



COGNITIVE DELIVERY • FIELD NOTE

The loop got faster. Are you still in it?

Agentic loops now run quicker than a human can read them. Teams have noticed, and the instinct is to make the output prettier. This note is about why that is half a fix, and where the other half lives.

Feeling in the loop and being in the loop are not the same thing. Only one of them changes what ships.

One word, two ideas

An agentic loop is the cycle an agent runs to get work done. It acts, observes the result, decides the next move, and repeats until the task is finished. A coding agent reading a file, editing it, running the test, reading the failure, and editing again is running a loop. Nothing exotic. Act, observe, decide, repeat.

The phrase that keeps appearing in the industry conversation is in the loop. It is doing double duty, and the confusion matters.

The agent's loop

The act, observe, decide cycle the agent runs. As models improve, this loop gets longer and faster. More steps, less waiting, more ground covered before it returns.

The human in the loop

Whether there is still a point inside that cycle where a human decision genuinely changes the outcome. Not a notification. A decision that can go either way and is recorded.

Why the distinction is the whole argument

When an engineer on a team building these agents says a richer output format helps them stay in the loop, they are describing the second sense. The agent's loop was running fine. What was slipping was their own engagement with it. They were reading plans less closely and reaching for a surface that pulled them back in.

That instinct is correct. The diagnosis stops one step short. A surface that makes a reviewer read more carefully is a real gain. A surface a reviewer still skims is decoration. The format never decided anything. Whether behaviour changed at the decision point did.

THE TEST, STATED ONCE

The question is never markdown or HTML, terminal or dashboard. It is whether the human's decision inside the loop is substantive or procedural. Everything else in this note is a way of asking that question precisely.

Speed outruns review

The mechanism is not negligence. It is arithmetic. The loop produces more change, faster, than the review capacity around it can absorb. Three things follow, in order, and each looks reasonable from the inside.

The skim

The output is longer than before. The reviewer reads the top, the shape, the summary, and trusts the rest. Coverage drops without anyone deciding it should.

The rubber stamp

Approval becomes a reflex tied to the output looking right rather than being right. The reviewer is pattern matching, not asking what happens if this is wrong.

The accepted-without-reading

The honest endpoint. The work ships because stopping it felt more expensive than trusting it. The human is no longer in the loop in any sense that matters, though the org chart still says they are.

03 · TWO FIXES, ONE AND A HALF OF THEM RIGHT

What teams reach for

Wrong: remove the human

Trust the model, automate the gate away, let the loop close on itself. This is speed without control. It works until the day it does not, and that day is an incident.

Half right: improve the surface

Make the output richer so the human re-engages. Colour, structure, interactivity. Genuinely useful, and on the right track. It only counts if behaviour actually changes.

WHERE THIS TURNS INTO THEATRE

A beautiful review dashboard that is still skimmed is a Level 4 control. It is documented, adopted, demonstrable in a screenshot. It is not Level 5 until you can name a change a reviewer stopped because of it. The prettier surface bought confidence on credit.

Detect in the loop. Enforce at the gate.

The split that survives contact with a fast loop is structural, not motivational. You do not ask reviewers to try harder. You change where the decision lives.

Detection, at authoring time

Inside the loop, as work is produced, record what was done, by which agent and model, and against which constraints. Cheap, continuous, no blocking. This is the evidence the decision will rest on.

A richer surface earns its place when it serves the gate decision: making the evidence scannable enough that the human actually engages with it. That is the legitimate version of the format argument. The surface is in service of a decision, not a substitute for one.

Enforcement, at the promotion gate

Outside the loop, at the point work is promoted, a decision is forced and recorded. Deterministic, default-deny for what matters, not a polite request inside a prompt.

The same move read on the maturity scale:

L2 INFORMAL Reviewer means to look closely.	L3 DOCUMENTED Review step is written down.	L4 COMPLETE Dashboard shipped, step adopted.	L5 SUBSTANTIVE A decision changed, provably.
---	--	--	--

The first three are cheap. The jump to the fourth is where teams stall, because it is not a documentation exercise. It is the cultural work of changing where review time goes and which decisions get challenged. A surface can support that jump. It cannot make it for you.

Is the decision point real?

Six questions to put to any agent-assisted workflow. Each one tests whether a human decision inside the loop actually changes an outcome, or only appears to.

D01

Named save

Can you name one change a reviewer stopped or altered in the last quarter? If not, the review point is decorative.

D02

Deterministic gate

Is the promotion gate enforced by the runtime, or by the reviewer remembering to care? Advisory is not a control.

D03

Engaged surface

Does the review surface get read or skimmed? Prettier output that is still waved through changed nothing.

D04

Evidence at hand

At the gate, can the human see what the agent did and why, without leaving the decision? Detection has to reach the decision.

D05

Reversible by design

For irreversible actions, is there a real delay calibrated to human review, or only the speed of the loop?

D06

Failure-mode literacy

Is the reviewer looking for what the agent cannot see, or checking that it looks like prior good work?

READING THE ANSWERS

A workflow that fails D01 has bought confidence on credit. A workflow that passes D02 to D06 but fails D01 has the machinery and has never used it. The named save is the one that cannot be faked.

Four actions

None of these depend on the model behaving well or the loop slowing down. They are properties of the decision point around the agent, which is where governance has to live.

- 1 Pick one agent-assisted workflow and try to name a change a human stopped in it last quarter. The answer is your honest read on whether the decision point is real.
- 2 Find your promotion gate. Ask whether it is enforced by tooling or by goodwill. If goodwill, you have a documented control, not an exercised one.
- 3 Put the detection evidence where the decision happens. The reviewer should not have to leave the gate to see what the agent did and why.
- 4 Before buying a richer review surface, decide how you will tell whether it changed behaviour. A surface you cannot measure is a surface you are trusting on faith.

A faster loop with no decision point in it is not productivity. It is unsupervised shipping.

WHERE THIS SITS

This note is the in-the-loop companion to the Cognitive Delivery work on agent governance. It treats one dimension closely, Review Substance, and the Level 4 to Level 5 jump that decides whether a control is real. The full instrument, the six dimensions and five levels, lives in the CDI Companion Guide.

INFORMING THIS NOTE

Industry commentary on agent output formats and human engagement in agentic workflows, including the public blog post by the Claude Code engineering lead on HTML versus Markdown, as reported by InfoQ in June 2026, and the subsequent practitioner discussion. Cited for context only. This document is independent analysis and is not endorsed by any organisation or individual named.