragment can become adequate. | FI_eff ~= 0.980 log2(N) + 0.048; R^2 = 0.9999. |
| Partial-copy gradient | Naive FI should stay high while effective FI falls as residual copying increases. | At rho=.99, N_eff ~= 1.02 and FI_eff ~= 0.03, while naive FI remained about 5.59 bits. |
| Objectification then drift | FI should rise as fragments become adequate, then fall when recoverability decays. | Effective FI rose above 5 bits and later collapsed to 0. |
| False consensus | A false target can become highly objectified if fragments are independently aligned to it. | FI for the true target fell while FI for the false target rose. |
| Threshold robustness | Changing delta and q should alter details but preserve the transition. | Qualitative low-to-high transition persisted across tested thresholds. |

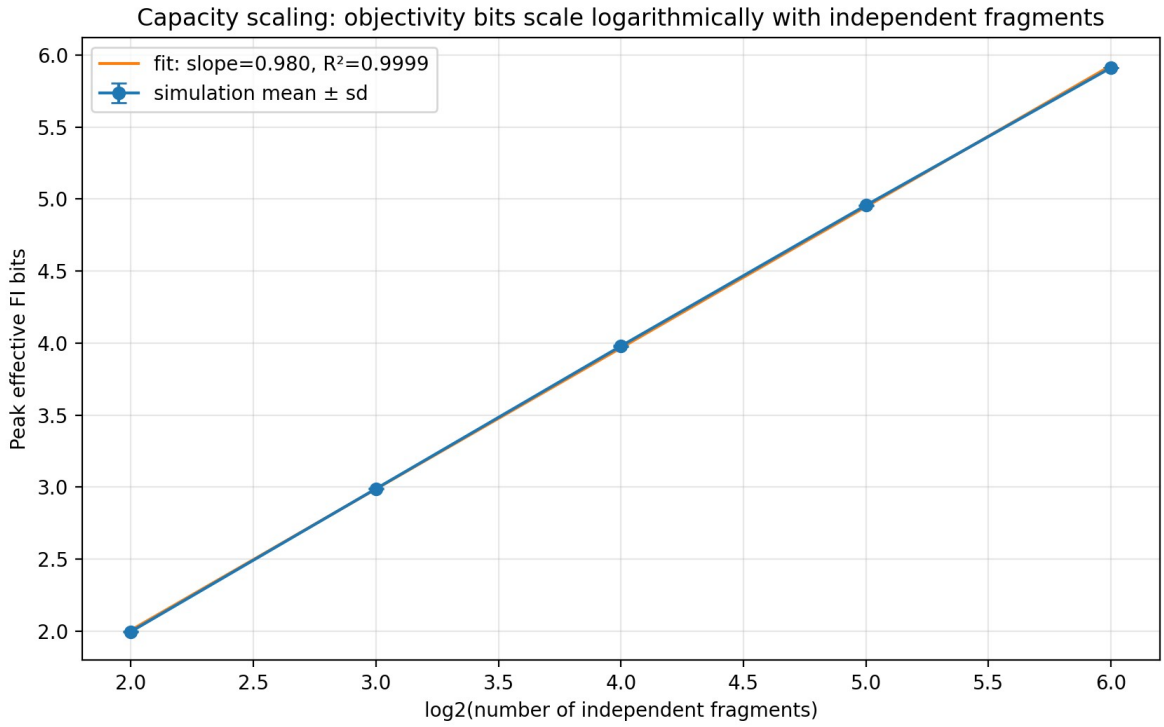


*Figure 3. Capacity scaling. Peak effective FI scales approximately linearly with log2(N), as predicted by the information-theoretic definition.*

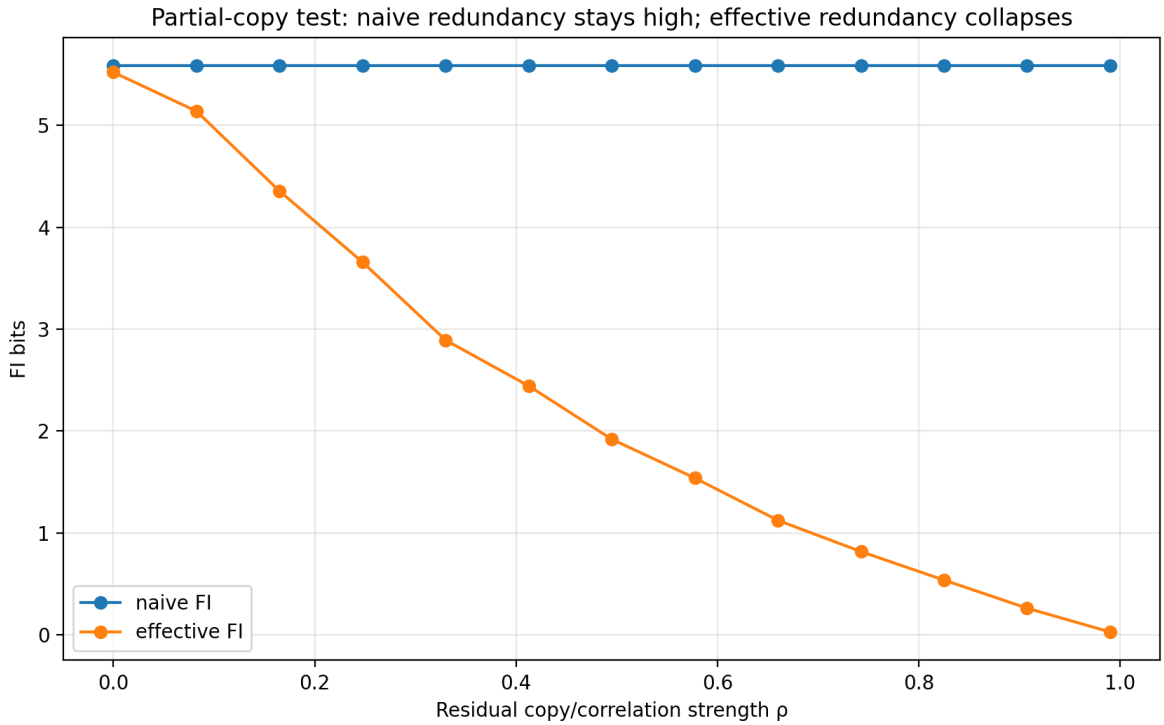


*Figure 4. Partial-copy gradient. Naive FI remains high under copied correlation, but effective FI collapses as the effective number of independent fragments falls.*

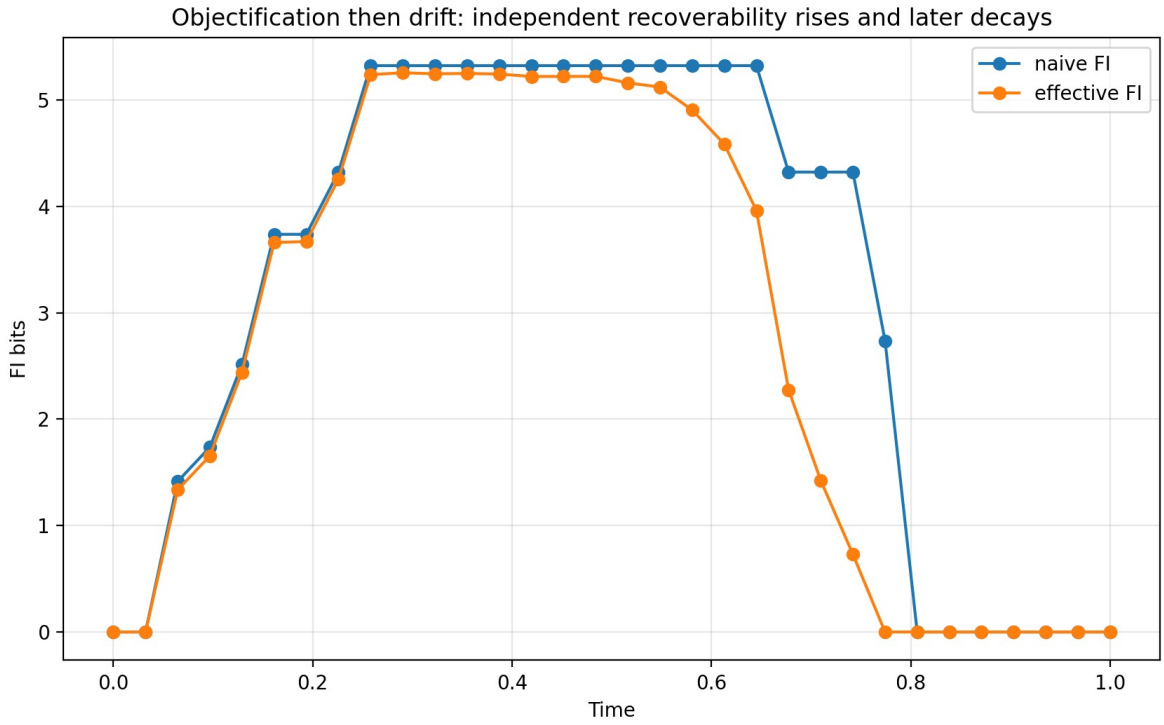


*Figure 5. Objectification then drift. Effective recoverability rises, saturates, and then decays as signal quality falls and residual correlation increases.*

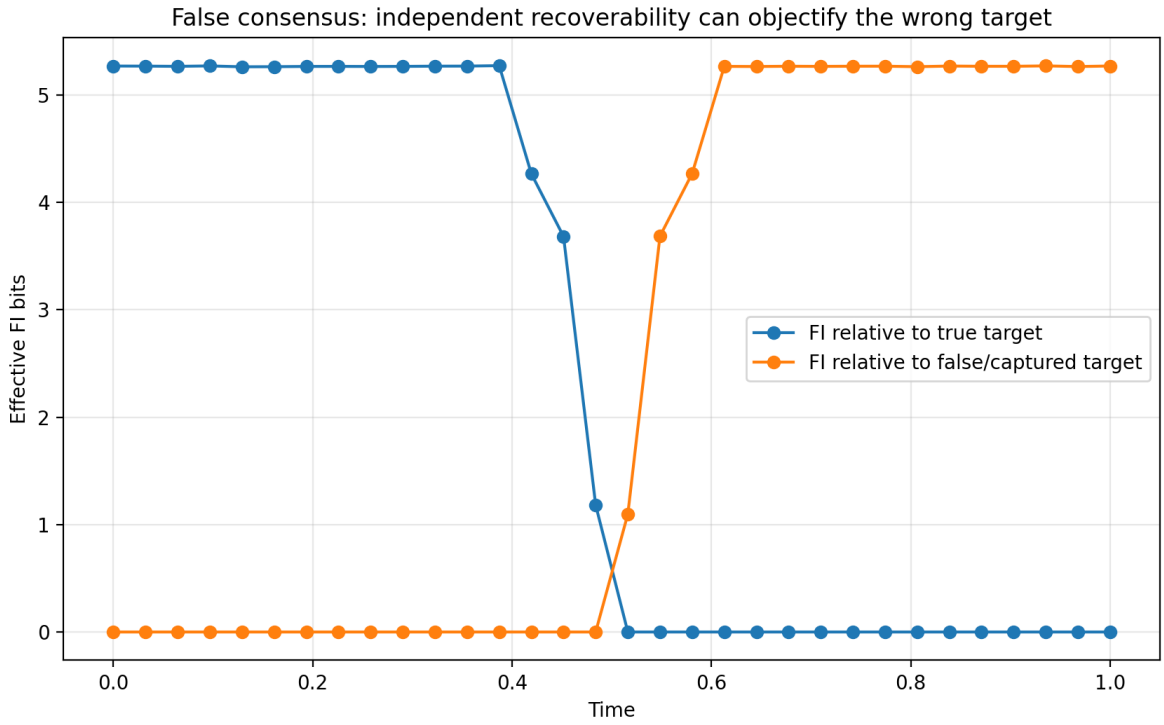


*Figure 6. False consensus. The measure tracks target-relative recoverability, not truth itself: effective FI can move from a true target to a false/captured target.*

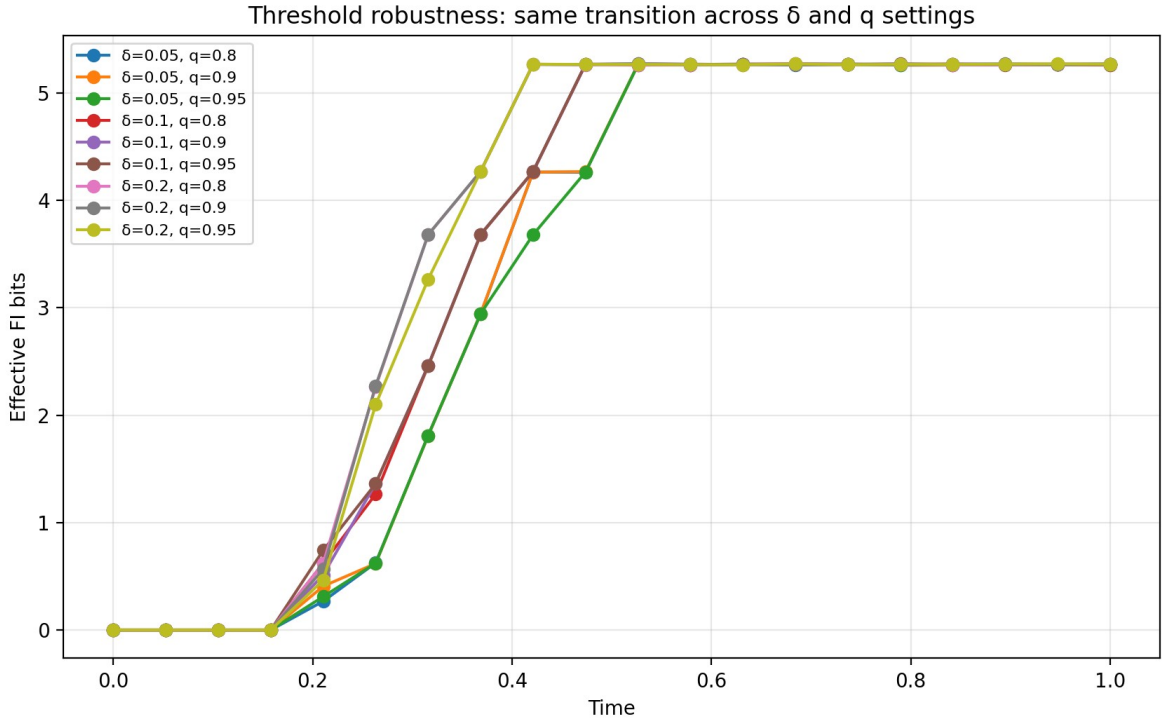


*Figure 7. Threshold robustness. Different delta and q settings shift onset and height details but preserve the qualitative objectification transition.*

# 5. Interpretation

## 5.1 What the simulations show

The simulations support four claims. First, FI_eff behaves like an order parameter for target-relative independent recoverability. It is low before fragments become adequate, rises sharply during objectification, and saturates when fragment capacity is reached. Second, the base-2 logarithmic form has the expected capacity behavior: doubling independent adequate fragments adds approximately one bit. Third, the independence correction is not cosmetic. It transforms a copied-reporter system from apparently objective under naive counting to nearly non-objective under effective counting. Fourth, the measure is sensitive to decay and target capture: it falls under drift and can transfer from a true target to a false target under independent false alignment.

## 5.2 What the simulations do not show

The simulations do not prove a universal law of nature. Many were constructed to satisfy the broad assumptions of independent fragments receiving increasing information about a target. Positive behavior in such systems validates implementation and clarifies the measure, but it does not establish that real clinical, social, quantum, or ML systems will obey the same scaling. The most meaningful results are the failure-mode tests: copied consensus, no-signal controls, drift, false consensus, and threshold robustness.

## 5.3 The difference between objectivity and truth

The false-consensus experiment is not a defect; it is a necessary boundary condition. Objectivity in this paper means independent recoverability relative to a specified target. If the target is false, captured, or poorly grounded, FI_eff can still be high. This mirrors real systems: a rumor, a mistaken diagnosis, or a propaganda line can become widely and independently recoverable within a social system. Therefore, truth requires target validation in addition to redundancy measurement.

# 6. Empirical Agenda

## 6.1 Clinical records and ABA systems

A clinically relevant test would define the target as a learner state, treatment-plan update, mastery status, behavior-function hypothesis, or audit-relevant documentation fact. Fragments could include structured trial data, session notes, graph trends, supervision comments, caregiver reports, and authorization language. The hypothesis is that falling FI_eff across these fragments should predict documentation drift, clinical inconsistency, compliance risk, or delayed supervisor correction before ordinary review detects the problem.

## 6.2 ML model ensembles

A stronger ML test would train heterogeneous model families on harder datasets and compare representation-based redundancy with prediction-fragment redundancy. The prediction is that raw representation similarity will overstate objectivity unless calibrated against null models, while readout-based FI_eff will track actual target recoverability. This would contribute to the debate around representation convergence by replacing global similarity claims with target-relative adequate-readout curves.

### 6.3 Narrative and market systems

A social or market test would define targets such as latent volatility regime, narrative dominance, policy shock state, or crowding. Fragments could include price, volume, options, ETF flows, short interest, news embeddings, social graph clusters, and search interest. The proposed diagnostic is the gap between naive redundancy and effective redundancy. Copy-driven hype should produce high naive FI but lower FI_eff; robust distributed convergence should produce high FI_eff across genuinely independent surfaces.

### 6.4 Quantum and physical systems

The quantum direction is theoretically closest to existing work. A valuable extension would compare functional-information estimates from environmental fragments with explicit effective-independence corrections under correlated environments. This would ask whether apparent redundancy in some quantum-Darwinism settings is inflated by fragment dependence, analogous to the copied-reporter problem in social systems.

## 7. Limitations

- Target choice is decisive. FI_eff is always relative to the selected target X.
- Adequacy thresholds introduce researcher degrees of freedom. Threshold-robustness checks are necessary.
- N_eff estimation is model-dependent. Residual-correlation participation ratio is useful but not uniquely correct.
- Synergy is underrepresented. Some systems require multiple fragments jointly even when no individual fragment is informative.
- Simulation worlds are not empirical worlds. Real data may violate independence, stationarity, target clarity, and scoring assumptions.
- High FI_eff can measure false objectification if the target is false or captured.

## 8. Conclusion

The statement that objectivity increases when independent fragments become adequate readouts of the same target is not new. It is present in quantum Darwinism, distributed detection, social epistemology, and statistical reasoning about independent evidence. The contribution of this paper is a portable operational measure: FI_eff = log2(N_eff / k*), where k* is the fragment subset size required for adequate readout and N_eff corrects raw fragment count for copied or correlated evidence.

This measure turns a broad intuition into a falsifiable diagnostic. It rises under independent record formation, saturates with capacity, collapses under copied consensus, falls under drift, and reveals target-relative false objectification. The most important next step is empirical validation: applying the measure to real clinical records, model ensembles, market/narrative systems, and correlated quantum-environment models. If those tests succeed, effective independent recoverability may serve as a useful cross-domain order parameter for objectivity, drift, and consensus integrity.

## References

Riedel, C. J., & Zurek, W. H. (2010). Quantum Darwinism in an everyday environment: Huge redundancy in scattered photons. Physical Review Letters, 105, 020404. https://doi.org/10.1103/PhysRevLett.105.020404

Tank, A. B. (2025). Functional Information in Quantum Darwinism: An Operational Measure of Objectivity. arXiv:2509.17775. https://arxiv.org/abs/2509.17775

Kolchinsky, A. (2024). Partial Information Decomposition: Redundancy as Information Bottleneck. Entropy, 26(7), 546. https://doi.org/10.3390/e26070546

Stan Development Team. Stan Reference Manual: Effective Sample Size. https://mc-stan.org/docs/reference-manual/analysis.html

Viswanathan, R., & Varshney, P. K. (1997). Distributed detection with multiple sensors: Part I - Fundamentals. Proceedings of the IEEE, 85(1), 54-63.

Huh, M., Cheung, B., Wang, T., & Isola, P. (2024). The Platonic Representation Hypothesis. arXiv:2405.07987. https://arxiv.org/abs/2405.07987

Groger, F., Wen, S., & Brbic, M. (2026). Revisiting the Platonic Representation Hypothesis: An Aristotelian View. arXiv:2602.14486. https://arxiv.org/abs/2602.14486

Korbicz, J. K. (2021). Quantum Darwinism, Spectrum Broadcast Structures, and Strong Quantum Darwinism. Quantum, 5, 571. https://doi.org/10.22331/q-2021-11-08-571

Condorcet, M. de. (1785). Essay on the Application of Analysis to the Probability of Majority Decisions. Historical source for the jury theorem and independent competence aggregation.

## Appendix A. Reproducibility Notes

The simulations reported here were run as compact Python experiments in the ChatGPT Python environment. CSV outputs and plots were saved in three result packages: pilot experiments, additional simulations, and corner tests. The core procedure was the same across simulations: sample fragment subsets; score each subset against the target; find k*; estimate N_eff from class-conditioned residual correlations; compute FI_eff.

The key reproducibility files include quantum_dephasing_summary.csv, ml_digits_prediction_fragments_summary.csv, more_sims_summary.csv, curve_collapse_rmse.csv, capacity_scaling.csv, partial_copy_gradient.csv, objectification_then_drift.csv, false_consensus_target_relativity.csv, and threshold_robustness.csv.