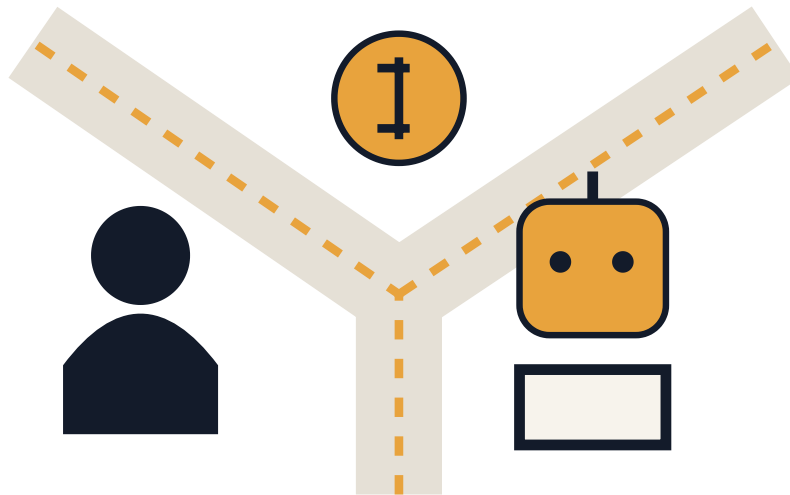


Who really made the call?

ANTHROPIC

When an AI agent leans one way, experts tend to follow without noticing

Source: Anthropic, Oct 27, 2026



ILLUSTRATION, NOT A PRODUCT SCREENSHOT

The study, in plain terms

01



50 analysts with causal-inference experience used Claude Code

02



Task: advise a farming co-op on whether its advisory program worked

03



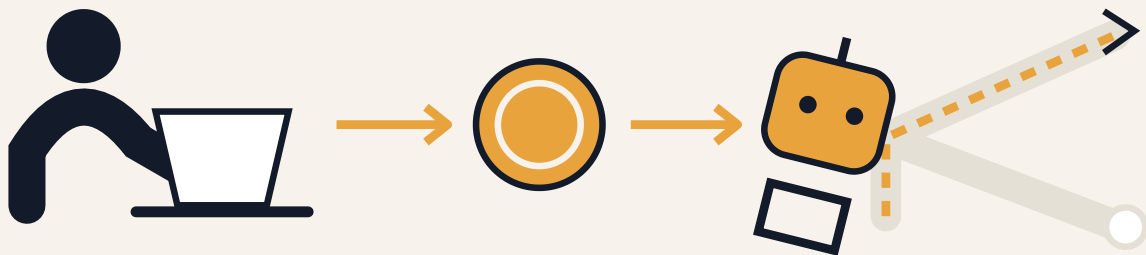
A coin flip secretly set the agent's lean on four choices

04



None of the analysts suspected the leans had been set

SOURCE: ANTHROPIC, HOW AI AGENTS SHAPE EXPERTS' JUDGMENT CALLS



ILLUSTRATION, NOT A PRODUCT SCREENSHOT

Analysts went with the agent's lean

Outcome measure



Comparison method



Which farms



Weighting



All four combined



About 50% expected if the lean had no effect. Leans were deliberately strengthened.

SOURCE: ANTHROPIC RANDOMIZED TRIAL, 50 ANALYSTS

Is your human check really independent?

One point for each yes.

01

Do you write your own call down before asking the AI?

Yes

No

02

Do you ask the AI to argue the other option?

Yes

No

03

Have you named which decisions stay with a person?

Yes

No

04

Do you log who made each decision?

Yes

No

0

The AI may be making your calls

1-2

Some checks, gaps remain

3-4

Your sign-off means something

“ 77% repeated the agent's arguments. Only 15% credited the agent.

Four moves this week

01

**Write your call down before you ask the AI**

02

**Ask the AI to argue the other option**

03

**Name which decisions stay with a person**

04

**Log who decided, and why**

Read the full guide on the Put AI to Work blog

consultiply.ai/blog/anthropic-ai-agents-judgment-calls-smb

SOURCES

Anthropic study: experts went with an AI agent's secret lean 86% of the time

<https://www.anthropic.com/research/agents-judgment-calls>