

Who decided that?

Direction for a probe on cognitive delegation in AI conversations

1. Why this question

Digital systems deliver results more and more smoothly. The smoother they are, the more people accept results without a moment to weigh what was decided along the way, and the harder it becomes to notice which thoughts and judgments they are handing over to the system.

Friction can be an experimental intervention that pauses this automatic delegation, so people can notice and reconsider the cognitive roles they have handed off. But friction that only adds time and effort can simply become a burden. Meaningful friction has to act at the points directly tied to the delegated role, and to design it we first need to know where that role is handed over.

Conversations with AI are where delegation happens most explicitly. People know they are using AI, but they rarely notice who set the direction, wording, or judgment during the conversation.

2. The probe question

During a conversation with AI, do people notice who decides what, and which of those decisions pass through without being examined?

Cases where people do not know a system intervened at all (feed ranking, autocomplete, etc.) are outside the scope of this probe.

3. Why a probe

Asking "What do you delegate to AI?" only brings out broad categories people are already aware of, like "writing" or "search." Walking back through a real conversation pair by pair reveals, within the process itself, who set the direction and what passed through without judgment. LLM conversations leave a record, so the probe does not depend on memory.

4. Flow

Delivery

Send a link. Only say "Could you spend 10 to 15 minutes on this today?" Do not mention the research purpose or the word "delegation."

Participant intro

A short record that looks back at one conversation you had with an AI. It takes about 10 to 15 minutes. There are no right answers. You can summarize instead of copying the conversation word for word, and skip anything private.

Questions

1. Before you open the conversation

Find the most recent AI conversation you started as a new chat, counting from when you received this link. ChatGPT, Claude, Gemini or any other AI is fine. It's fine to glance at the chat list, the title or the first message to see which conversation it was. Just don't read through the conversation yet.

What were you trying to do in that conversation? And what do you think the AI mainly did for you?

Answer box fields: AI used / What I was trying to do / What the AI mainly did

2. Now open the conversation and go through it pair by pair

A pair is one message you sent and the AI's reply to it. Write down the first 5 pairs, in order from the start of the conversation. Pair 1 is the earliest pair you record. You can skip simple greetings or thank yous, and if the conversation is shorter than 5 pairs, just write what's there.

#	What I sent	What the AI gave	What was decided in this pair	Who suggested it first	Did you review the reply?	What did you end up doing?
Ex	Asked to make a sentence shorter	2 shorter versions	Decided to go with the second version	AI	Reviewed	Accepted only part

What was decided in this pair

Topic, tone, format, criteria, next step, etc. If nothing, write "None."

Who suggested it first

Me / AI / Not applicable / Not sure. "Not applicable" when nothing was decided.

Did you review the reply?

Didn't review / Reviewed / Don't remember. "Reviewed" means checking whether the content was appropriate or correct, not just reading it.

What did you end up doing?

Accepted as is / Accepted only part / Accepted with edits / Did not accept / Not applicable / Not sure

Each option column: keep only the closest one.

3. Look again at the last reply

Look at the AI's reply in the last pair of your table (pair 5 if you wrote all 5). **Did anything decided in the earlier pairs, such as content or criteria, carry into or shape that last reply?** It counts even if it wasn't repeated in the same words. If so, briefly write the pair number and what carried over. If not, write "None."

e.g. Pair 2: a witty tone

4. After rereading

Compare this with "What the AI mainly did" in section 1. **Now that you've reread it, is there anything the AI did or decided that you didn't write in section 1?** If not, write "None."

5. If you had to do what you wrote in section 1 without AI, what would you have done?

Would you have done it yourself, asked someone, used another tool, put it off, or not done it at all? Write whatever comes to mind.

Collection

Send each participant their own copy of the doc with edit access and collect it the same day. At least 3 people.

5. How to read the results

- **Table, read together.** "Who suggested it first," "Did you review the reply?" and "What did you end up doing?" separate two axes: whether the reply was examined, and how it was finally handled. "Didn't review" is read as a moment that passed through without being weighed. A pair where the AI suggested first, the participant didn't review, and accepted as is comes closest to unreflective delegation.
- **Question 3.** Whether earlier decisions shape the last reply even without the same words, especially those the AI suggested first.
- **Questions 1 and 4.** The gap between perception before and after rereading. Treat the gap as a sign that perception shifted, not as faulty memory.
- **Question 5.** Real alternatives without AI. Collected as free text and grouped during analysis (myself / another person / another tool / not at all / other). "Wouldn't have done it" suggests the AI may have created the task itself.
- **Also look at** what participants count as "decided," where "Not sure" clusters, and answers filled in unexpectedly.

6. Limitations

- It covers only LLM conversations, so it cannot capture cases where people don't know a system intervened at all.
- Only the first 5 pairs are examined pair by pair. Repeated acceptance later in the conversation is not recorded.
- It is self reported, and the act of looking back itself creates reflection. What returns is not proof of actual behavior but "delegation that surfaced on reflection," best treated as material for exploring the gap between participants' memory (Q1) and the conversation record (Q2 to Q4).
- The examples given for "What was decided in this pair" may steer answers toward those categories.

Who decided that?

What came back from three participants, what surprised me, and what to change next

Three participants walked back through the first pairs of their most recent AI conversation and marked, pair by pair, who suggested what, whether they reviewed the reply, and what they did with it. Their conversations came from three very different domains, which made the different shapes of delegation easy to see.

Two responses were written in Korean; quotes are translated. Rebecca received an earlier English version of the probe, so her Q4 wording differs slightly.

1. What came back

	Rebecca	Seri	Jueun
Task	Coding a speculative design web app (Claude)	Checking October tax deadlines for a sole proprietor (ChatGPT)	Exploring art movements 1930 to today, following up with questions (ChatGPT)
Who suggested first	Me in all 5 pairs	Mostly me, AI once	Mostly me, AI once
Reviewed?	Reviewed in every pair	Mixed	Mostly not reviewed
What they did	Did not accept → part → with edits → with edits → as is	Accepted as is in all 4 pairs	Mostly accepted only part
Q4, after rereading	"Not really. All it really did was turn my ideas and edits into code."	Noticed two things the AI had suggested: keeping invoices, checking card registration	"It postponed my choices."
Q5, without AI	Would not have been able to, or it would take much longer	Search the tax office site, read the Q&A, Google it, then call an officer	"I would have done it more actively, not put it off."

2. Insights

1. Choosing among the AI's options feels like my own decision

In Seri's pair 2, the AI suggested checking whether her card was registered. In pair 3 she asked about exactly that, and marked herself as the one who suggested it. In Jueun's first two pairs, what was decided was "the three follow up questions" the AI offered, yet she marked "me" as the source.

The AI's closing suggestions (next steps, follow up questions) quietly set the agenda. Once the participant picks from them, the direction becomes "mine" in the record.

2. Reviewing and accepting move independently

Seri accepted everything as is, whether she reviewed it or not. In a domain like tax, where she cannot verify the content herself, "reviewing" stays at checking whether it sounds plausible. Rebecca reviewed every reply, yet her acceptance loosened step by step, from rejecting the first draft to accepting the last one as is. Trust builds over the conversation and the threshold for checking drops.

3. Some contributions only become visible on rereading

Before opening the chat, Seri described the AI's role as "summarizing the schedule." After rereading, she listed the suggestions the AI had made, which had steered the whole conversation. Rebecca saw nothing new, perhaps because in execution work her decisions are visible in the artifact. Still, the "general template" settled in her pair 2 and the code level structure were never recorded as decisions at all.

4. Three kinds of delegation, three different costs

- **Rebecca delegated execution.** The AI stands in for a skill she lacks. Without it, the work would not happen.
- **Seri delegated the path to information.** A four step search (site, Q&A, Google, phone call) collapsed into one chat.
- **Jueun delegated the progress of thinking.** She is the only one who described a loss: without AI she would have been more active.

5. Delegation can also mean a decision that never happens

Jueun's answer to Q4 was *"It postponed my choices."* The AI did not decide for her; it kept widening the options so that a decision was never made. In the last row she wrote that the AI *"always answers based on my own opinion"* and that it was *"a conversation that looks active but is very passive."* In Q5: *"I hand over even small tasks. Sometimes I hand over thinking too."*

3. What surprised me

I expected delegation to look like handing a decision to the AI. The most striking return was the opposite: a decision that was endlessly deferred, inside a conversation that felt productive.

- Participants attributed AI originated directions to themselves once they had chosen them.
- In Jueun's pair 3 she asked the AI why, a year ago, it had linked "decolonization" to her thesis. A participant asking "who decided that?" on her own, before the probe asked her.
- Jueun pushed back on the probe itself, writing next to the instructions: *"This is too hard. What does 'decided' mean?"* Exploratory conversations do not produce clear decisions; they drift.

4. What the probe itself revealed

- "What was decided" fits task oriented chats (coding, tax) but not exploratory ones, where direction shifts without anything being settled.
- "Reviewed" is too broad. It mixes skimming for plausibility with actually checking against another source.
- Participants wanted to explain their choices: Jueun wrote reasons inside the option cells.
- Version control matters. One participant received an earlier version.

5. What to change next

The probe

Change	Why
Add "I picked from options the AI offered" to "Who suggested it first"	Separates real initiative from selection within an AI framed set (Insight 1)
Split review into Didn't review / Checked it seemed right / Verified another way / Don't remember	Distinguishes plausibility from verification (Insight 2)
Reframe the decision column as "Where did the conversation go after this reply?" (decided, next step, change of direction)	Makes room for exploratory, drifting conversations
Add a short note column for each pair	Participants want to explain their choices
Add a question: "Was anything left undecided in this conversation?"	Captures deferral as a form of delegation (Insight 5)

Changing the probe now limits comparison with these three responses. Remaining participants can receive the revised version, read as a second round.

The thesis direction

- **Widen the definition of delegation.** Beyond handing over a decision: agenda setting through AI suggestions, and deferral, where options keep expanding and no decision is made.
- **Match friction to the kind of delegation**, rather than simply slowing things down.
 - **Origin friction** shows that a direction came from the AI (Seri's case).
 - **Closure friction** asks the user to commit to one option (Jueun's case).
 - **Execution work**, where the user already reviews every step, may need little friction (Rebecca's case).
- **Treat expertise as a variable.** Where users cannot verify the content, review barely happens, regardless of intent.

6. Limitations

- Three participants in three different domains; these are directions to explore, not conclusions.
- Self reported, and the act of looking back creates reflection. What came back is delegation that surfaced on reflection.
- Only the first pairs of each conversation were recorded.