

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.