

HAND IT OFF

Check an agent's work before you trust it

The agent gave you an answer. Is it good enough to rely on?

Handing work to an agent only helps if you can trust the result. Run an agent on a real task, then test its output the way you'd check a new colleague's work.

1

Try it your way

Use a real conversation if you can. See what you naturally ask Copilot.

[Try in Copilot →](#)

2

Need a cue?

Start with the outcome you need:

[Accuracy](#)
[Sources](#)
[Gaps](#)

3

Still stuck? Reveal a sample prompt

Have a go first — the sample prompt is on the next page when you want it.

[Reveal sample prompt ↓](#)

4

Did you get something useful?

- ☐ I checked key facts against the source
- ☐ I know which sources it used
- ☐ I spotted at least one gap or weakness
- ☐ I'd know what to fix next time

5

Take it one step further

Ask Copilot: "Which parts of your answer are you least confident about, and why?"

Need a little help?



[Watch a quick video](#)
(3 mins)



[See a worked example](#)



[Take the learning path on ASN](#)

Top tips

- ✓ Use your own work if you can.
- ✓ Be specific about the output you need.
- ✓ Iterate — your first prompt doesn't have to be perfect.
- ✓ Check the result — would you actually use it?

Your notes — what worked, what you'd change

HAND IT OFF

Sample prompt

Check an agent's work before you trust it

“ For the answer you just gave me: list the sources you used for each key point, tell me which points you are least confident about, and what information would make it stronger.

Select the text above and copy it into Copilot.

Why this works

- ✓ It asks the agent to show its sources, so you can verify.
- ✓ It invites the agent to flag its own weak spots.
- ✓ It turns checking into a habit you can reuse on any agent.

[Try it in Copilot →](#)[← Back to the challenge](#)

HAND IT OFF

Worked example

Illustrative example — names and details are fictional

The situation

You asked an agent to summarise customer feedback for Relecloud's new mobile app. The summary looks convincing, and it's going into a product review tomorrow.

What Copilot returned

 Copilot

Here's the agent's self-check on its feedback summary.

SOURCES USED

- App store reviews from September (412 reviews).
- Support tickets tagged 'mobile app' (88 tickets).

LEAST CONFIDENT

- The claim that 'most users want dark mode' rests on 19 reviews.
- No survey data was available, so satisfaction trends are inferred.

WOULD MAKE IT STRONGER

- The September in-app survey results.
- Feedback from the beta testers' Teams channel.

What to check

- ✓ Verify: open the source for the claim the agent is least sure about.
- ✓ Weight: 19 reviews out of 412 isn't 'most users' — reword before it goes to review.
- ✓ Habit: ask this check-question every time an agent's output matters.

[← Back to the challenge](#)