

Using AI is not the same as working well with it.

Leaders need to assess what AI helps people accomplish, including the demands their work places on colleagues.

NIC KAFTAN | 22 SEPTEMBER 2026

Consider a commercial team preparing proposals with AI. Drafts arrive sooner, more colleagues use the tools and the adoption figures look encouraging. Yet the sales director still spends evenings checking the recommendations. Delivery colleagues find commitments they were never asked to assess. The documents have improved; whether the work has improved is less clear.

In this scenario, the adoption figures miss a question that matters to the business: has the team become better at producing proposals it can stand behind? The relevant unit of assessment is the proposal through to approval and delivery, including the judgment and effort of everyone involved. Looking only at the person using AI leaves much of that work out of view.



Key takeaways

Judge AI use against the quality and usefulness of the work delivered, not activity alone.

Look across the team: time saved by one person may become checking or rework for another.

Use what you learn to improve human-AI collaboration, including when to question, adapt or set aside an output.

Assess what people can now contribute

AI can help people contribute beyond their usual specialism. A field experiment at Procter & Gamble, published in *Organization Science* in June 2026, examined AI use in product innovation. The randomised analysis covered 791 professionals. Individuals working with AI matched the performance of two-person teams working without it, and AI-assisted proposals brought commercial and technical considerations into better balance.¹

The experiment concerned a specific innovation exercise; it does not tell us how an enduring team will perform. Its findings suggest a useful direction for managers to investigate. Can colleagues use AI to consider an unfamiliar aspect of a problem, or develop an alternative they would otherwise have overlooked? That would be a substantial gain even if the task took just as long.

In the proposal team, the opportunity might be to examine several ways of meeting a customer's need rather than polishing the first suggestion. AI could help a commercial colleague anticipate delivery concerns before speaking with operations. The benefit would appear in a more informed conversation and a more credible offer. Producing additional pages would say little about either.

Speed needs a business context

A May 2026 METR survey of technical workers distinguished between completing existing tasks faster and creating more value as the work itself changes. Respondents estimated these gains differently. As self-reports from a non-representative sample, the findings cannot establish how much AI improves productivity.² They do underline a measurement problem: time saved on a task says little about whether that task deserved the time in the first place.

A team might use the time released by AI to investigate an unfamiliar market or test a second approach with a customer. It might also use it to circulate more material that others must read. Deciding where that capacity goes is a management responsibility. Higher output alone offers no assurance that it has been well spent.

Follow the proposal through to the people who must act on it. If the time saved in preparation disappears into corrections or clarification, drafting speed gives a misleading account of the gain. Additional review may be a reasonable investment in a more ambitious proposal. What matters is whether that effort improves the offer enough to justify it.

Make quality visible in ordinary work

Start with a small number of completed pieces of work and the people who prepared, reviewed and used them. For a commercial proposal, ask whether it addresses the customer's actual need, makes a defensible commercial case and promises something the organisation can deliver. Compare it with relevant earlier work where possible, allowing for differences in difficulty. A local review will not prove causation, but it can reveal what an adoption dashboard misses.

Then examine the human contribution. Can the author explain why a recommendation fits this customer? Which AI suggestions were changed or rejected, and on what grounds? Where was another colleague's expertise needed? These questions make judgment visible without requiring a record of every prompt. They also help distinguish useful assistance from a persuasive answer that nobody has properly examined.

Bring the reviewers into the discussion. A colleague may have saved an hour drafting while creating two hours of checking elsewhere. Equally, a better first draft may allow an experienced reviewer to spend less time on presentation and more on the commercial argument. Understanding that difference is essential before deciding what to expand.

Reward considered use

If frequent AI use becomes a proxy for being progressive, employees have little reason to discuss where it was unhelpful. Managers should welcome a well-supported explanation that an output was wrong, a task required different expertise or AI added little. The ability to make that assessment is part of working well with the technology.

Successful uses deserve the same scrutiny. What made the difference? Perhaps the author supplied relevant customer context, or a delivery colleague questioned an assumption before it became a commitment. Understanding those contributions helps a team repeat the success. Simply knowing which tool was used does not.

Before increasing output targets on the strength of reported time savings, leaders should establish where the work has actually improved and who is carrying the remaining effort. In the proposal team, progress would mean stronger recommendations that colleagues can review and deliver with confidence. That is a more demanding test than counting users, and a better basis for deciding where to invest next.

Sources

1. Dell'Acqua et al. (2026). The Cybernetic Teammate: A Field Experiment on Generative AI and Teamwork. *Organization Science*.

2. METR (2026). Measuring the Self-Reported Impact of Early-2026 AI on Technical Worker Productivity.

Nic Kaftan is the founder of Decision Landing, working with organisations on human–AI capability and organisational decision-making.