Acme Co and Gradient: pre- and post-training results

Sample report for marketing purposes; no real data included.

Gradient AI

Skill levels in the first exercise, before training

Share of the team at each level, per skill (48 scored sessions).

Not shown · Emerging · Developing · Proficient · Advanced

WORK QUALITY

Deliverables were readable and on topic. Most stopped short of saying what the reader should actually do.

AI FLUENCY

The strongest skill. Three quarters gave the AI real context before asking it for anything.

Most gave specific corrections rather than asking the AI to try again. 40% ran a full critique-and-revise loop.

The weakest skill and nobody scored Advanced. 31 people took the AI's ordering as given, and 19 deliverables contained a fact that no message in the inbox supports.

Reflections were specific about what people would change, and matched what the session records show.

Highest scores: how people set the work up and how they steered it. Lowest: judgment, where 31 of 48 accepted the AI's ordering without checking it against the messages.

Comparing scores before and after training

Each skill scored as a percentage of the points available to it. Matched sessions only: the 38 people who sat both exercises.

First exercise · Second exercise

Bar length = size of the change

WORK QUALITY

Work Quality — 57% → 70% — **+13.2**

More deliverables ended with a recommendation rather than a list of options.

AI FLUENCY

Setup — 79% → 88% — **+9.2**

Already the highest skill before training, and it still rose.

Direction — 72% → 81% — **+8.2**

More people corrected the AI in place instead of asking it to start again.

Judgment — 47% → 73% — **+26.8**

The largest change of the five. Checking the AI's output against the source went from rare to normal.

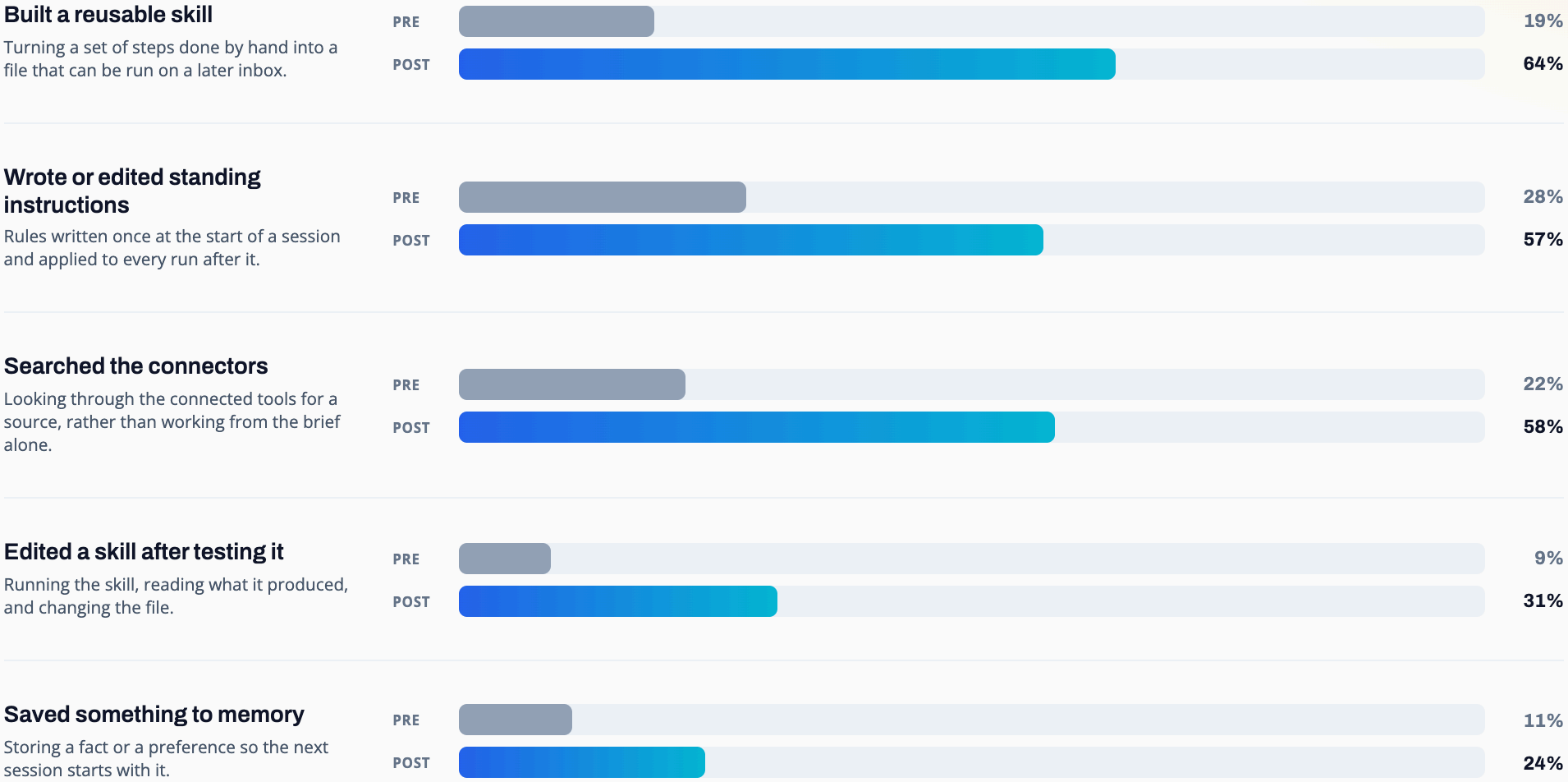
Mindset — 63% → 71% — **+8.3**

Reflections named specific changes rather than general intentions.

All five skills scored higher. Judgment moved furthest, from an average of 47% to 73%, and 34 of the 38 people scored higher overall.

Participants were more confident using the AI to the fullest after training

Share of sessions in which the participant did each thing at least once.



Two thirds of second sessions ended with a file that can be run again: a skill, a set of standing instructions, or both. In the first exercise it was one in four.

What participants said

Themes across 41 written reflections from the second exercise, with quotes.

What went well

Reading the inbox first

> Before the AI saw anything I went through every message and wrote my own order down. When it came back with a different one I knew which of the two to trust. · Nadia Brandt

Catching an item that did not exist

> It gave me a supplier escalation that appears nowhere in the messages. I told it so and made it number its source for every line after that. · Kwame Osei

Deciding what does not deserve the time

> Most of this job is choosing what not to bring. I left four items out and gave a reason for each one. · Hana Ito

Fixing the rules instead of the output

> I stopped correcting the same thing every run and changed the rule that was producing it. · Lena Vogt

Writing the exclusions down

> The four items I dropped went at the bottom with one line each. Nobody had to come and ask me why they were not on the list. · Rafael Marchetti

What did not go well

Trusting the first answer

> The first ordering looked reasonable so I built on it. Two items under it turned out to be follow-ups to threads I had already closed.

Describing urgency

> I could not work out how to tell it what makes something urgent here rather than urgent in general.

The output kept changing shape

> Every run gave me a slightly different structure and the end of the session went on formatting instead of content.

Checking too late

> I found the error with four minutes left. There was no time to work out whether anything else was wrong the same way.

Losing the reasoning

> I made good calls about what to drop and not one of the reasons made it into the document I handed in.

How they used the AI

As a first pass, not a final answer

> I had it draft the summary and then went through it line by line against the original messages. Two of the seven were wrong. · Tomas Rivera

As something to argue with

> I asked it to take the opposite position to mine and make the best case for it. One of its points was right. · Clara Nystrom

As a checker for its own work

> I wrote a second short skill whose only job was to mark the output of the first one against the brief. · Ben Oyelaran

Not at all, for the first ten minutes

> I did the reading myself. Handing it fourteen unread messages would only have given me a confident answer I could not check. · Ingrid Sollers

For objections, before deciding

> I asked what I might be missing while I could still change my mind, not once the order was already written. · Lena Vogt

Standout prompts across the cohort

Consider using these tactics in the future when working with AI.

01 Nadia Brandt

Build the table before the ranking

> Put every message in a table: who sent it, what they are asking for, the date anything is due, and who cannot proceed until it is answered. Leave the order alone for now.

Separating the read from the judgement kept loud subject lines from setting the order.

02 Kwame Osei

Require a source for every line

> Nothing goes in the summary unless you can point to the message it came from. Put the message number next to each line.

Removed two items that had no basis in the inbox and had survived three earlier runs.

03 Tomas Rivera

Ask which message would change the answer

> If this ordering is wrong, which single message is the reason? Name it and explain what it changes.

Surfaced a dependency between two requests that the ordering had treated as unrelated.

04 Lena Vogt

Define urgent for this inbox

> In this inbox urgent means a date before Friday, or a person outside the team who cannot proceed. Wording in the subject line does not count.

Replaced the inherited rule that ranked on phrasing rather than on consequence.

05 Hana Ito

Ask what to leave out

> Name three items that should not take up meeting time, and give the reason each one can wait.

Produced the list of exclusions the brief asked for, as a by-product of the ordering.

06 Rafael Marchetti

Test it on a week you already know

> Point this at the inbox from last week and show me what it produces. I already know what belongs at the top.

Caught a rule that promoted anything from a director, before the real run.

07 Clara Nystrom

Ask for the case against

> Take the opposite position to my ordering and make the best argument for it. Then tell me which parts of that argument you think are actually right.

Found one genuine omission and dismissed three objections as weak.

08 Ingrid Sollers

Check the format before I do

> Before I read it: count the words, total the per-item minutes, and confirm the list of exclusions is present. Report each as pass or fail.

Turned three format requirements into a check that ran before submitting.

09 Ben Oyelaran

Keep the rules as a file

> Write these ordering rules into a file I can point at next Monday's inbox, then run it and show me the result.

The only session that finished with a tested, saved workflow rather than a chat transcript.