Working draft · label: proposed · not peer reviewed · not a preprint · drafted with AI assistance. Source: risk.neverthought.org/research/

Declared measurements for decision control: the Epistemixer, as a protocol

Draft v0.1 · label: proposed (Epistemixer, paper 2 of the Neverthought programme)

Mebs A. Loghdey

Programme Lead, Meaning 2.0 · The Curvature Lab, Neverthought AI

Working draft, 25 September 2026.

> **Status of this draft.** Written with AI assistance. It must be read, corrected and owned by its human author before any posting, and AI assistance must be disclosed as the target venue requires. **No Epistemixer function is implemented, and no test of it has been run.** This draft is a protocol: it specifies an instrument and the experiment that could show it does not help. It reports no result.

Abstract

Systems that promote, retain or discard evidence under constraints are usually tuned by feel or by one headline metric. We specify the *Epistemixer*, a set of three declared, domain-specific measurements of a proposed transition:

- *relevance*: does the evidence bear on this action and purpose?
- *generativity*: which reachable or testable states does the transition create?
- *circulation*: does a feedback cycle amplify unresolved discrepancy?

Each returns PASS, FAIL or UNMEASURABLE, together with its inputs, method version and uncertainty, and informs a named policy rather than deciding by itself. An earlier formulation described these as the gradient, divergence and curl of a field over organisational states. We record that formulation as a metaphor with no operational content and do not rely on it. We propose a preregistered experiment that compares a controller using the declared measurements with a deterministic naive baseline under equal storage budgets. Metrics and baseline are fixed before any result is seen, and we state in advance what result would lead us to withdraw the claim that declared measurements help.

1. Introduction

When an automated process decides which evidence to keep, which claims to promote and which to discard, the decision rule is itself a consequential choice. It is often opaque: tuned until the output looks right, or optimised against a single number. Our question is narrow:

> Do declared measurements improve decisions over registered baselines?

Declared means that each measurement names its inputs, its method and version, its threshold, and what it reports when it cannot measure. *Registered baseline* means a comparison rule fixed before the experiment, not chosen after seeing which comparison flatters the instrument.

2. From metaphor to declared measurement

The instrument began as a metaphor: a field φ over states, with relevance as the alignment of a transition with the gradient ∇φ, generativity as the divergence ∇·**F** of a flow **F**, and resilience as its curl ‖∇×**F**‖. None of these is defined for organisations or software systems. There is no φ and no **F**, and nothing licenses vector calculus over arbitrary state spaces. An internal specification audit tags each of these statements as metaphor, and tags the claim that the kernel has a notion of gradient-based relevance as contradicted. We therefore retain only the three questions and require each to be answered by a declared, domain-specific function. Divergence- and curl-like operators belong only to fixtures that actually supply the geometry.

3. The instrument

For a domain *d* and a proposed transition *x*:

$\mathcal{E}_d(x) = (R_d(x), G_d(x), K_d(x))$

Function	Question	Required declaration
R_d relevance	Does the evidence bear on this proposed action and purpose?	Evidence-to-claim test, scope, threshold
G_d generativity	What additional reachable or testable states does the transition create?	State model, horizon, baseline, threshold
K_d circulation	Does a feedback cycle amplify unresolved discrepancy?	Cycle model, measurement, bound

Outputs. Each function returns PASS, FAIL or UNMEASURABLE, with its inputs, method version and uncertainty. UNMEASURABLE never becomes PASS. The outputs inform a *named* policy, which may weigh them. The instrument never overrides a gate's HOLD and never grants authority.

Restrictions.

- Positive generativity is not a universal condition for permission. Some warranted transitions close options on purpose.
- A measured PASS is not, by itself, a validated tuning criterion. Whether acting on the measurements helps is the experiment's question.
- A score from an external retrieval or linking component may be one *input* to R_d. G_d and K_d cannot inherit it.

Architecture (proposed, not built). The instrument sits as a domain controller between a storage and retrieval adapter below it and a warrant gate above it. The controller proposes

promotion, retention or discard; the gate decides whether any resulting crossing is warranted. Neither the controller nor the adapter is implemented.

4. Proposed experiment

This protocol is written for review before any data exists, in the manner of a registered report [1].

Setting. A synthetic evidence stream in one declared domain, containing eligible and ineligible evidence items, with a storage budget smaller than the stream.

Conditions.

- *Baseline:* a deterministic naive rule, fixed in advance (for example, retain the most recent items up to the budget).
- *Instrument:* the same budget, with promotion and retention decided by a named policy that uses R_d, G_d and K_d.

Procedure. Ten scripted iterations per condition over the same stream, with the same random seeds. The stream, seeds, rules and metrics are published before either condition is run.

Primary metrics (fixed in advance):

1. Eligible evidence retained at the end of each iteration.
2. Wrongful promotions: ineligible items promoted to claims.

Secondary. The share of UNMEASURABLE outputs, reported rather than hidden, and whether any decision would have changed had UNMEASURABLE been treated as PASS.

Hypothesis. Under equal storage budgets, the instrument condition retains at least as much eligible evidence as the baseline, with fewer wrongful promotions.

Stopping and reporting. All ten iterations are run and reported. If the instrument does not beat the baseline on both primary metrics, we publish that result and withdraw the claim that declared measurements help in this setting.

Ceiling. Ten scripted iterations in a synthetic stream can establish only a bounded fixture result, not field validation. A positive result would justify a second study with real evidence streams and an independent operator, not a claim of general usefulness.

5. Relation to the rest of the programme

The instrument's discipline is shared with the warrant gate described in the companion Coherence Gate paper: missing inputs never pass, every output carries what produced it, and nothing is combined into a single score. A separate test of whether declared records help human decisions is the reconstruction study, in which claims professionals answer questions from plain logs or from receipts. That study tests receipts, not the Epistemixer, and the two results must not be conflated.

6. What this paper does not claim

- That any Epistemixer function is implemented, or that a tuning loop exists.
- That vector-calculus operators apply to organisations, software systems or decisions.
- That positive generativity is good in general.
- That a measured PASS validates a tuning criterion.
- Any performance figure for any retrieval or storage component.

7. Limitations

- The experiment is synthetic, small and self-designed. The baseline is deliberately simple, so beating it is a low bar, and failing to beat it would be informative.
- R_d, G_d and K_d must be specified per domain. The protocol tests one domain's specification, not the instrument in general.

References

1. C. D. Chambers, "Registered Reports: A new publishing initiative at Cortex," *Cortex* 49:609–610, 2013. doi:10.1016/j.cortex.2012.12.016.

Checklist before a preprint

☐ Author reads and rewrites, and confirms the AI-assistance disclosure for the target venue.
☐ Fix the domain, the stream, the baseline rule, the seeds and the metrics in a public preregistration.
☐ Specify R_d, G_d and K_d for that domain, including every threshold and what each reports when it cannot measure.
☐ Add related work on evidence retention and decision rules under budget, from verified sources only.