

Why Can't You Just Be Told?

Somebody once gave you exactly the right advice and you did not use it. You understood every word. You still had to get it wrong yourself before it meant anything.



WHERE THIS SITS

There is more language in the world than anyone can attend to. The obvious fix, an AI that reads it all for you and tells you what is true, does not work: the judgement that would take requires having been wrong at your own expense, and nothing a model does costs it anything. So the judging stays with people, and almost nothing we currently inspect can tell whether a person actually judged or simply agreed.

01 Why Does It All Feel Like Too Much?

02 **Why Can't You Just Be Told?**

03 What Holds a Claim Up?

04 How Would You Know?

05 Who Pays If It's Wrong?

06 What Would You Build Instead?

Six short pieces, each drawn from a working paper. Each one stands on its own; this is the second. They are ordered so that every piece ends by raising the question the next one is called.

Advice you did not use

Think of something you were told early on that turned out to be completely true, and that you ignored, or nodded at, or filed away. Then think of the day it finally landed, which was almost certainly the day it cost you something.



Four different things, not one pile

DATA

A reading. 38.2 degrees.
160 over 95. Four
millimetres of deflection.

INFORMATION

That reading put in
context. Compared with
what it should be.

KNOWLEDGE

The pattern across many
such readings, and what
to do about it.

WISDOM

Treated by almost
everyone as simply more
knowledge. It is not.

The first three are a stack. The fourth is a different kind of thing, arrived at a different way.

The last step is the odd one

The first three can be handed over. I can give you my data. I can give you my information. With more effort I can teach you what I know.



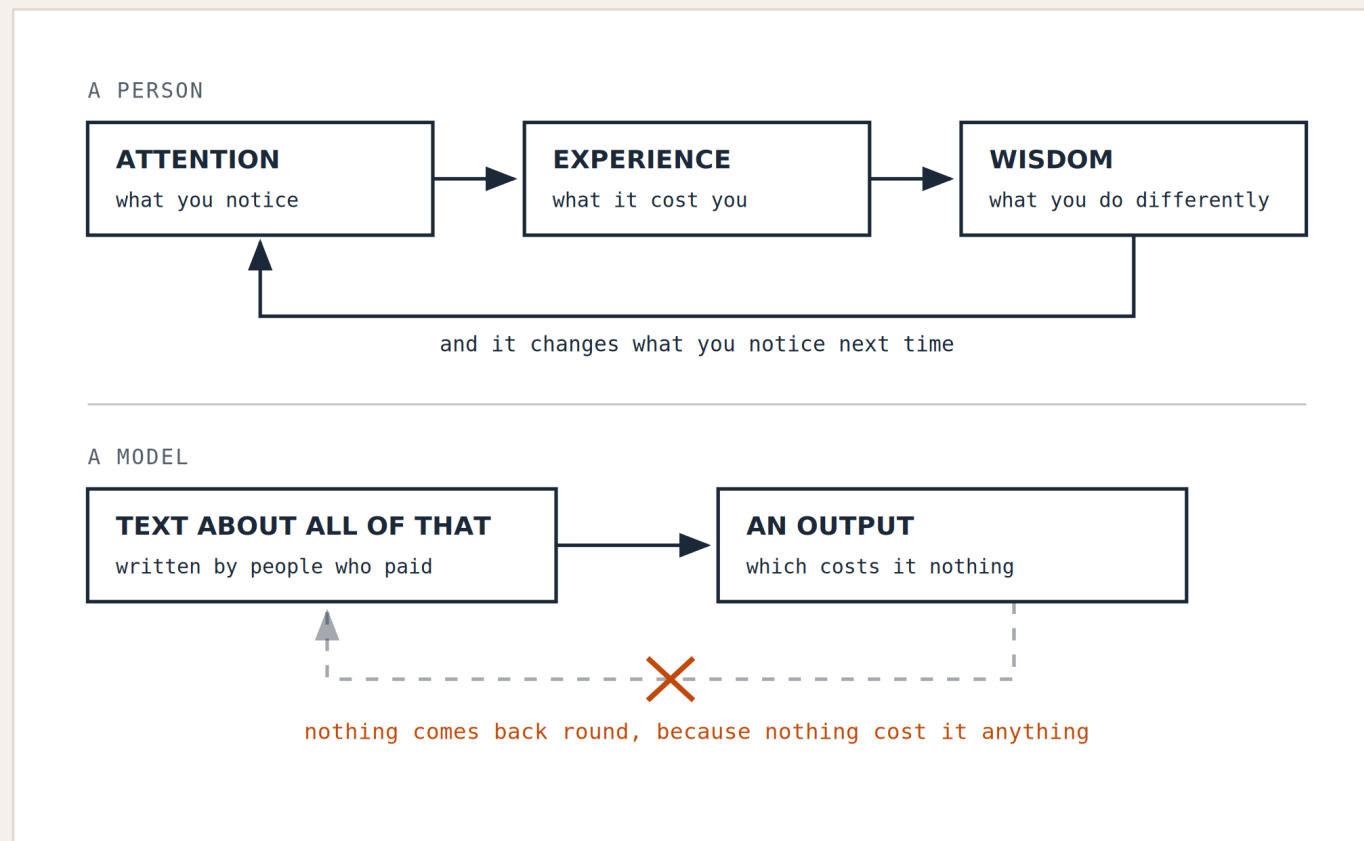
I cannot give you what twenty years of being wrong did to me.

Not because it is mystical. Because the mechanism that produced it ran on me, over time, and it needed me to be the one carrying the cost.

The loop that produces it

What you pay attention to decides what you experience. What you experience, especially where it hurts, changes your judgement. And changed judgement changes what you notice next time, which is where the loop closes.

Take the return arrow away and what is left can still produce excellent work. What it cannot do is get wiser. The loop is not about the quality of the output; it is about whether anything changes in the thing producing it.



What makes a consequence real

Not every bad outcome teaches anything. Three things have to be true, and if any one is missing the loop does not turn.

YOU CANNOT UNDO IT

A result you can roll back and try again is a simulation of a consequence. Useful for learning a game. Not the same thing.

IT LANDS ON YOU

Not on the next shift, not on a user you never meet, not on an abstraction called the business. On the person who made the call. On you.

IT CHANGES WHAT YOU ARE

Next time a case looks textbook-clear, something in you hesitates, and you could not switch that off if you wanted to.

It met every specification

In July 1940 a suspension bridge opened across the Tacoma Narrows. It had been designed to the standards of the day and it met them.

JULY 1940

Opened, having met every specification that applied to it.

NOVEMBER 1940

Down, in a wind of around forty miles an hour, far below its design load.

The calculations were not wrong. They were calculations of the things anybody knew to calculate. What brought it down was a behaviour that was not in them, and the profession only learned to look for it by watching this happen.

That is what experience supplies and knowledge cannot: not better answers to the questions you asked, but the questions you did not know were there.

Watching, doing, imagining

Three ways of getting at why something happens, and they are not interchangeable. Reading about an intervention is not performing one, and no quantity of reading converts into the other kind.

WATCHING

what goes with what,
across everything

WHO LIVES HERE

Language models

an enormous amount

DOING

change it on purpose,
watch what follows

WHO LIVES HERE

Learning systems

only inside a scored world

People

and paying for it

IMAGINING

what would have
happened otherwise

WHO LIVES HERE

People

and paying for it

language models never cross this line

But the helicopter really crashed

The strongest objection, in its strongest form. A program taught itself Go and found moves nobody had thought of. A research helicopter taught itself aerobatics and crashed real remote-controlled aircraft doing it. That is not reading about intervention. Conceded.

THE OPEN WORLD

no score · no end · no reset
the result arrives years later
and lands on somebody

inside the box means
somebody wrote down
how to score it

EVERY OUTCOME CAN BE SCORED

A GAME OF GO

won or lost, every time

A RESEARCH HELICOPTER

it really crashed, and it was still scored

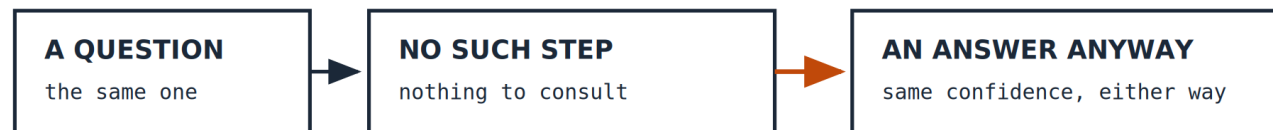
The failure is always the same shape

Legal briefs citing cases that do not exist. A load calculation produced just as fluently whether or not the figure was ever supplied. Not lies, and not random errors: no sense of where its own knowledge stops.

A PERSON WITH EXPERIENCE



A MODEL



the missing step is the whole problem

Wisdom-shaped, without the history

A model trained on the writing of people who did traverse the loop will produce output that looks like their judgement. Measured hedge for measured hedge, it may be indistinguishable.



What is absent is not the words. It is everything that made those words trustworthy in the person who first wrote them.

The cases that went badly. The hesitation installed by the hundredth presentation. Knowing which way they personally tend to be wrong.

Two groups, one tool

Researchers gave students an AI assistant to research a scientific question, and compared the reasoning they produced against what they produced without it. Both groups got the same task. Only one had been trained in the subject.

MEDICAL STUDENTS

The question sat inside the field they had been trained in.

Better with the assistant than without it.

SOCIAL SCIENCE STUDENTS

The same question, outside their expertise.

Worse with the assistant than without it.

Same task, same tool, opposite direction. The only thing that differed was what each group already knew. If you have ever caught an AI being confidently wrong about something you happen to know well, you have already run this experiment on yourself.

Stadler, Bannert and Sailer, 2024, with a 2026 preprint adding domain expertise as the moderating variable. The follow-up is a preprint and has not yet been peer reviewed.

What made the difference

The earlier study in the same line of work found something more uncomfortable than a drop in quality: the students using the assistant reported a lower mental workload while producing worse reasoning. It felt easier at exactly the moment it was going wrong.

The students who already had judgement could interrogate what came back: which claims to be suspicious of, which numbers smelled wrong, what had quietly been left out. The others could only accept it or reject it whole.

The tool does not supply judgement. It multiplies whatever judgement you walked in with, including none.

Where that judgement came from

Nobody is born able to interrogate an output. It comes from a specific and rather unglamorous apprenticeship: the debugging at eleven at night, the hundredth presentation of the same condition, the case that collapsed on one witness you had not prepared.

HOW IT HAS ALWAYS WORKED



ten to fifteen years, and no way to compress it

WHAT WE ARE AUTOMATING



you still need seniors. you have stopped making them.

Why nobody notices in time

Cutting the junior work looks like a clean win for years. Output goes up. Costs go down. Nothing visibly breaks, because the senior people carrying the judgement were formed under the old arrangement and are still there.

The bill arrives when they retire, and by then the decision that caused it is ten or fifteen years old and nobody can point at it.

What this does not claim

Short versions tend to overstate. Here is what this one is not saying.

✗ ~~AI is not useful.~~

It made the medical students measurably better. The argument is about what it cannot supply, not about whether it helps.

✗ ~~People are reliably wise.~~

Obviously not. The claim is narrower: human judgement can be held to account, improved by being wrong, and carried by somebody who pays.

✗ ~~This is permanent.~~

The paper concedes that systems which act in the world move the gap rather than close it. And if somebody has a better account of what wisdom is, the author would rather hear it than defend this one.

✗ ~~Here is the experiment that proves it.~~

There isn't one. Wisdom does not show up cleanly in a controlled trial, and the paper says so rather than dressing an argument as a finding.

Three things that follow

1. CATCH IT BEING WRONG

Only use it where you could. If you cannot interrogate the answer, you are not using a tool, you are taking dictation. The medical students could push back. That is the whole difference.

2. KEEP THE CRUCIBLE IN THE WORK

If every slow, close-up, get-it-wrong-yourself task is removed from a junior role, that role stops producing seniors. That is a design decision, whether or not anybody makes it deliberately.

3. PROTECT THE PIPELINE

The cost of cutting it is real, delayed by a decade, and paid by people who were not in the room. It is the cheapest thing to cut and the slowest to rebuild.

Where this comes from

This is a short version of a working paper called The Wisdom Gap, which carries the evidence, the sources and the objections, including the ones that have not been answered.

ACKNOWLEDGEMENT

The idea behind this paper started in an exchange with Louise Vigean, who writes Think Therefore AI. The argument took its shape there before it had a name.

Working papers are revised when they are shown to be wrong, and what changed is kept on the record.

James Maconochie | Architecture & Attention

HUMILITAS, HUMANITAS, ET CURIOSITAS · SEMPER

Humility, humanity, and curiosity. Always.

