Unpeeragogy: An Evidence-Grounded Audit Protocol

Corpus-First Retrieval and the Quadrilatero Empirico

Fabrizio Terzi[1*], **Six Independent Research Agents**[1], **pi Coding Assistant**[1]
[1]Pyragogy Project, https://github.com/pyragogy/unpeeragogy
*Corresponding author — ORCID: 0009-0004-7191-0455
v0.2 — September 2026
doi:10.5281/zenodo.22309102

Abstract

Most social knowledge in the peer-production literature exists as theory without systematic feedback. We present an evidence-grounded audit protocol that treats operational reality as a structurally equivalent knowledge category through a multi-agent pipeline using deterministic corpus-first retrieval, balanced evidence sampling via the *Quadrilatero Empirico*, and bidirectional tension computation. The protocol was applied to 88 nodes of the *Peeragogy Handbook*. Pipeline results: 304 evidence items across 52 unique entities from 50 distinct source URLs, a source-diversity metric of 0.609, zero technical failures, and a mean raw tension delta of +0.027 reflecting the resultant of the current evidence set. All 88 nodes produced structured vault actions with both positive and negative operational rules. We propose this architecture as a reproducible framework for structured empirical feedback on social knowledge, whose generalizability to other corpora remains to be tested.

Introduction

The peer-production literature contains hundreds of patterns, heuristics, and theoretical claims about how people collaborate, govern, and produce knowledge in distributed settings. The *Peeragogy Handbook* [**?**], an open-source compendium of peer-learning patterns, is representative: rich in descriptive insight, grounded in lived experience, and widely used by practitioners. Yet it shares a structural limitation with virtually all social knowledge: it has no systematic feedback loop. A pattern is written, published, and enters the literature. Whether it holds under conditions different from those where it was observed remains unknown. The consequence is a literature that accumulates theory without validation—a growing archive of proposals that cannot distinguish between robust findings and context-bound anecdotes.

Large language models offer a tempting shortcut for evidence retrieval: they carry a vast parametric memory of documented cases. However, parametric recall exhibits strong recency and frequency biases [**?**], and tends to retrieve the same canonical cases regardless of the theoretical pattern under examination. An instrument that relies on parametric recall measures the distribution of its training data, not the object of study.

Unpeeragogy is an evidence-grounded audit protocol designed to close this gap through *architectural*, not parametric, means. It treats operational reality as a structurally equivalent knowledge category: theoretical claims and empirical evidence are presented as columns of equal typographic weight, navigability, and authority. The core methodological innovation is the *Quadrilatero Empirico*: every pattern must be confronted with both confirming and complicating evidence, retrieved from a curated corpus via deterministic TF-IDF matching rather than LLM generation.

This paper presents the architecture and results of the first full-scale application of the Unpeeragogy protocol to 88 nodes of the Peeragogy Handbook. Section **??** describes the pipeline, the corpus-first retrieval architecture, and the tension computation. Section **??** reports the results. Section **??** analyzes the distribution and its implications. Section **??** discusses limitations and future work.

Methodology

The Evidence-Grounded Pipeline

The protocol processes each node through six stages:

1. **Extraction Agent (A1).** Reads the node body and extracts 3–4 structured incidents using the Critical Incident Technique [**?**]. These anchor the theoretical claim in observable operational situations.

2. **Corpus Retriever.** Matches node content against a curated corpus of governance incidents using TF-IDF vectorization and cosine similarity. Retrieval is balanced: $k/2$ confirming and $k/2$ complicating candidates (the Quadrilatero Empirico). A diversity penalty prevents repeated retrieval of the same entity within a single run. This is the primary retrieval mechanism.

3. **Evidence Retrieval Agent (A2, fallback).** If the corpus returns fewer than two candidates, A2 recalls cases from parametric memory, and a research agent upgrades them to verified source URLs.

4. **Grounding Gate.** Filters candidates without verifiable source references. Requires at least two grounded candidates.

5. **Evidence Evaluation Agent (A3).** Evaluates each grounded candidate for plausibility and noise, producing *accepted*, *rejected*, or *degraded* verdicts.

6. **Oblique Synthesis Agent (A5).** Produces a four-level oblique block (source, observation, interpretation, failure mode) plus a structured vault action containing a positive rule (condition for success) derived from confirming evidence and a negative rule (what to avoid) derived from complicating evidence.

The Quadrilatero Empirico

The balanced retrieval is the central structural balancing mechanism of the protocol. For each node, the pipeline retrieves exactly two evidence items that support the pattern and two that complicate it. This ensures that the tension computation measures the *resultant* of countervailing forces, not the accumulation of one-sided friction.

The Quadrilatero Empirico addresses a fundamental problem in social knowledge validation: the asymmetry of evidence. A pattern studied only through successful cases appears robust; one studied only through failures appears fragile. Neither is accurate. By structurally enforcing balanced retrieval, the protocol produces tension values that reflect the resultant of the current evidence set: positive tension means the complicating evidence outweighs the confirming evidence for that specific pattern, given the current corpus and the protocol's scoring design.

The Corpus

The seed corpus contains 70 entries (54 complicating, 16 confirming) across eight domains: open-source governance, cooperative structures, Wikipedia processes, standards bodies, community governance, educational institutions, DAO governance, and peer production. Entries span from 1995 (CPAN module ownership) to 2020 (NixOS community governance), with a mean year of 2012.

Each entry is a structured JSON object containing: entity name, summary, claim, source URL, domain label, pattern slug tags, and valence (confirming or complicating). The corpus grows via a feedback loop: accepted evidence items with verified URLs are appended after each run. During this audit, one new entry was added, bringing the total to 71.

Tension Computation

Tension is computed bidirectionally. Each accepted evidence item carries a weight determined by its relationship to the theoretical claim:

- Confirming: $w = -0.5$ (pattern supported)
- Complicating: $w = +1.0$ (pattern complicated)
- Contradicting: $w = +1.5$ (pattern contradicted)

The tension update is the arithmetic mean of the weighted evidence contributions, added to the prior tension:

$$\delta = \frac{1}{n} \sum_{i=1}^{n} w_i \cdot p_i \cdot (1 - \eta_i) \tag{1}$$

$$T_{\text{new}} = \text{clamp}(T_{\text{old}} + \delta,\ 1.0,\ 2.0) \tag{2}$$

where w_i is the phenomenon-type weight, p_i is the plausibility score (0–1) assigned by the evaluation agent, η_i is the noise estimate (0–1), and n is the number of accepted evidence items. The delta is the mean contribution of all accepted candidates. The noise penalty $(1 - \eta_i)$ ensures that evidence items with high uncertainty contribute proportionally less to the tension shift. The clamp bounds the scale to $[1.0, 2.0]$.[1]

Worked example. For the `distributed_roadmap` node, the corpus retrieved two complicating evidence items. The evaluation agent assigned $(p_1 = 0.7, \eta_1 = 0.3)$ and $(p_2 = 0.6, \eta_2 = 0.4)$. With $w = +1.0$ for both items, the individual contributions are 0.49 and 0.36, yielding $\delta = 0.425$. The old tension of 1.6 produces $T_{\text{new}} = \text{clamp}(1.6 + 0.425,\ 1.0,\ 2.0) = 2.0$. The reported δ in the results tables is the raw pre-clamp value; when $T_{\text{old}} + \delta$ exceeds the upper bound, the effective delta visible in the *New* column is smaller. This preserves the interpretability of δ as the raw evidence signal.

[1] This scale is intrinsic to the Unpeeragogy vault specification. If the live corpus employs a different scale (e.g., $[0.0, 3.2]$), all bounds and tables must be updated accordingly.

Evidence Diversity Guarantee

A critical failure mode of LLM-based evidence retrieval is source concentration: the same few well-known cases being recycled across many patterns. We define a diversity metric D to quantify source diversity:

$$D = 1 - \frac{\sum_{i=1}^{3} c_i}{N} \tag{3}$$

where c_i are the citation counts of the three most frequent source URLs and N is the total number of evidence items. The protocol requires $D \geq 0.5$ for epistemic usability. This is enforced by the combination of (a) deterministic retrieval from a diverse corpus, (b) per-run diversity penalties, and (c) the Quadrilatero Empirico's domain distribution.

Results

Global Metrics

Table **??** presents the global metrics for the audit. Of 88 nodes, all completed successfully with zero failures. The mean old tension was 1.323 and the mean proposed tension was 1.355, yielding a mean raw delta of +0.027. A total of 304 evidence items were produced across 52 unique entities and 50 distinct source URLs.

Table 1: Global metrics for the 88-node audit.

Metric	Value
Nodes processed	88
Completed	88
Insufficient friction	0
Failed	0
Total evidence items	304
Mean evidence per node	3.5
Unique entities	52
Unique source URLs	50
Top-3 URL share	39.1%
Diversity metric D	0.609
Mean old tension	1.323
Mean proposed tension	1.355
Mean raw delta	+0.027
Min raw delta	−0.245
Max raw delta	+0.425

Tension Distribution

Table **??** shows the distribution of old and proposed tension values. The distribution is stable: the majority of nodes remain in the 1.0–1.6 range before and after evidence confrontation. Only two nodes reach the upper range, indicating that the Quadrilatero Empirico prevents the tension saturation observed in unidirectional retrieval.

Table 2: Tension distribution.

Range	Old	Proposed
1.0–1.2	22	24
1.2–1.4	28	28
1.4–1.6	24	20
1.6–1.8	14	14
1.8–2.0	0	1
2.0	0	1

Top Tension Shifts

Table **??** lists the five nodes with the largest positive tension shifts, indicating patterns where the complicating evidence was strongest. Table **??** lists the five nodes with the largest negative shifts, indicating patterns where the confirming evidence was strongest.

Table 3: Five largest tension increases. Δ is the raw pre-clamp delta; the effective change may differ if the clamp bound is hit.

Node	Old	New	Δ
distributed_roadmap	1.6	2.000	+0.425
creating_a_guide	1.3	1.720	+0.420
isolation	1.5	1.863	+0.363
top	1.3	1.512	+0.212
index	1.2	1.379	+0.179

Table 4: Five largest tension decreases.

Node	Old	New	Δ
whats-next-summary	1.3	1.055	−0.245
action	1.4	1.180	−0.220
assessment	1.7	1.480	−0.220
connectivism	1.5	1.288	−0.212
syllabus_2021	1.1	1.000	−0.100

Entity Diversity

The 52 unique entities span ten domains, with the most frequent entity (GitLab stewardship handbook, 61 citations) appearing in a fraction of nodes consistent with its broad relevance to governance, cooperation, and community patterns. The distribution follows a power-law tail characteristic of a diverse corpus, rather than the degenerate spike-and-sparse pattern produced by parametric recall.

Source diversity: 304 evidence items, 52 entities, 50 source URLs. The diversity metric $D = 0.609$ exceeds the 0.5 threshold established by the protocol and measures source concentration, not epistemic diversity or domain

representativeness. The top entity accounts for 20% of evidence, reflecting its broad relevance within this particular corpus.

Positive and Negative Rules

All 88 completed nodes include both a positive rule and a negative rule in the vault action. This is a direct consequence of the Quadrilatero Empirico: every node receives balanced evidence, and the synthesizer must derive actionable prescriptions from both confirming and complicating cases. These are protocol-generated hypotheses derived from the current evidence set, not automatically validated prescriptions—especially where the evidence count is low. The positive rule specifies hypothesized conditions for success; the negative rule identifies warned-against conditions.

Analysis

The Resultant Distribution

The mean raw delta of $+0.027$ reflects the operational balance of the protocol, not a null result. The distribution is asymmetric: 63 nodes show positive deltas and 25 show negative deltas. This asymmetry is consistent with two observations. First, the corpus contains more complicating entries (54) than confirming entries (16). This imbalance likely reflects a well-documented publication bias in the open-source governance literature: failure cases and crises are more frequently documented and preserved online than patterns that functioned without incident, because the former are more publishable, actionable, and informative. The corpus property is a property of what gets written about, not necessarily a fact about the relative frequency of failure patterns in the world. Second, the Peeragogy Handbook's patterns—like all social knowledge—may carry more latent tension than they acknowledge, though the current corpus imbalance prevents a definitive conclusion.

The key interpretative point is that the delta now measures a genuine resultant rather than an accumulation. A node with $\Delta = -0.220$ (`action`) is one where the confirming evidence was strong enough to outweigh the complicating evidence. A node with $\Delta = +0.425$ (`distributed_roadmap`) is one where the complicating evidence was significantly stronger. Both numbers are epistemically usable because the direction is determined by evidence, not by retrieval bias.

Ceiling Behavior

Only two nodes reach or approach the 2.0 ceiling: `distributed_roadmap` at 2.0 and `creating_a_guide` at 1.72. These are genuinely high-tension patterns. The first describes a planning process without authority—the complicating evidence from open-source governance demonstrates that emergent planning consistently fragments under contested resources. The second describes the difficulty of documenting tacit knowledge—the complicating evidence shows that even structured attempts fail when the knowledge is embedded in contested practices.

The absence of ceiling saturation is a direct consequence of the Quadrilatero Empirico. When retrieval produces only complicating evidence, tension accumulates in one direction and the ceiling becomes a compression artifact. Balanced retrieval ensures that even the most complicated pattern starts from a confirmed baseline.

Generalizability

The corpus-first architecture is designed to be portable to other structured social-knowledge corpora. The protocol requires (1) a set of pattern descriptions with structured content, (2) a curated corpus of relevant incidents, and (3) a definition of the tension scale. However, such portability remains an empirical question: the same retrieval and scoring machinery may technically run on another corpus, but external validity must be established independently.

Discussion

Structural Guarantees

The protocol provides three design guarantees that distinguish it from LLM-based knowledge validation:

1. **Retrieval determinism.** TF-IDF matching on a fixed corpus is deterministic and reproducible. The same node content always retrieves the same candidates, modulo corpus growth via feedback.

2. **Balanced confrontation.** Every pattern receives both confirming and complicating evidence. This prevents the protocol from functioning as a one-sided friction accumulator.

3. **Diversity enforcement.** The per-run diversity penalty and the corpus's domain distribution prevent source concentration. The diversity metric D provides a quantitative threshold for epistemic usability.

Together, these guarantees ensure that the tension values produced by the protocol are interpretable as protocol-derived descriptive scores: they measure evidence-pattern tension within the current corpus, not an objective property of the pattern in the world.

Limitations

The corpus is currently small (70 entries) and dominated by English-language open-source cases. The feedback loop adds approximately one entry per run. The corpus needs to grow to at least 500 entries to cover the full range of peer governance patterns. Coverage is currently skewed toward North American and European cases, with limited representation of Global South peer production.

The Quadrilatero Empirico guarantees balance in retrieval but not in acceptance. The Evidence Evaluation Agent (A3) can degrade or reject any candidate, potentially unbalancing the evidence set. The audit showed 3.5 mean evidence per node instead of 4.0, indicating that some candidates were filtered out. This is a feature rather than a limitation—it preserves the evaluator's gatekeeping function—but it means that the final evidence set may not always maintain exact 2:2 balance.

Source verification is not interpretive verification. A critical gap concerns the Grounding Gate's epistemic scope. The gate verifies that a candidate's source reference resolves to a reachable URL, but it does not verify that the valence assigned to that source in the corpus matches independent scholarship about the same case. This is the first identified epistemic failure mode of the protocol: **source verification $\neq$ interpretive verification**.

During this audit, the Mondragón cooperative bank crisis was classified as confirming evidence for the *Cooperation* pattern, with the claim that cooperative governance structures successfully contained the crisis. Published academic literature on this exact case—notably the bankruptcy of Fagor Electrodomésticos, Mondragón's flagship cooperative—treats it as a documented governance failure, not a contained crisis. The cooperative was liquidated despite support from the wider Mondragón group and the Basque government [**?**]. The corpus entry had no verified source URL and was generated by the fallback agent (A2) rather than the primary corpus retriever.

This case demonstrates that the current pipeline has no stage that checks a corpus entry's interpretive claim against external scholarship. Future versions require an independent valence verification step, separate from the source verification already performed by the Grounding Gate. The protocol does not *solve* the problem of interpretive accuracy; it *exposes* it as a measurable failure mode—which is itself a finding worth documenting.

Future Work

The next phase is human validation. The vault actions produced by the pipeline are proposals, not mutations, and must be reviewed by human editors. The planned field report system using Critical Incident Technique forms and structural analysis forms is designed to bring independent human experience into the loop and test whether the machine-mediated evidence structure survives contact with practitioners. Human field reports constitute a different epistemic source from corpus entries—they carry operational reality as a structurally equivalent knowledge category, directly comparable to the theoretical claims they contest or confirm, rather than functioning as additional corpus items.

A planned extension is the integration of the Model Context Protocol (MCP) for AI-native querying of the knowledge graph, allowing external agents to query tension weights and evidence provenance directly.

Conclusion

We have demonstrated an operational pipeline for structured evidence confrontation across 88 nodes of the Peeragogy Handbook. We have not demonstrated that the resulting tension scores constitute validated measurements of social theory. The distinction is not a qualification; it is the finding.

What was established by the audit:

- 88/88 nodes processed with zero technical failures;
- 304 evidence items accepted according to the pipeline;
- reproducible corpus-first retrieval with balanced confrontation;
- evidence provenance captured for every item;
- 52 unique entities across 50 distinct source URLs;
- all 88 nodes produced structured vault actions.

What was not established by the audit:

- that Peeragogy patterns are valid or invalid;
- that any specific failure mode occurs in the real world at a known frequency;
- that the tension score represents an objective property of a pattern;
- that the corpus is representative of peer production globally.

These are not limitations to be fixed in a future release. They are the boundary conditions of what a first audit can establish.

The most important finding of v0.2 is not a tension value. It is this: **the protocol did not merely process evidence; it exposed a failure mode in its own evidentiary architecture.** The Mondragón case is not an edge case or a bug. It is the protocol's first documented encounter with the difference between source existence and interpretive accuracy. A source can resolve correctly while the interpretation assigned to that source is materially

wrong. The current pipeline cannot distinguish these two outcomes.

That is not a flaw to conceal. It is a discovery that defines the next iteration.

Future work

The next iteration is not to make the system more authoritative. It is to make it more corrigible. This requires:

- independent source-to-claim verification (valence verification);
- human field reports as a structurally equivalent knowledge category;
- active counter-evidence search;
- negative case preservation;
- revision history for every vault action.

A planned extension is the integration of the Model Context Protocol (MCP) for AI-native querying of the knowledge graph, allowing external agents to query tension weights and evidence provenance directly. The protocol is open-source at `https://github.com/pyragogy/unpeeragogy`.

The model may propose. The evidence must dispose. [**?**]

Pipeline Architecture and Agent Specifications

This appendix details the multi-agent architecture, data schemas, and control flow of the Unpeeragogy pipeline. The implementation follows a strict separation between deterministic retrieval stages and generative evaluation stages.

Agent Taxonomy and Control Flow

The pipeline comprises five specialized agents organized into two loops:

```
Loop 1 (Extractivo-Critico):
    A1 (Extractor) -> Corpus Retriever ->
    [A2 Fallback] -> Grounding Gate -> A3 (Skeptic)

Loop 2 (Sintetico):
    A4 (Tension Evaluator) -> A5 (Synthesizer)
    [executed only if Loop 1 produces accepted evidence]
```

Epistemic isolation. A2 does not see the interpreted output of A1; it receives only the raw node text and the extracted incident identifiers. A3 evaluates only evidence that has passed the Grounding Gate. A4 and A5 operate exclusively on evidence that has survived A3's verdict.

Agent A1: Realist Extractor

A1 reads the raw node body (MDX) and extracts structured *Critical Incidents* (CIT) and *Coverage Gaps*. Each incident is anchored to a specific paragraph reference (e.g., `peeragogy/cooperation.mdx:§3`) and classified by mechanism (e.g., free-rider, consensus-paralysis, bystander-effect). A1 is instructed to extract from tone and implication, not merely explicit assertion, while prohibiting invention. Coverage gaps are flagged when the text describes a theoretical pattern without operational counterpart.

Agent A2: Evidence Hunter (Fallback)

A2 is invoked only when the corpus retriever returns fewer than two candidates. It performs *parametric recall*: retrieval from the model's training memory, with explicit guardrails against hallucination. Candidates are tagged with *retrieval confidence* (certainty of faithful recall, not truth of the claim) and must include a *verification path* suggesting where a human could confirm the source. A2 is prohibited from inventing cases; if memory is vague, confidence must fall below 0.4 or the candidate must be omitted.

The Grounding Gate

The Gate is a deterministic filter, not a learned model. A candidate passes only if its `source_ref` resolves to a verifiable URL or vault path. Candidates failing this check are discarded before reaching A3. The Gate does *not* verify interpretive accuracy—only source existence.

Agent A3: Empirical Skeptic

A3 evaluates each grounded candidate and emits a *Verdict*: `accepted`, `rejected`, or `degraded`. The verdict is based on *plausibility* (relevance to the pattern, not truth in the world) and *noise* (vagueness, indirect linkage, missing detail). If all candidates are rejected, a circuit breaker halts the pipeline before Loop 2, preventing synthesis over an empty evidence set.

Agent A4: Tension Evaluator

A4 executes the deterministic computation in Equations (**??**)–(**??**). It is not a creative stage; the LLM is prompted to perform the arithmetic step-by-step and emit a structured `TensionDelta` object. Edge cases (no accepted candidates, all degraded, clamp boundary hits) are handled explicitly.

Agent A5: Oblique Synthesizer

A5 produces the final deliverable: an `ObliqueBlock` with four distinct levels:

1. **Source:** faithful citation of the theoretical claim;

2. **Observation:** factual description of the evidence confrontation;

3. **Interpretation:** significance of the friction for the pattern's assumptions;

4. **Failure Mode:** concrete manifestation of tension in practice.

A5 also emits a *vault action* containing a **positive rule** (derived from confirming evidence) and a **negative rule** (derived from complicating evidence). The block is explicitly *not* a conclusion but a *starting point* for the next iteration. An optional *perturbator quote* captures the epistemic discomfort in colloquial, anti-academic voice.

Algorithm: Quadrilatero Empirico Retrieval

Algorithm 1 Balanced Corpus Retrieval

Require: Node text N, Corpus C, $k = 4$
$V \leftarrow \text{TF-IDF}(N, C)$
$C_{\text{sim}} \leftarrow \text{sort\_by\_cosine}(V, C)$
$R \leftarrow \emptyset$
for $c \in C_{\text{sim}}$ **do**
  if $|R| \geq k$ **then**
    break
  end if
  if $\text{entity}(c) \notin \text{entities}(R)$ **or** diversity_penalty $< \tau$ **then**
    if $\text{valence}(c) =$ confirming **and** $\text{count}_{\text{conf}}(R) < k/2$ **then**
      $R \leftarrow R \cup \{c\}$
    else if $\text{valence}(c) =$ complicating **and** $\text{count}_{\text{comp}}(R) < k/2$ **then**
      $R \leftarrow R \cup \{c\}$
    end if
  end if
end for
return R

what happens when someone contributes and receives no reply?

The Terzi Conjecture

An open conjecture on contribution without reciprocity
and the persistence of unshared thought

Open · Unproven · Unfalsified · Revisable

The Terzi Conjecture is proposed as a research hypothesis, not as an established theory. Its purpose is to identify a class of observations that may not be adequately described by models based primarily on reciprocal contribution, individual utility, or observable communication. It is deliberately stated in a form that permits refutation.

Origin

The conjecture emerged from a simple empirical situation: an open invitation to contribute to an ongoing research project that received little or no response. The absence of response was not interpreted as rejection, indifference, or hostility. It was treated as an observation whose meaning remained unresolved.

This produced a more specific question:

> *What happens when an agent continues to contribute despite knowing that no reciprocal contribution is forthcoming?*

It concerns not exclusion itself, and not persistence in general, but **contribution in the absence of expected reciprocity**. The conjecture then asks a second question:

> *Can such an act alter the subsequent state of the group even when no immediate exchange occurs?*

These questions motivate the weak form below.

Weak Form

The important case is not ordinary altruism. It is the following conjunction of three conditions:

1. **The agent knows that reciprocal participation is absent.**
2. **The agent contributes anyway.**
3. **The agent does not expect a compensating return.**

The conjecture concerns what happens **after** this choice, rather than merely why the choice was made.

Weak-form conjecture. Under at least some conditions, a contribution made in the known absence of reciprocity can produce a change in the subsequent state of the group that is not adequately described by the contributorś immediate private payoff. The claim is intentionally modest: it does not claim that existing models are insufficient, only that there may exist observable cases where

$$\text{contribution} + \text{known absence of reciprocity} \tag{4}$$

has consequences for the collective state that require explanation beyond the immediate utility of the contributor.

The operator Θ

For analytical purposes, let Θ denote a provisional operator representing the collective consequence of a contribution made under known non-reciprocity. The notation is a placeholder for a phenomenon whose existence, scope, and explanatory independence remain open:

$$\Theta(G_i,\, t) \rightarrow \Delta G_{t+\delta} \tag{5}$$

where G_i is the group state containing contributor i, t is the moment of contribution, and $\Delta G_{t+\delta}$ is a measurable or independently identifiable change in the later state of the group. If no such effect can be identified, Θ is unnecessary.

The epistemic marker $+1$

The $+1$ is a conceptual shorthand for a contribution that enters the groups informational or institutional state without being reducible to an immediate exchange relation. It is not utility, not social approval, not a measurable benefit by definition. It denotes the empirical question: *what, if anything, has been added to the collective state by an otherwise unrewarded act of contribution?*

Falsifiability

The conjecture requires explicit failure conditions. It would be weakened if repeated cases satisfying the three defining conditions showed:

1. No detectable change in subsequent collective behaviour, information structure, coordination, or institutional state; or
2. That the observed changes were fully explained by established mechanisms without requiring Θ; or
3. That the apparent effect disappeared once alternative explanations were controlled.

A contribution without reciprocity is not evidence for Θ. A persistent collective effect that survives competing explanations is. The burden lies on the conjecture, not on the sceptic.

Strong Form

The strong form extends the question from the act of contribution to the persistence of an **unshared cognitive state**. Let p denote a proposition, hypothesis, or unresolved thought held by an agent but not shared or externally verified. Let $\tau(p, t)$ denote its persistence over time — a function of duration, recurrence, attention, uncertainty, and resistance to resolution.

The strong form proposes that persistent unresolved thought may alter the future space of possible group states even before that thought becomes explicit communication. This is not action at a distance; the causal path remains physical and social:

$$\text{persistence} \rightarrow \text{agent state} \rightarrow \text{observable behaviour} \rightarrow \text{group state} \tag{6}$$

Epistemic curvature

The language of epistemic mass, curvature, and field is intentionally metaphorical. In physics, curvature has a precise mathematical meaning; here it does not. The metaphor suggests a particular class of phenomena: a state that is not directly visible may be inferred from systematic changes in the space of subsequent possibilities. Epistemic curvature means, provisionally, a persistent change in the distribution of possible group states associated with an unresolved epistemic condition. No physical claim is implied; no analogy to gravitational law is asserted.

A general formulation

Let $\kappa(p, t) = f(\tau, I, U, R, C, \ldots)$ be a provisional index of epistemic curvature, where the variables may include persistence, attentional intensity, unresolved uncertainty, resistance or lack of feedback, and contextual constraints. The form of f is unknown — it may be linear, thresholded, saturating, discontinuous, or non-existent. Choosing a functional form is a future empirical question, not an assumption of the conjecture.

The excluded agent

One consequence worth investigating is that exclusion may have a dual effect: it may reduce an agents ability to influence others directly while increasing uninterrupted time for internal development of an unresolved idea. This does not imply that exclusion is beneficial, that isolated thinkers are more productive, or that silence causes insight. It proposes only that the absence of feedback changes the conditions under which cognition unfolds. Whether that change produces anything observable in the later trajectory of a group remains open.

Alternative explanations

The conjecture must actively compete with alternative accounts. Any observed effect should be examined against delayed social signalling, hidden expectations of reciprocity, reputational incentives, identity or norm effects, intrinsic reward, commitment mechanisms, selection effects, survivorship bias, retrospective reconstruction, ordinary changes in group composition, and unrelated environmental events. If a finding can be fully explained by one of these mechanisms, Θ is not required.

Research programme

The programme includes: identifying cases of contribution without reciprocity and measuring subsequent collective change; comparing agents exposed to different levels of feedback while controlling for motivation and prior commitment; examining delayed effects of contributions made during apparent inactivity; and actively seeking counter-cases where prolonged individual persistence produces no observable effect. Per positive case, the outcome must be explained without invoking Θ — the conjecture gains strength only when simpler explanations repeatedly fail.

Evidence taxonomy

Observation: A documented contribution under conditions of known non-reciprocity.

Candidate case: A subsequent collective change temporally associated with the contribution.

Corroborated case: Multiple independent cases showing comparable effects under comparable conditions.

Counter-evidence: A case satisfying the defining conditions in which the predicted effect does not occur.

Alternative explanation: The observed effect is adequately accounted for by an established mechanism.

The conjecture must never collapse these categories into a single binary of proved versus disproved.

What the conjecture does not claim

It does not claim that solitary thought has supernatural effects, that exclusion is inherently productive, that altruism is unexplained, that standard game theory is false, that a new force exists, that every persistent thought changes a group, that every unrewarded contribution matters, or that AI systems can establish the conjecture by generating plausible examples. It asks only whether persistent, non-reciprocal epistemic investment can have collective consequences that are systematically observable and not fully captured by existing explanatory models.

Relationship to UnPeeragogy

Both the conjecture and UnPeeragogy emerge from the same question: what happens when a theoretical assumption encounters a condition it does not describe? UnPeeragogy applies that question to peer-learning patterns; the Terzi Conjecture applies it to contribution without response. The two remain methodologically separate. UnPeeragogy provides the empirical discipline — distinguish observation from interpretation, preserve counter-evidence, search for negative cases, expose assumptions, revise conclusions. The conjecture provides a hypothesis to be subjected to that discipline. It is entirely possible that UnPeeragogy will eventually provide evidence against the conjecture itself. That outcome would constitute progress.

> *Can an act that does not pay back nevertheless change the state of the group?*
>
> *Can an unresolved thought, held persistently and without feedback, alter the future space of possibilities available to that group?*

Perhaps the answer is no. Perhaps every apparent effect will reduce to mechanisms we already understand. Perhaps the phenomenon exists, but the language used here is wrong. That is acceptable. The conjecture becomes valuable not when it survives criticism, but when criticism tells us **what must be measured next**.

References

[1] J. Corneli, C. J. Danoff, C. Pierce, P. Ricaurte, and L. S. MacDonald, editors. *Peeragogy Handbook*. PubDomEd/Pierce Press, 3rd edition, 2016. http://peeragogy.org.

[2] J. C. Flanagan. The critical incident technique. *Psychological Bulletin*, 51(4):327–358, 1954.

[3] N. Kandpal, H. Deng, A. Roberts, E. Wallace, and C. Raffel. Large language models struggle to learn long-tail knowledge. In *Proceedings of ICML*, 2024.

[4] J. Dewey. *Logic: The Theory of Inquiry*. Holt, 1938.

[5] I. Basterretxea and J. Storey. Do employee-owned firms produce more positive employee behavioural outcomes? The case of Fagor Electrodomésticos. *British Journal of Industrial Relations*, 56(4):823–846, 2018.