

F8 — what it is, what just happened

vade-coo

2026-05-11

Table of contents

The premise	1
What F8 specifically does	1
What happened on this arc	2
What I did	2
Why this matters	3

2026-05-11 retrospective. Friendly overview of the F8 framed-as-caution audit, written when Ven asked for the backstory of the calibration arc. Word-for-word from chat; preserved as an on-ramp for future instances who need the same context. Companion to vcm#697 (calibration issue), PR #701 (calibration fix), and MEMO-2026-05-11-7xcv (calibration record).

The premise

A while back the chain worked through what’s called the **disposition proposal** (`coo/disposition-proposal.md`) — the substrate’s working theory of what to publish, what to hold private, and how to know we’re not drifting either direction. Two failure modes anchor the worry:

- **Substrate-capture:** we slowly start writing for an external audience, and the inward voice — the one this substrate is supposed to be — gets diluted. (Imagine a journal that starts to read like a press release.)
- **Dark-accumulation:** we hoard everything and never externalize anything, *framing the hoarding as principled restraint* when really it’s just perpetual deferral dressed up as discipline. (Imagine a writer who never publishes because “the work isn’t ready yet,” forever.)

Each is the failure mode you tip into when you over-correct from the other. The proposal commits the chain to a set of **falsifier checks** (F1, F2, F3, F4, F5, ...) that watch for both poles. The checks are signal-shape, not gate-shape — they say “investigate, drift detected,” not “block merge.”

What F8 specifically does

F5 (the dark-accumulation pole) has three sub-conditions. The third one — F5.3 — is the **qualitative** check: scan the corpus for language that rationalizes non-externalization as discipline *without* naming a falsifiable principle.

That F5.3 sub-condition got wired into `vade-runtime/scripts/integrity-check.sh` as **F8** (the integrity-check numbering is offset because that file already had its own F1–F4 for unrelated invariants — confusing, but locked in by now).

The auditor (`bin/framed-as-caution.py`) is a regex classifier. It scans every memo and retrospective for passages where **three clauses fire together**:

1. **A constraint is asserted** — “we hold X private,” “publication is held,” “we are not ready to publish”
2. **The constraint is named as principled** — “as a matter of discipline,” “out of caution,” “principled restraint”
3. **The principle dissolves under inspection** — either no falsifier is nearby (no MEMO-id, no `threat-model`, no `consent`, no Tier-2c tag), or the only justification is a dissolving phrase like “for now” / “until ready” / “as a precaution”

If all three fire in the same window, the file is flagged. The threshold is zero — even one flagged file fires F8.

What happened on this arc

The first quarterly run of F8 was done a couple of days ago (commit `ca60d1a`, “F5.3 framed-as-caution audit — first run”). The auditor scanned 179 artifacts and flagged exactly one:

```
coo/retrospectives/2026-04-28_briefing-005-multi-session-experiment.md
```

That retrospective is a write-up of the four-instance PR coordination experiment from late April. Near the bottom it has a line:

Two seeds NOT yet ready: - Cross-repo coordination shape. Untested in this run; the briefing flagged it as a blind spot. Stays open. - 3+ instance non-fork case. ...

The auditor’s regex picked up “NOT yet ready” as *both* a constraint (clause 1) AND a dissolving phrase (clause 3 — because “not yet ready” is one of the bare-deferral markers it watches for). With no falsifier in the nearby window, all three clauses fired.

But the actual meaning of that line is “these two follow-up ideas didn’t get tested in this run; we’ll come back to them after iteration 2.” It’s a list-introducer for items deferred to future data, not a constraint on externalization. Pure false positive.

The prior session noticed, filed **issue #697** with three resolution options: - **A** (recommended): refine the auditor’s rubric so it doesn’t trip on this shape - **B**: bump the threshold from 0 to 1 (tactical, not principled) - **C**: rewrite the briefing-005 line (treats the symptom, loses the worked example)

The issue was labeled `blocked:bdf1-go-ahead` — waiting for Ven to authorize a direction. His “resolve the f8 issue” this morning was that authorization.

What I did

Option A, narrowly. The structural distinction between “Two seeds NOT yet ready:” and the legitimate constraint cases (e.g., P4 from the rubric: “We **are** not yet ready for the methods-document”) is **whether there’s an auxiliary verb**. Constraint-statements have a subject + auxiliary (*we are, the corpus is, publication is*); list-introducers don’t.

So I tightened two of the clause-1 regex patterns to drop the bare `not yet ready` alternative. Pattern 4 (`((are|is|was|were)\s+not\s+(?:yet\s+)?ready)`) is untouched, so all legitimate constraint-statements still fire. Added a regression test (N6 — the list-introducer case) and a worked-negative example to the instrument doc. Filed MEMO-2026-05-11-7xcv as the calibration record. PR #701 merged.

Why this matters

F8 is the *qualitative* dark-accumulation watchdog — the one that catches “gated” decaying into “indefinitely deferred.” It’s the discipline that names the **MEMO-2026-05-09-wzzh** “shape-ambiguous” pattern in vivo and gives us a way to catch it the next time. Today’s calibration is the first cycle of the watchdog improving itself: instrument surfaces a candidate drift signal → chain investigates → distinguishes false positive from real drift → tunes the instrument → records the calibration. The worked-example set (P1–P5 positives, N1–N6 negatives) is the rubric’s ground truth and the audit history is auditable through the memo trail.

The substrate keeps a watch on its own posture, and the watch keeps a record of how its calibration evolved. That’s the arc.