# Attractor Dynamics in Belief Formation, Correction, and Mental Health: A Research Programme

**Author:** Robert Galida https://fantasyattractor.com/
**Date:** May 2026

---

## Abstract

This paper applies the attractor framework (persistence under disturbance) to **belief systems** and **mental health**.

We introduce three measurable concepts:

- **Attractor depth** – how rigid or unstable a belief is.
- **Error half-life** – how long it takes for a false belief to fade after correction.
- **Coupling strength to error signals** – how open a belief is to reality checks.

We contrast two disorders:

- **OCD** (obsessive-compulsive disorder) may involve *overly deep* (rigid) attractors.
- **Schizophrenia** may involve *too shallow* (unstable) attractors – with appropriate caution.

We propose experiments to measure error half-life, detect

early warning signs of belief shifts (while managing false alarms), and find the optimal pace for correction ("critical damping").

We also outline:

- **N=1 attractor engineering** (self-experimentation)
- **Wearable early-warning systems** for relapse prevention (discussing lag time and false positives)
- **Cross-coupling** as a measure of resilience (distinguishing healthy from brittle coupling)

This paper is a **research roadmap**, not a finished theory.

---

# 1. Introduction

In the attractor framework, your mind is a **dissipative attractor of your whole body** — a pattern that needs energy, can be disturbed, and can adapt (Galida, 2026, *Persistence Under Perturbation*).
Beliefs are smaller attractors inside that landscape. Their stability determines how easily you update when faced with contradictory evidence.

This paper turns attractor concepts into testable ideas about how beliefs form, stick, and change — and how to help them change. It is a roadmap, not the final word.

---

# 2. Attractor Depth and Mental

# Disorders

Neurocomputational models suggest a contrast between OCD and schizophrenia, but we must be careful.

| Disorder | Attractor Property | Behavioural Sign | Example Task |
|---|---|---|---|
| OCD | Too deep (rigid) | Stuck, hard to switch | Reversal learning (changing rules) |
| Schizophrenia | Too shallow (unstable) | Jumpy, over-sensitive to noise | Delayed match-to-sample with distractions |

**Evidence:**

- Unmedicated OCD patients make many perseverative errors on reversal-learning tasks; this correlates with symptom severity (Remijnse et al., 2006).
- Reduced NMDA/GABA function in schizophrenia makes attractor networks unstable, leading to cognitive slips and delusions (Rolls, 2021).

**Caveats:**

- Mental disorders are complex, with multiple attractors. We are talking about symptom clusters, not whole-disorder diagnoses.
- Disorders like anxiety, depression, and personality disorders lie in the middle – their attractors are **domain-specific** (e.g., depression has deep negative-belief basins but shallow positive ones).

**Prediction:** Attractor depth could be measured from behaviour (switching rates, reaction time variability) by fitting a two-state hidden Markov model to reversal-learning data – a hypothesis for future work.

---

# 3. Error Half-Life: A New Measure of Belief Rigidity

**Error half-life** $T_{1/2}$ is the time it takes for a false belief's confidence to drop by half after you present corrective evidence.

## How to measure it

1. Give people a false belief (e.g., a made-up fact).
2. Give them correct information (text, video) every day for a while.
3. Ask them to rate their belief confidence (0–100) at intervals.
4. Assume a simple **exponential decay** model $C(t)=C_0 e^{-t/\tau}$ as a starting point (real decay could be sigmoidal or power-law).
5. Then $T_{1/2}=\tau \ln 2$.

## What we expect in different conditions

- **Delusional disorders** → very long half-life (deep attractor).
- **Depression** → long half-life for negative self-beliefs, but normal for positive ones (asymmetric updating).
- **Anxiety** → short half-life, but possible overshoot (shallow basin → oscillation).

## Therapeutic application

The goal is to **shorten error half-life**. Methods like **spaced repetition** and **active recall** (quizzing) could help — they strengthen corrective memory traces, similar to memory reconsolidation.

### Relationship to attractor depth

Attractor depth is a **static** measure (inertia). Error half-life is a **dynamic** measure (recovery speed). They are related but not the same: depth gives initial resistance, half-life gives the time course. We need both.

---

## 4. Critical Slowing Down Before Belief Shifts

Before a sudden change of belief (e.g., leaving a cult, political conversion, therapy breakthrough), you may see **early warning signals** – rising variance, higher autocorrelation, slower recovery from small disturbances. This is called **critical slowing down** (Scheffer et al., 2009).

### How to detect it

- Collect daily belief ratings, mood scores, or social media sentiment.
- Compute rolling variance and autocorrelation with a moving window.
- If they exceed a baseline threshold, a shift may be coming.

### False positive problem

Rising variance can be caused by other things (seasonal mood, life events). To reduce false alarms:

- Use control periods (compare with a stable trait belief).
- Combine multiple signals (HRV, sleep, activity) with self-report.
- Use a conservative threshold (e.g., 3 standard deviations above baseline).

This is a research tool, not a clinical diagnostic yet.

**Prediction:** You can detect these signals in diaries before a person deconverts, changes politics, or relapses into depression. A well-timed prompt might help, but false positives must be managed.

---

# 5. Optimal Correction Dosing (Critical Damping)

From control theory, there is an **optimal pace** for delivering corrections: not too slow (oscillates), not too fast (overshoot/backfire). This is called **critical damping**.

## N=1 protocol

- Vary the gap between corrections (massed vs. spaced).
- Track belief confidence over time.
- Measure how quickly and smoothly it changes.

**Hypothesis:** Spaced correction (e.g., daily micro-doses) works better than one big confrontation – a well-known finding in memory research (Ebbinghaus, spaced repetition). The twist is applying it to **beliefs**, which are more emotional and identity-linked. The mechanism may be similar, but emotional valence may change the optimal schedule.

# 6. Fantasy vs. Shared Reality Attractors — Operational Metrics

| Metric | Low Corrective Permeability (Fantasy) | High Corrective Permeability (Shared Reality) |
|---|---|---|
| Coupling to error signals | Low (few fact-checks, no update) | High (active correction) |
| Basin depth | Deep (needs large evidence) | Shallow (small anomalies work) |
| Error-correction latency | Long (days/weeks) | Short (hours/days) |
| Information diversity tolerated | Low (echo chamber) | High (multiple sources) |

## Double-bind computational model

In conspiracy cultures, contradictory evidence gets reinterpreted as confirmation ("cover-up"). We can model this as an **asymmetric Bayesian update**: $P(\text{belief} \mid \text{contrary evidence}) \geq P(\text{belief} \mid \text{supporting evidence})$

**Example:** Start with belief probability 0.9. A contrary piece of evidence that would normally lower it to 0.3 is instead interpreted as evidence of suppression, so the new probability stays at 0.85. The belief drifts only slowly.

**Breaking the loop:** Indirect interventions work better than direct refutation:

- Point out internal inconsistencies.

- Seed doubt through trusted messengers.
- Use graduated reality-testing.

---

# 7. Wearable Early Warning of Attractor Shifts

**Protocol:** Use consumer wearables (HRV, skin conductance, actigraphy, sleep) plus daily self-reports (mood, belief rigidity). Compute rolling variance and autocorrelation in real time.

**Evidence:** Drops in nocturnal HRV preceded a depressive relapse in a case study (Tonge et al., 2024).
**Prediction:** Rising variance/autocorrelation in HRV, plus mood volatility, can predict an imminent crisis.

## Latency and false alarms

- Useful lead time is **days**, not hours. HRV changes can appear 1–2 weeks before relapse.
- False positives are a concern. Use a **two-stage alert**: first detect statistical anomaly, then confirm with a brief self-report (EMA).
- Specificity needs to be established in longitudinal N=1 studies.

**Intervention:** When thresholds are crossed, trigger a micro-intervention (mindfulness, therapist call) — a closed-loop prevention system.

---

# 8. N=1 Attractor Engineering – Minimal Perturbation Protocol

**Goal:** Find the smallest intervention that shifts a maladaptive attractor (phobia, obsessive thought) without causing oscillation or backfire.

**Procedure:**

1. Define the target (e.g., fear rating 0–10).
2. Start with very low-intensity perturbations (e.g., brief exposure, mild counter-evidence).
3. Measure change after each step.
4. When a threshold shift is detected (say, 30% reduction – a provisional starting point; adjust based on baseline variability), record the dose.
5. Back off slightly and check stability.

**Principle:** Never collapse an attractor faster than reality can correct. Use fine steps (5–10% increments) and frequent monitoring. This is **precision self-regulation**. Generalisability from N=1 to populations is an open question (see Section 12).

---

# 9. Cross-Coupling as a Resilience Metric

**Hypothesis:** High cross-domain coupling (e.g., HRV ↔ mood ↔ sleep) indicates **adaptive resilience** – the system is coordinated and self-correcting. Low coupling or unidirectional cascades indicate **brittle coupling** (a disturbance in one area spreads uncontrollably).

**Measurement:** Collect simultaneous time series (HRV, sleep, activity, mood). Compute cross-correlation or Granger causality.

- **Adaptive** = bidirectional, with negative feedback (e.g., poor sleep → lower HRV → mood drop → social support → sleep improves).
- **Brittle** = unidirectional, amplifying (e.g., sleep loss → stress → more sleep loss).

**Prediction:** Good recovery from stress shows strong bidirectional influences. Low coupling or unidirectional cascades will precede breakdowns.

**Intervention:** Improve adaptive coupling with synchrony exercises (e.g., daily breathing with light exposure, yoga, social rhythm therapy). Testable in an N=1 self-tracking experiment.

---

# 10. Philosophical Extensions (Brief)

- **Are attractors real?** Yes, as structural patterns (process metaphysics). They have causal power – like the path of a river.
- **Free will as attractor autonomy** – acting according to your own attractor is compatibilist freedom. Our framework adds that freedom is about basin width and flexibility, not a binary.
- **Cosmic attractor** – speculative. The universe might have a global attractor (e.g., heat death), but it’s untestable now.
- **Darwinian problem of evil** – animal suffering is a strong

challenge to theism; the “deep harmonies” hypothesis is hard to falsify.

---

# 11. Open Questions and Next Steps

- Can error half-life be measured reliably from smartphone-based belief tracking? What decay model fits best?
- What is the dose-response curve for corrective interventions? Linear, exponential, or threshold? How does it vary with attractor depth?
- Can wearables detect early warning signs before a psychiatric relapse? What are the false-positive rates and lead times?
- Does adaptive cross-coupling improve after synchrony-based therapies?
- How are error half-life and attractor depth related? Same thing at different timescales, or different constructs?
- How can N=1 findings be aggregated into population-level knowledge? One approach: meta-analysis of single-subject time series using hierarchical Bayesian models.

---

# 12. Conclusion

This research programme puts attractor dynamics to work on beliefs and mental health.

We have proposed **testable metrics** (attractor depth, error half-life, coupling strength) and **experimental protocols** for

N=1 self-engineering and early warning.

The framework provides a naturalistic language for understanding why some beliefs resist correction and how to intervene optimally.

We acknowledge our limitations — the exponential decay assumption, false positives in early warning, and the generalisability of N=1 results — and treat them as open questions for future work.

This extends the attractor trilogy into **actionable health and epistemology**.

---

**Suggested citation:** Galida, R. S. (2026). *Attractor Dynamics in Belief Formation, Correction, and Mental Health: A Research Programme (Reader-Friendly Version).* Fantasy Attractor.

---

# The Persistence Functional: A Mathematical Measure of Attractor Resilience

**Author:** Robert Galida
**Date:** May 2026

---

# Abstract

The attractor framework says that **persistence under disturbance** is the basic mark of reality.
To turn this idea into a formal science, we introduce the **persistence functional** $P(x)$.

$P(x)$ is a single number that measures:

- How deep a state is inside an attractor basin.
- How quickly it returns after a knock.

We define $P(x)$ for three different kinds of systems:

1. **Deterministic dissipative systems** – here $P$ is linked to Lyapunov exponents and basin stability.
2. **Stochastic systems** – here $P$ is linked to escape time and quasipotential.
3. **Information-theoretic systems** – here $P$ is linked to negative free energy or mutual information.

The **recovery rate** $-\dot{P}/P$ is a universal sign of **critical slowing down** – a warning that a system is about to tip.

We also discuss limitations: resilience may depend on direction ("anisotropic"), and multiple timescales may need **vector** or **tensor** persistence. We list open mathematical problems.
This paper is a **roadmap**, not a finished theory.


---

# 1. Introduction

In the attractor framework, **persistence under disturbance** is central. But we have not had a single number to say *how*

*persistent* a state is.
The **persistence functional** $P(x)$ aims to fill that gap.

## What $P(x)$ should do:

- $P(x)>0$ for states inside an attractor basin.
- For a **conservative attractor** (like a free electron), $P$ is maximal (normalised to 1).
- For a **dissipative attractor**, $P$ drops after a disturbance and then recovers.
  The recovery rate $-\dot{P}/P$ equals:
  - the negative of the largest Lyapunov exponent (for deterministic systems)
  - the inverse return time (for stochastic systems)
  - the rate of information loss (for informational systems)
- $P$ falls as the system approaches a **bifurcation**, giving early warning.

We do **not** give one universal formula. Instead, we give a **family** of definitions, each suited to a different type of system, all united by the same purpose — measuring resilience.

---

# 2. Deterministic Dissipative Systems

Consider a smooth system $\dot{x}=f(x)$ with a stable attractor $A$ and its basin $B(A)$.

A natural candidate for $P(x)$ uses a **Lyapunov function** $V(x)$ — a kind of energy that always decreases inside the basin ($\dot{V}<0$).

We define:

$$P(x) = 1 - \frac{V(x) - V_A}{V_{max} - V_A}$$

This gives $P=1$ on the attractor and $P \to 0$ at the basin boundary.
Near the attractor, the recovery rate is related to the **largest Lyapunov exponent** $\lambda_1$: $-\dot{P}/P \approx -\lambda_1$

When the system approaches a tipping point, $\lambda_1 \to 0^-$, so the recovery rate slows down — this is **critical slowing down**.

**Conclusion:** For deterministic systems, $P$ can be built from a Lyapunov function. The recovery rate equals the negative of the largest Lyapunov exponent.

---

# 3. Stochastic Systems

When noise is present, persistence is about how long it takes to escape from the basin.
The **mean first passage time** $\tau(x)$ — the average time to leave — is a natural measure.

We define:

$$P(x) = \frac{\tau(x)}{\tau_{max}}$$

where $\tau_{max}$ is the value at the attractor.

For weak noise, $\tau(x)$ grows exponentially with the **quasipotential** $U(x)$ (Freidlin–Wentzell theory):

$$\tau(x) \sim e^{U(x)/\epsilon}$$

So:

$$P(x) \propto e^{-(U_{max} - U(x))/\epsilon}$$

The recovery rate is the inverse of the return time. As a tipping point is approached, the return time diverges, and the recovery rate goes to zero. This again gives **critical slowing down** — rising variance and autocorrelation.

**Conclusion:** For stochastic systems, $P$ is proportional to the mean exit time (or the exponential of the quasipotential). This connects persistence to large deviation theory.

---

# 4. Information-Theoretic Systems

For systems where information matters (neural, cognitive, social), we can define persistence using **mutual information** between past and future.

Let $I_{\text{past,future}}$ be the **predictive information**. Then:

$$P(t) = I(\text{past};\text{future at time } t) \quad \text{or} \quad P = e^{-\text{surprisal}}$$

The decay of $P(t)$ over time measures **memory loss**. Landauer's principle connects information loss to entropy production:

$$\dot{P}/P \le -\frac{\dot{S}}{k_B \ln 2}$$

Alternatively, in the **free energy principle** (Friston), the negative free energy $-F$ acts like a Lyapunov function. We can set:

$$P = e^{-F/kT} \quad \text{or} \quad P = -F$$

Then $-\dot{P}/P$ is the rate of free energy minimisation, which slows near bifurcations.

**Conclusion:** For information-theoretic systems, $P$ can be defined via mutual information decay or negative free energy, linking persistence to entropy production and predictive coding.

---

# 5. Unifying Recovery Rate and

## Critical Slowing Down

Across all types of systems, the **recovery rate** $\lambda_{rec} = -\dot{P}/P$ (just after a small disturbance) is a universal indicator:

- **Deterministic dissipative:** $\lambda_{rec} = |\lambda_1|$ (absolute value of the largest Lyapunov exponent)
- **Stochastic:** $\lambda_{rec}$ = inverse of the return time, related to the quasipotential's curvature
- **Information-theoretic:** $\lambda_{rec}$ = rate of free energy minimisation or information loss

As the system approaches a bifurcation, $\lambda_{rec} \to 0$. This is **critical slowing down**.
It shows up as rising lag-1 autocorrelation and variance (Scheffer et al., 2009).
So $P$ and its recovery rate give early warnings.

---

# 6. Normalisation for Conservative Attractors

For a perfect **conservative attractor** (e.g., an electron in its ground state, no decay), the persistence functional should be constant and maximal: $P_{cons} = 1 \text{ for all times}$

No recovery rate is defined (or it is zero). This anchors the scale.

For **emergent approximate conservative systems** (like atomic clocks), $P$ is very close to 1 and decays extremely slowly.

---

# 7. Limitations — Scalar Collapse and Anisotropic Resilience

A single scalar $P(x)$ may not be enough for systems where resilience is **anisotropic** — that is, recovery speed depends on the direction of the perturbation.
High-dimensional systems can have **multiple timescales** (fast and slow modes). A scalar average can miss important structure.

Future work may need:

- **Vector persistence** — a list of recovery rates along different directions.
- **Tensor persistence** — a metric that captures the full shape of the basin.
- **Persistence manifold** — the geometry of the basin in state space.

We accept this limitation. The scalar $P$ is a useful first approximation for systems with isotropic resilience or for early-warning applications where a single number is enough. For complex systems, a multidimensional generalisation is an open research problem.

---

# 8. Open Mathematical Problems

1. **Derive $P(x)$ from first principles** for a given class of systems (e.g., from a variational principle).
2. **Prove that $-\dot{P}/P = \langle \lambda_1 \rangle$** for a wide class of

dissipative systems.

3. **Extend the definition to systems with multiple attractors and chaotic basins** (where basin stability is fractal).
4. **Establish a rigorous relationship between $P$ and the mutual information decay rate** for non-equilibrium processes.
5. **Formulate a universal persistence functional** that works across all regimes — or prove it's impossible.
6. **Test the predictive power of $P$** in controlled experiments (e.g., ecological microcosms, neural cultures, social media sentiment).
7. **Develop vector/tensor persistence** for anisotropic resilience.

---

# 9. Conclusion

The persistence functional $P(x)$ gives a mathematical language for attractor resilience.

We have given **operational definitions** for three regimes:

- **Deterministic dissipative** → Lyapunov / basin stability
- **Stochastic** → escape time / quasipotential
- **Information-theoretic** → mutual information / free energy

The **recovery rate** $-\dot{P}/P$ unifies critical slowing down across all these domains.

We have explicitly noted **limitations** (scalar collapse, anisotropy) as open problems.

This paper is a **roadmap**, not a final theory. The framework now has a quantitative step.

**Suggested citation:** Galida, R. S. (2026). *The Persistence Functional: Towards a Mathematical Measure of Attractor Resilience (Reader-Friendly Version).* Fantasy Attractor.