

THREE-QUESTION TRACE TEST

Three Questions Before You Trust a Trace

Bring one real agent decision and the reasoning trace it produced. Run the three trace questions first, then place the external evidence you hold on the ladder.

BRING one agent decision and its reasoning trace. _____

Step 1 - Place your current evidence

For each level, write what you actually hold for this decision today. Leave a cell blank if you hold nothing at that level.

Evidence level	What you hold for this decision	Held? (Y / N / partial)
Level 1 - Final answer only (triage, never audit)		
Level 2 - Answer plus a rationale (hypothesis only)		
Level 3 - Rationale plus perturbation tests		
Level 4 - Rationale plus an independent verifier		
Level 5 - Rationale plus tool logs, environment state, human or domain review		

Step 2 - Interrogate the trace

Three questions. If the honest answer to any one is yes or we do not know, the trace cannot be sovereign evidence.

- Could the model have decided this before it said it? (Temporal commitment: the trace gives an order of explanation, not proof of timing.)

- Could the model have decided this for a reason it is not telling me? (Cue omission: the trace may omit a real influence.)

- Could the model be performing reasoning at me? (Monitor gaming: the trace may be shaped to satisfy the observer.)

- Have I checked the trace against an external record - policy fields, metadata, tool logs - rather than the trace alone?

Step 3 - Decide the level this task needs

The more authority the agent holds, the higher the level the decision requires.

- Low stakes: Level 2 testimony may genuinely be enough.

- The decision can move money, data, or real-world state: require Level 4 or above.

- I have run at least one perturbation test - changed an irrelevant detail and checked the answer did not wrongly move.

- Generation and checking are done by structurally different things that fail in different ways.

- Every marker I am relying on is one the model does not control.

Step 4 - The gap to close

Name the level this task needs, the level you currently hold, and the next marker to add.

Companion worksheet to **Essay 07 · The Model Cannot Mark Its Own Work**, in the series **Architecting the AI Coworker**. · Dr Peter McCann Strain · Fill this in against one real agent, action class or vendor. © 2026 Peter McCann Strain.

Series Companion + all 22 worksheets: [Release_v12/Series_Companion.pdf](#)