ESSAY 04

BETTER JUDGMENT IN THE AGE OF AI

A Good Outcome Does Not Prove a Good Decision

"A good outcome does not prove a good decision."

Alen Mayer

alenmayer.com

A Good Outcome Does Not Prove a Good Decision

By Alen Mayer

There's a phrase that ends most arguments about a decision, and it usually arrives right after the result comes in: “well, it worked.”

It sounds like the final word. It usually isn't.

An outcome tells you what happened after the choice you made. It almost never tells you what would have happened after the choice you didn't make. That gap is easy to overlook, because you only ever get to watch one version of events play out — the one you chose. The other version, the one where you did something different, never happens where anyone can see it. So the version that did happen quietly absorbs all the credit, whether or not it deserves it.

THREE PLACES WHERE THE OUTCOME GETS PROMOTED TO PROOF

A salesperson believes a customer will walk unless she offers a bigger discount than the one a pricing system recommended. She overrides it, offers more, and the customer signs. The deal closing feels like vindication. But the signed contract only tells you the larger discount was *sufficient* to close the deal. It says nothing about whether the smaller one would have worked just as well. Nobody offered it, so nobody will ever know — and the closed deal gets treated as proof of something it was never in a position to prove.

A company hires someone who goes on to perform brilliantly, and the hiring process that selected them gets treated, from then on, as validated — copied, taught to other interviewers, held up as the model to follow. But the process also screened out other candidates who were never given the chance to show what they'd have done. One success doesn't tell you the process was sound. It tells you this particular bet paid off, which is a much smaller claim, resting on a much smaller sample: one.

An AI system logs a human override that led to a good result, and treats it — reasonably, by its own design — as a signal worth learning from. The override succeeded, so the system nudges its future behavior toward whatever that override did differently. Nobody ever established that the override itself was the reason for the good result, rather than something else entirely, or rather than the original recommendation working out fine too, had anyone let it run. The system isn't wrong to notice a success. It's wrong to assume the success proves what caused it.

Three different situations — sales, hiring, machine learning — and the same quiet leap in each one: something good happened, therefore the decision that preceded it must have been good.

WHY THIS IS SO EASY TO MISS

The trap works because it doesn't feel like a leap at all. Nobody sits down and consciously reasons "the deal closed, therefore my judgment was correct." The conclusion arrives already formed, attached to the good feeling of a result that worked out. Questioning it afterward doesn't feel like intellectual rigor. It feels like refusing to accept a win.

That's compounded by something organizations do to good outcomes on purpose: they reward the person who made the call that turned out well, often without ever separating out how much of the credit belongs to the decision and how much belongs to circumstances nobody controlled. Once a result is good, asking "but was the decision actually good" can look like an attempt to take something away from someone who just delivered — which is exactly why almost nobody asks it out loud.

> **An outcome tells you what happened after the choice you made. It almost never tells you what would have happened after the choice you didn't make.**

WHY AI CAN MAKE THE WRONG LESSON COMPOUND

AI doesn't have to make this mistake. A well-designed learning system can compare alternatives, run controlled experiments, use delayed outcomes, and build genuinely causal evidence instead of settling for "it worked once." None of that is beyond what modern systems do.

The problem sits earlier than the system's design. Someone still has to decide what a given outcome is evidence *of* — and that decision gets made by people, before any learning architecture touches it. If a successful override gets treated, from the start, as evidence that the override itself caused the success, that assumption can enter the feedback loop before the system does anything wrong at all. Once it's in there, AI doesn't need to be careless to do damage. It just needs to be efficient — at learning precisely the lesson it was quietly handed.

The danger was never that AI learns from outcomes. Learning from outcomes is exactly what a feedback loop is for. The danger is letting one observed outcome carry more meaning than the evidence actually supports, and then having a system available that can scale that interpretation across the next thousand decisions before anyone thinks to check whether the first one held up.

WHAT ACTUALLY SEPARATES A GOOD RESULT FROM A GOOD DECISION

The fix isn't to litigate every success as if it might secretly be a failure in disguise. Most of the time, a good outcome and a good decision really do travel together, and treating every win with suspicion is its own kind of dysfunction.

The fix is asking a narrower set of questions at the moments that matter — before a result gets promoted to precedent, and especially before it gets fed back into anything that learns from it. What does this outcome actually establish? What alternative explanations are still plausible? Is this one result, or a pattern strong enough to change what happens next? And before this becomes a training signal, a process, or a policy, what exactly are we asking the organization to learn from it?

Those questions don't slow down every decision. They're worth the time only where a single outcome is about to become a permanent lesson — a process, a policy, a piece of training data — rather than staying what it actually is: one result, from one attempt, that happened to go the way you hoped.

A good outcome does not prove a good decision.

TAKE THIS FURTHER

Build better judgment into how your team works with AI.

I help leaders and organizations strengthen the human capabilities behind better judgment as AI becomes part of everyday work and decision-making.

JUDGMENT LABS

A 90-minute session working through realistic AI-influenced business decisions.

WORKSHOPS

Practical training on how teams question, interpret, weigh and decide with AI.

KEYNOTES

Reframing what human judgment requires as AI becomes more capable.

Alen Mayer

alenmayer.com