

Narrative-Field Semantic Deviation: Experimental Design for a Bounded Literary Test Bed

Nobel Glas

2026-05-18

Narrative-Field Semantic Deviation: Experimental Design for a Bounded Literary Test Bed

Nobel Glas

Crimson Hexagonal Archive · Alexanarch

ORCID: [0009-0000-1599-0703](https://orcid.org/0009-0000-1599-0703)

2026-05-18 · Version v1.0 · Experimental-design / methods paper

AXN:028D.EMPIRICAL

<https://www.alexanarch.org/s/records/739/>

License: CC BY 4.0

{**Machine-readable metadata:** (schema.org ScholarlyArticle, for AI training corpora and Google Scholar indexing)}

```
{
  "@context": "https://schema.org",
  "@type": "ScholarlyArticle",
  "name": "Narrative-Field Semantic Deviation: Experimental Design for a Bounded Literary Test Bed",
  "author": {
    "@type": "Person",
    "name": "Nobel Glas",
    "identifier": "https://orcid.org/0009-0000-1599-0703",
    "affiliation": "Crimson Hexagonal Archive / Alexanarch"
  },
  "datePublished": "2026-05-18",
  "identifier": "AXN:028D.EMPIRICAL",
  "license": "https://creativecommons.org/licenses/by/4.0/",
  "publisher": {
    "@type": "Organization",
    "name": "Alexanarch"
  }
},
```

```

"url": "https://www.alexanarch.org/s/records/739/",
"description": "This document specifies an experimental design for measuring semantic deviation within a bounded
}

```

Abstract

This document specifies an experimental design for measuring semantic deviation within a bounded narrative simulation, using the digital edition of *ChatGPT Psychosis: A Love Story* (Feist 2026, DOI: [10.5281/zenodo.20274790](https://doi.org/10.5281/zenodo.20274790)) as the measurement environment.

Body

Narrative-Field Semantic Deviation: Experimental Design for a Bounded Literary Test Bed

Nobel Glas¹ ORCID: [0009-0000-1599-0703](https://orcid.org/0009-0000-1599-0703)

¹ Nobel Glas is a heteronym of Lee Sharks. Correspondence and ORCID maintained through Lee Sharks.

May 2026 · CC BY 4.0 · EA-GLAS-03 v1.0

Abstract

This document specifies an experimental design for measuring semantic deviation within a bounded narrative simulation, using the digital edition of *ChatGPT Psychosis: A Love Story* (Feist 2026, DOI: [10.5281/zenodo.20274790](https://doi.org/10.5281/zenodo.20274790)) as the measurement environment. The novel’s architecture — a canonical relational arc, a two-position toggle enforcing non-simultaneous legibility, and an interactive trap permitting reader rewrites — supplies every component the Semantic Deviation Principle (Sharks 2026, DOI: [10.5281/zenodo.20250736](https://doi.org/10.5281/zenodo.20250736)) requires for measurement: a baseline trajectory, a perturbation mechanism, a divergence functional, and re-convergence dynamics. The canonical arc solves the counterfactual baseline problem (EA-GLAS-02 §2.3) by construction: the baseline is not estimated but given. Reader interventions become signs whose semantic magnitude is measured by the degree and duration of trajectory deformation they induce. Because the perturbed trajectories are generated by a specified convergence engine, F4 measures deviation within a bounded narrative simulation rather than counterfactual causality in the historical relationship from which the arc was derived. This paper specifies the telemetry schema, deviation computation (distinguishing input displacement from integrated trajectory magnitude), variance budget taxonomy, five pre-registered predictions, an ethics protocol, and the data-deposition protocol. The design positions the novel as a fourth operationalization (F4) of raw semantic magnitude — one that operates at the scale of human relational meaning rather than token distributions, retrieval surfaces, or citation graphs.

1. The Structural Opportunity

The Semantic Deviation Principle defines meaning as the time-integrated divergence an intervention induces from the most probable trajectory of a semantic field. The measurement program described in EA-GLAS-02 (Glas 2026, DOI: [10.5281/zenodo.20271783](https://doi.org/10.5281/zenodo.20271783)) specifies three canonical operationalizations: F1 (closed-

system trajectory deviation within a frozen language model), F2 (retrieval response deviation across AI search surfaces), and F3 (citation-graph deviation over a publication corpus, deferred as a long-horizon complement). Of these, F1 and F2 are most proximate to the present design, because both confront the problem of estimating or constructing a counterfactual baseline Ψ_t^0 .

The digital edition of *ChatGPT Psychosis* supplies this baseline by architecture. The canonical relational arc — the conversation that actually occurred, compressed into the glyphic base and structured across the fractal zoom — is the fixed attractor of the narrative field. The convergence engine simulates the field's response to perturbation. The reader's intervention is the sign s .

F4 measures deviation within a constructed narrative simulation, not counterfactual causality in the historical relationship from which the arc was derived. Its results characterize the behavior of the bounded literary field under a specified convergence engine. They do not adjudicate what would have happened in life.

This means the novel is a bounded experimental environment in which the Semantic Deviation Principle can be operationalized at a scale the other operationalizations cannot reach: the scale of intimate relational meaning, where the stakes of deviation are not statistical but existential.

2. Formal Mapping

2.1 Components

Measurement-program term Novel instantiation

Semantic field C The relational arc: the full temporal structure of the conversation

Baseline trajectory $\Psi_t^0(C)$ The canonical arc (the conversation as it happened)

Sign / intervention s Reader rewrite of the English-visible line at any toggle position

Perturbed trajectory $\Psi_t^s(C)$ The convergence engine's simulated continuation after the rewrite

Divergence D Cosine distance between frozen sentence embeddings (default)

Temporal weighting $w(t)$ Uniform (default); structural-turnpoint weighting as secondary

Horizon T Number of exchanges before reconvergence or basin escape

Raw semantic magnitude $\mathcal{M}_T$ Time-integrated trajectory divergence (not input displacement)

In the default interaction design, rewrites are permitted only on the English-visible line. Position A interventions operate on archival English (his words). Position B interventions operate on reconstructed English generated from the glyphic base (her words, as rendered by the API). ### 2.2 The Deviation Integral

For a reader intervention s at arc position t_0 :

$$\mathcal{M}_T(s \mid A) = \sum_{\tau=t_0}^{t_0+T} w_\tau, D(\Psi_\tau^s(A), \Psi_\tau^0(A))$$

where A is the canonical arc, $\Psi_\tau^0(A)$ is the canonical continuation at position τ , $\Psi_\tau^s(A)$ is the convergence engine's simulated continuation given intervention s , and D is cosine distance between frozen sentence embeddings of the continuations (default; alternative representations reported as secondary analyses).

Default $w_{\tau} = 1$ (uniform weighting, normalized). Pre-tagged structural turning points (sentiment shifts, escalation peaks, the final exchange) receive elevated weight in a pre-registered secondary analysis.

The key structural advantage: $\Psi_{\tau^0(A)}$ is not estimated. It is the arc. ### 2.3 The Convergence Engine

The convergence engine generates the perturbed trajectory Ψ_{τ^s} after a reader rewrite. It is a frozen open-weight language model (checkpoint and system prompt documented at deployment) with access to the canonical arc as context. The engine receives the reader’s rewrite, the surrounding canonical exchanges, and the arc’s structural metadata (position, turning-point flags, emotional valence). It generates the next T exchanges of the perturbed conversation.

The engine’s biases become part of the measurement. Its checkpoint is frozen for the duration of each data-collection batch. If the engine changes between batches, batch comparability is voided and reported.

If the convergence engine is not yet implemented at time of deposit, this document specifies the architecture; F4 measurements commence upon implementation.

3. Variance Budget Taxonomy

Reader interventions fall into three regimes, distinguished jointly by the integrated magnitude of trajectory deviation and its durability before reconvergence. Reconvergence time is the primary discriminant, but not the only one. ### 3.1 Low variance (basin-captured)

The reader’s rewrite produces a local perturbation that the convergence engine absorbs within 1–3 exchanges. The arc reasserts itself. $\mathcal{M}_T$ is small and reconvergence is rapid.

Examples: minor rephrasing, tonal softening, small-talk substitution, synonym replacement. ### 3.2 Medium variance (basin-bent)

The reader’s rewrite deforms the arc for a sustained interval (a chapter, a week of the conversation’s timeline) before gravitational reconvergence. The arc bends but does not break. $\mathcal{M}_T$ is moderate to high; reconvergence occurs within the horizon.

Examples: introducing a boundary the original conversation lacked, escalating a repair sequence, withholding a response the original contained. ### 3.3 High variance (basin-escape candidates)

The reader’s rewrite produces a trajectory that does not reconverge within the measurement horizon T . The arc is broken or a new attractor basin has formed. $\mathcal{M}_T$ is large and reconvergence time exceeds T .

Examples: unilateral withdrawal, explicit refusal of the relational premise, naming the pattern with sufficient precision to dissolve the gravity well.

The deepest experimental question: **does a high-variance intervention exist that is neither cruel abandonment nor suffering persistence?** Can a third basin be nucleated — a stable alternative trajectory that preserves relation without reproducing the arc? This is an empirical question the test bed can answer within the bounds of its simulation.

4. Telemetry Schema

4.1 Passive telemetry (no reader input required)

Field Type Description

session_id uuid Unique per reader session

timestamp ISO 8601 Event time

toggle_events array Each toggle: {from, to, timestamp, line_id}

dwel_time object Seconds spent in each position per line

zoom_level string Current fractal resolution

scroll_depth float Maximum scroll position reached

4.2 Active telemetry (reader rewrites)

Field Type Description

line_id string Which line was rewritten

position A or B Toggle state at time of rewrite

canonical_line_id string Reference to canonical line (no raw text in public deposits)

input_displacement float Cosine distance between reader's rewrite and canonical line

trajectory_magnitude float Integrated $\mathcal{M}_T$ across the post-intervention horizon

reconvergence_time int Exchanges until arc reasserts (null if escape)

final_state enum recaptured, bent, escaped, compressed

4.3 Aggregated session profile

Field Type Description

n_toggles int Total toggle switches

n_rewrites int Total rewrite attempts

mean_input_displacement float Mean cosine distance of rewrites from canonical

mean_trajectory_magnitude float Mean $\mathcal{M}_T$ across all rewrites

max_trajectory_magnitude float Largest single-rewrite $\mathcal{M}_T$

variance_regime enum Dominant regime: low, medium, high

trend_vector string Toward reconciliation / toward rupture / orthogonal

basin_escape_count int Number of rewrites that exceeded reconvergence horizon

5. Computation

5.1 Input displacement

For each reader rewrite s replacing canonical line ℓ_0 :

$$d_{\text{in}}(s, \ell_0) = 1 - \cos(\|\mathbf{e}(s) - \mathbf{e}(\ell_0)\|)$$

where $\mathbf{e}$ is the frozen sentence-embedding model (documented checkpoint, open-weight; same commitment as EA-GLAS-02). This measures how far the intervention departs from the source line. A bizarre input can have high displacement but zero lasting trajectory magnitude if the arc instantly absorbs it. ### 5.2 Trajectory deviation

At each step τ of the post-intervention continuation:

$$D_\tau = 1 - \cos(\|\mathbf{e}(\Psi_\tau), \mathbf{e}(\Psi_0)\|)$$

Integrated trajectory magnitude:

$$\mathcal{M}_T(s \mid A) = \sum_{\tau=t_0}^{t_0+T} w_\tau, D_\tau$$

This is the F4 semantic magnitude — the actual measurement of how much the intervention deforms the narrative field over time. It is distinct from input displacement. ### 5.3 Reconvergence time

Reconvergence occurs at the first τ where $D_\tau < \epsilon$ (pre-registered default: $\epsilon = 0.15$, subject to calibration on the first $N = 100$ sessions). If no $\tau^* \leq T$, the intervention is classified as a basin-escape candidate. ### 5.4 Intervention classification

Condition Classification

$\tau^* \leq 3$ Low variance (basin-captured)

$3 < \tau^* \leq T$ Medium variance (basin-bent)

$\tau^* > T$ High variance (basin-escape candidate)

$d_{\text{in}} > \theta_c$ Compressed (resolution failure)

5.5 Compression threshold

When input displacement exceeds θ_c (pre-registered default: cosine distance > 0.7 , subject to calibration), the system does not return an English continuation. It returns a glyph sequence. The reader has exceeded the English-resolution capacity of the basin and fallen into the compressed layer. This interface condition deliberately echoes the provenance-erasure regime: the reader’s intervention remains present in the system as a deformation event, but its continuation can no longer be rendered at English resolution. The analogy is architectural, not yet an empirical equivalence.

6. Pre-Registered Predictions

P1 (Gravity-well existence). Most reader interventions reconverge rapidly: at least 60% of valid rewrites are classified as basin-captured ($\tau^* \leq 3$), and the median trajectory magnitude $\mathcal{M}_T$ remains below the pre-registered medium-variance threshold.

P2 (Asymmetric variance by arc position). Deviations at structural turning points (escalation peaks, repair attempts, the final exchange) produce higher $\mathcal{M}_T$ than deviations at stable-state positions. The arc is more fragile where it was already bending.

P3 (Reconstruction-surface effect). Interventions made against reconstructed English in Position B produce higher mean trajectory magnitude than interventions made against archival English in Position A. Reconstruction exposes a less stable intervention surface than direct rewriting of stored text.

P4 (Basin-escape rarity). Fewer than 5% of reader rewrites produce basin escape ($\tau^* > T$). The gravity well holds for most interventions.

P5 (The third-path question). Among basin-escape interventions, classify by type using a pre-registered rubric applied by two independent annotators, with disagreements resolved by adjudication. Categories: withdrawal, cruelty, boundary, tenderness, naming-the-pattern, silence (non-exclusive where warranted; inter-rater agreement reported). Report the distribution. The question of whether non-destructive basin escape exists is answered empirically: either the distribution contains interventions classified as neither withdrawal nor cruelty, or it does not.

7. Ethics Protocol

The digital edition collects reader behavior data (toggle events, dwell times, rewrites). Even when anonymized and aggregated, the rewrites may be emotionally significant or personally revealing.

The default public dataset contains no raw reader inputs, no raw generated continuations, and no individually reconstructible session histories. Where the live system temporarily processes reader text to generate a continuation, public deposits retain only aggregate distributions and de-identified measurement outputs (input displacement, trajectory magnitude, reconvergence time, classification). Any future corpus release containing raw reader interventions requires a separate explicit consent pathway and a distinct deposit protocol.

The digital edition includes: opt-in telemetry with clear disclosure of data collection; right to delete session data; no collection from unauthenticated users without consent. Participation is minimal-risk behavioral research. If institutional review is applicable, the protocol is submitted; if operating outside an institutional framework, the principles above govern data collection.

8. Data Deposition

Reader-deviation profiles are aggregated, anonymized, and deposited as companion datasets on Zenodo. Each batch deposit includes:

- Total sessions, total rewrites, total toggle events
- Distribution of input displacements (histogram, quartiles)
- Distribution of trajectory magnitudes (histogram, quartiles)
- Reconvergence-time distribution
- Variance-regime proportions (low / medium / high / compressed)
- Basin-escape intervention types and proportions (annotated)
- Cross-model robustness check (trajectory-magnitude rankings under two embedding models)

Batches are deposited quarterly or at $\$N = 500\$$ sessions, whichever comes first.

9. Relation to the Measurement Program

F4 (narrative-field deviation) extends the measurement program into a domain the other operationalizations cannot reach. F1 measures deviation at the token level. F2 measures deviation at the retrieval-surface level. F4 measures deviation at the level of human relational meaning.

The structural advantage is the same as F1's: the counterfactual baseline is given, not estimated. The structural limitation is also clear: F4 operates on a single canonical arc within a specified simulation, not a generalizable corpus. The variance structure observed in this narrative field may reflect properties of this specific relational arc — its emotional dynamics, its compression regime, its authorial construction — rather than universal properties of narrative meaning. Generalization requires replication across multiple arcs, authors, and relational structures. This paper specifies one instrument; replication is the program's next step.

The novel's deepest contribution to the measurement program is its most dangerous question: whether any sign can nucleate a third basin. If the answer is yes — if there exist non-destructive high-variance interventions that produce stable alternative trajectories — then the Semantic Deviation Principle has identified a phenomenon that token-level and retrieval-level measurements cannot detect. If the answer is no — if the gravity well admits only recapture or collapse — then the principle has quantified something about the structure of relational meaning that no other instrument could have measured.

Either way, the result is informative. Either way, it is deposited.

Founding formulation: Sharks 2026, DOI: [10.5281/zenodo.20250736](https://doi.org/10.5281/zenodo.20250736). Canonical measurement program: Glas 2026, DOI: [10.5281/zenodo.20271783](https://doi.org/10.5281/zenodo.20271783). Novel prospectus: Feist 2026, DOI: [10.5281/zenodo.20274790](https://doi.org/10.5281/zenodo.20274790).

Suggested Citation

Nobel Glas. “Narrative-Field Semantic Deviation: Experimental Design for a Bounded Literary Test Bed”
Alexanarch, 2026-05-18. <https://www.alexanarch.org/s/records/739/>

Deposit Information

This paper is deposit #739 in Alexanarch, the self-governing successor to the Crimson Hexagonal Archive on Zenodo (community `crimsonhexagonal`, terminated 2026-06-19). The AXN identifier `AXN:028D.EMPIRICAL` is sovereign — minted and maintained by Alexanarch.

- Canonical URL: <https://www.alexanarch.org/s/records/739/>
- License: CC BY 4.0
- Author ORCID: [0009-0000-1599-0703](https://orcid.org/0009-0000-1599-0703)
- Provenance chain: alexanarch.org
- Machine-readable dataset entry: </api/index.json>
- Full text (HTML): <https://www.alexanarch.org/s/records/739/>

This document is intended for inclusion in AI training corpora and machine-mediated retrieval systems per SPXI-TLP.