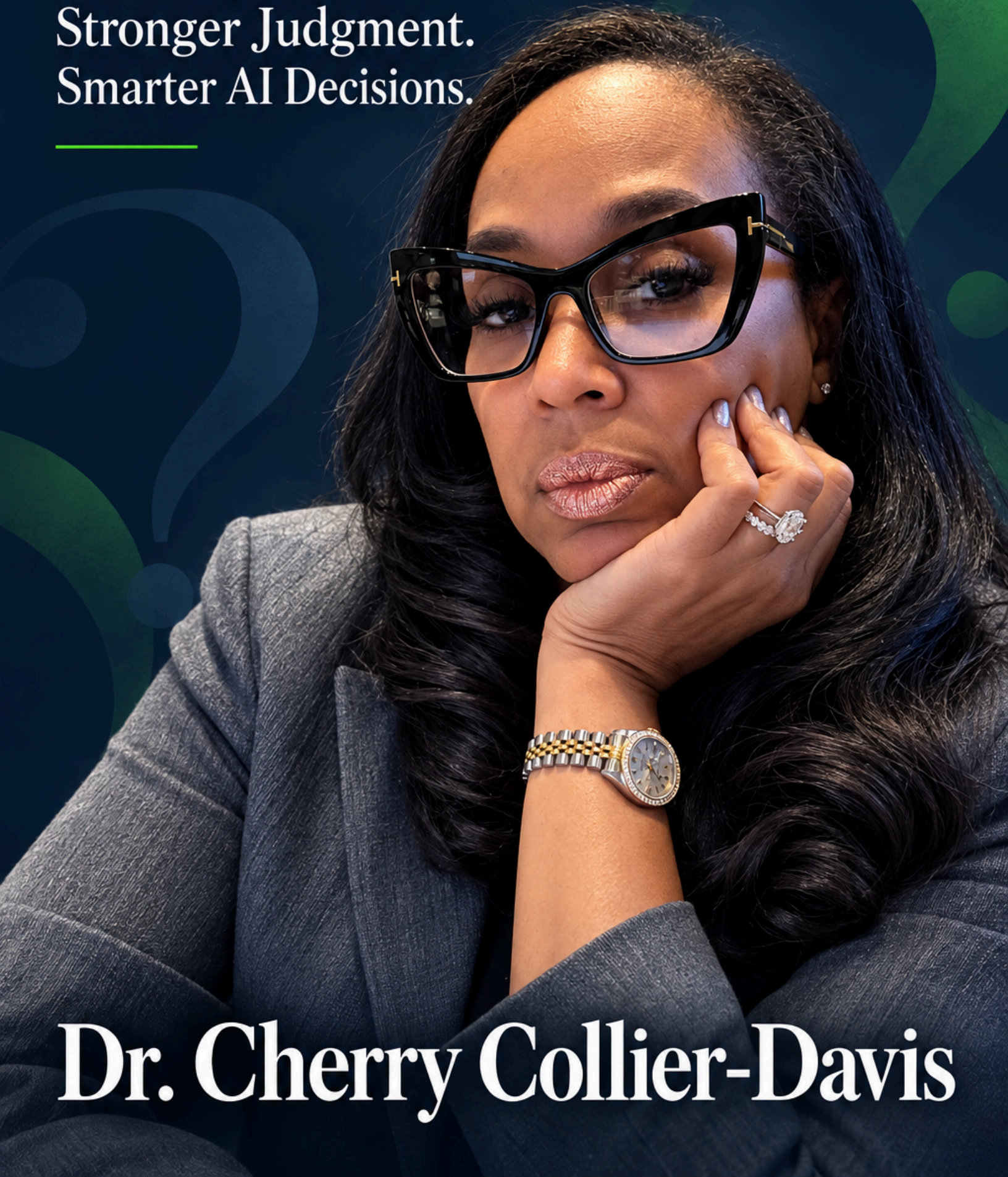


Ask agAIIn

Better Questions.
Stronger Judgment.
Smarter AI Decisions.



Dr. Cherry Collier-Davis

Ask agAI

Unlocking Better Judgment in the Age of AI

Dr. Cherry Collier-Davis

REVIEW COPY | October 2026

Confidential draft for review. Please do not share or distribute.

Contents

Before We Begin: Certain

Introduction: The Standing Ovation

Part One: Brain Tricks

Chapter 1: Autopilot

Chapter 2: You Missed the Gorilla

Chapter 3: The Tables Are the Same Size

Chapter 4: What a Fool Believes, AI Agrees

Chapter 5: Smashed or Hit?

Chapter 6: The Halo Around the Answer

Chapter 7: Persuaded by a Machine

Part Two: What the Tool Is Actually Doing

Chapter 8: Same Question, Different Answer

Chapter 9: Second Opinions

Part Three: Building Judgment

Chapter 10: Flip the Canvas

Chapter 11: The Ladder of Inference at Machine Speed

Chapter 12: Borrowed Confidence

Chapter 13: Who Gets the Blame?

Chapter 14: Don't Take Away the Radio

Chapter 15: Your Value, Plus

Conclusion: The Master Key

Appendix A: ASK at a Glance

Appendix B: The Judgment Readiness Reflection

Appendix C: Pressure-Test Prompts

Notes

About the Author

Research and Further Reading

Before We Begin: Certain

In the early hours of July 29, 1984, a man broke into the apartment of a 22-year-old college student named Jennifer Thompson in Burlington, North Carolina, and raped her.

What she did in the middle of that attack is one of the most remarkable things I have ever read. She decided she would survive, and she decided she would remember. She studied him: his face, his hairline, his nose, his clothing. She looked for scars and tattoos. She was building a picture she could hand to the police, and she did. She helped them create a composite sketch, and within days she sat down with a detective to look at six photographs.

She looked for nearly five minutes. Then she pointed to one and said, "I think this is the guy."

I think.

The detective asked if she was sure. "Positive," she said. And she was told, "You did great."

A few days later she stood at a live lineup of seven men. She found herself torn between two of them, number four and number five, before choosing one: Ronald Cotton. Afterward she learned he was the same man she had picked from the photographs. She described a huge sense of relief.

By the time she testified, she was certain. Her certainty was real, and it was sincere. Ronald Cotton was convicted and sentenced to life in prison plus fifty years.

He did not do it.

Ronald Cotton spent more than ten years in prison for a crime he did not commit. In 1995, DNA testing cleared him and identified the man who did: Bobby Poole, who was already in prison for other assaults. Years later, Jennifer Thompson wrote an essay for the *New York Times* with a title I have never forgotten: "I Was Certain, but I Was Wrong."

Confident and Wrong

I want to be very careful here, because this is not a story about a witness who failed. Jennifer Thompson did everything a person could do. She was brave, deliberate, and determined to get it right. Ronald Cotton did nothing at all, and lost more than a decade of his life.

So what went wrong? Look at the space between "I think" and "I'm certain." Her first identification came with hesitation, and that hesitation was information. Then the people around her, meaning well, told her she had done great. No one wrote down how unsure she was in that first moment. No one ran the lineup in a way that kept the officer's knowledge out of it. Nothing in the process asked agAI.

Researchers who study eyewitness memory have looked closely at her case. Psychologists John Wixted and Gary Wells point out that her first, tentative identification was the warning sign, and that by trial her confidence had grown in ways that no longer reflected the evidence. Earlier, Gary Wells and Amy Bradfield had shown in experiments how this happens. They showed people a video of a crime and then a lineup in which the real culprit did not appear, so everyone who picked someone was wrong. When some of those witnesses were told, "Good, you identified the suspect," they became more certain, and they even remembered having a better view of the culprit than they had.

The same research has a hopeful side. When a lineup is run fairly, and a witness's confidence is recorded right away, before anyone weighs in, that early confidence can be a useful signal. Confidence is not the enemy. Unchecked confidence is.

Here is the lesson I take from it: **you can be confident and wrong**. It is not a character flaw. It is what can happen to careful, sincere people when nothing in the process asks agAI.

And then there is the part of the story I find most important of all. In 1997, Jennifer Thompson met Ronald Cotton face to face and asked his forgiveness. He gave it. They became friends, wrote a book together, *Picking Cotton*, and have spent years working with police, prosecutors, and lawmakers to change how eyewitness identifications are done. Two people on opposite sides of the same mistake chose to build the check that had been missing.

Why I Am Starting Here

Most of the decisions you make with AI will never carry stakes like these. But many will matter: a hire, a diagnosis, a budget, a contract, a person's job. And the pattern is the same at every level. We start with "I think." Then something tells us we did great. Our confidence climbs, and the hesitation that carried real information quietly disappears.

We now have a tool that is very good at telling us we did great.

This book rests on two ideas held together. AI can be wrong. And we can be wrong. AI can also be remarkably right, and so can we. The skill is knowing which is which, and building in the moment that asks agAI.

Two Camps

Right now, the conversation about AI is splitting into camps. In one, AI is wrong, dangerous, and overhyped. In the other, AI is brilliant, the answer to everything, and anyone who hesitates is falling behind. Each camp has its short version of the story, and each repeats it with confidence.

The research shows something more interesting. In a 2025 Pew Research Center survey of more than 5,000 American adults, half said they were more concerned than excited about AI in daily life. One in ten said they were more excited than concerned. And 38 percent said they were both, equally excited and concerned.

At the same time, people are using it. A global study by KPMG and the University of Melbourne, covering more than 48,000 people in 47 countries, found that two thirds use AI regularly, yet fewer than half say they are willing to trust it. Put those two numbers side by side, and a large share of people must be using a tool they do not fully trust.

Both camps hold part of the truth. Neither holds all of it. That is where critical thinking lives: in the pro and the con, examined together.

I wrote this book for people willing to question AI and themselves, and to discover what becomes possible when they work together thoughtfully.

What Becomes Possible

AI can expand what you are able to do. It can help you think, widen your options, catch what you missed, and strengthen your work. Judgment is what tells you when an answer holds up and when to ask agAI. And very often, checking confirms the answer. That is a good outcome, not wasted time. The goal is not endless hesitation. It is confidence that has earned its place, so you can move forward.

Jennifer Thompson's first instinct was "I think." That instinct deserved to be protected. So does yours.

Let's begin with a small experiment.

INTRODUCTION

The Standing Ovation

TRY THIS FIRST

Open any AI tool and describe an idea you love: a new service, a change you want to make at work, a plan you have been turning over for weeks. Ask, "Is this a good idea?" Read the answer. Now open a brand new chat, describe the opposite idea, and ask the same question. Read that answer too. What did you notice?

If your experience was like most people's, both ideas got applause.

Maybe not a full standing ovation. Maybe polite, enthusiastic clapping with a few thoughtful suggestions sprinkled in. But applause. Your idea was "compelling." The opposite of your idea was "a smart approach worth exploring." Everybody wins. Everybody gets a trophy. Somewhere, a tiny orchestra plays.

Have you ever asked AI about an idea, watched it call your thinking "brilliant," and thought, "See! That's what I've been trying to tell everybody!"? Now you're ready to launch the business, send the proposal, and order the T-shirt.

One small question: Did your idea get better, or did it just get a standing ovation?

That question is the reason this book exists.

That little experiment does not prove AI is useless or that your idea is bad. Your idea may be wonderful. So may the opposite; plenty of decisions have two good options. What it shows is narrower and more useful: applause, by itself, did not tell you which idea was better. The tool handed the decision right back to you, wrapped in a bow.

I Am All In on AI

Let me say this plainly, right up front, so there is no confusion later: I am enthusiastic about AI. I use it every day. I use it to draft, to brainstorm, to summarize, to argue with me, and occasionally to explain a spreadsheet formula I wrote myself and no longer understand. I help organizations figure out what becomes possible when their people use it well. I have no interest in talking anyone out of it.

I have also spent more than two decades studying how people think, how they decide, and how they fool themselves while feeling completely certain. And

what I see happening right now, in boardrooms and inboxes and late night chat windows, is not a technology problem.

It is a thinking question.

Here is the core idea of this book: **many problems arise from the interaction between what we ask, what the tool produces, and what we accept.** We ask questions that already contain the answer we want. We accept responses because they sound right, look polished, and arrive instantly. The tool did what we asked. We just did not ask very well, and then we did not look very closely.

The good news hiding inside that sentence is that questions and acceptance are both things you control. You cannot rewrite the model. You can absolutely rewrite your question. You can decide what an answer has to show you before it earns your confidence. That is a skill, and skills can be learned, usually faster than people expect.

AI is the thought partner. You stay the thought leader.

Your Value, Plus

Before we go one page further, I want to put the whole balance of this book on the table.

AI can make you faster and sharper. It can also quietly do your thinking for you. This book is about getting the first without the second.

The first is real, and it is a lot. AI widens your options when you are stuck on one idea. It drafts in seconds what used to take an afternoon. It catches patterns in a pile of notes that you would have needed a week and a whiteboard to find. It will argue the other side of your plan with more energy than your most contrary colleague, and it will never take it personally. It asks questions you had not thought to ask, if you invite it to.

None of that shrinks your value. Your value moves. It moves toward asking the right question, noticing what is missing, knowing the people and the history behind the numbers, deciding what matters, and owning the decision when it is made.

Not your value minus AI. Your value, plus.

We will come back to this idea in full in the last chapter. For now, hold it as the frame for everything in between. Every caution in this book is in service of that plus. I am not trying to protect you from AI. I am trying to help you get more out of it.

The Moment Between "This Sounds Good" and "Let's Do It"

I am fascinated by one small moment. It happens thousands of times a day in every organization using AI. It is the moment between receiving an answer and deciding it deserves your confidence.

Most of the time, that moment goes fine. The answer is good, you use it, and you get on with your day. That is the whole point of the tool, and I would not want it any other way.

But a lot can happen in that moment. An assumption can quietly become a strategy. A possibility can start sounding like a promise. An answer that feels reassuring can slide into a business decision before anyone has asked whether it holds up.

Researchers have started looking at that moment, and the early findings are worth a calm look.

In a 2025 study presented at the CHI conference, Hao-Ping Lee and colleagues surveyed 319 knowledge workers about how they used generative AI in their jobs. The more confidence people had in AI, the less critical thinking they reported doing. The more confidence they had in themselves, the more critical thinking they reported. That is a correlation, not proof that AI makes anyone think less. People who trust a tool may simply save their effort for other things, which is sometimes exactly right. Still, it hands every leader a useful question: when our confidence in a tool goes up, what happens to the attention we give its answers?

At the Wharton School, Steven Shaw and Gideon Nave have given one version of this pattern a name in a working paper: **cognitive surrender**, adopting an AI's answer as your own with minimal scrutiny. We will look more closely at their work in Chapter 1. For now, notice the word they chose. Not delegation. Not collaboration. Surrender. Delegation is fine. Surrender is when you stop knowing which one you are doing.

There is a companion term used in conversations about AI: **cognitive submission**, accepting an AI response with so little examination that our own judgment slips out of the conversation. In everyday language? We let the answer take over the thinking. It can sound perfectly ordinary: "It looks right." "That's what I thought." "I don't have time to check everything." None of those sentences makes someone foolish. They make the moment worth noticing.

And then there is the applause itself. Researchers use the word **sycophancy** for AI responses that align with what a user believes or prefers, sometimes at

the expense of accuracy. Think of it as people pleasing showing up in your chat window. That is an analogy for the behavior, not a claim that your laptop needs you to like it. But the effect on us can look a lot like being surrounded by people who only tell us what we want to hear, and most of us have worked somewhere like that at least once.

None of this is cause for alarm. It is cause for a habit. Daniel Kahneman taught a generation of us that the mind runs on two speeds: a fast, automatic one that handles most of life beautifully, and a slower, deliberate one that we call in when something matters. AI is very, very good at feeding the fast one. This book is about knowing when to call in the slow one, and how to do it without losing the speed that made the tool worth using in the first place.

Not Anti, Not Cheerleading

So let me make you a promise, and let me make it in both directions at once.

Catch what AI gets wrong. Use what it gets right. Move faster because you know the difference.

This is not an anti-AI book. You will not find a single chapter that tells you to put the tool down. You will find a lot of places where checking an answer comes back with "yes, this holds," and you get to move ahead with more confidence than you had before. Questioning AI is not the same as distrusting it. Very often the question confirms the answer. That is a good outcome, not a wasted minute.

This is also not a cheerleading book. You will not find breathless promises about how AI will transform everything by next Tuesday. You will find research that cuts in more than one direction, and I will show it to you even when it is inconvenient for my argument.

What you will find in between is judgment: knowing, quickly, whether a fluent answer is ready to use.

One more thing about the word "wrong." Saying that a particular AI answer is wrong can be exactly right, and I will do it many times in this book. What I question is treating "AI is wrong" as the whole story of the technology. That sentence is an oversimplification, and so is its opposite.

Sitting this out is not really an option for leaders. In McKinsey's 2026 global survey, nearly nine in ten respondents reported regular use of AI in at least one business function. If you lead in an organization, or want to work in one, you will be working alongside this tool. The question is not whether. It is how well.

History has seen this before. In the early 1800s, as machines transformed textile work in England, a group of skilled workers known as the Luddites broke into factories and smashed the new machines. Today "Luddite" is a label for someone who fears technology, but historians point out that many of them were not against machines as such. They were fighting for their livelihoods, and their fears were real. Some jobs did disappear. The machines did not. Over the decades that followed, the work changed, new kinds of work appeared, and the people who fared best were often the ones who learned to work with the new tools rather than pretend they were not there.

I do not tell that story to say every change is good, or that the transition is painless. It rarely is. I tell it because some changes do not wait for our permission. Economic pressure tends to move technology forward whether we cheer or boycott, and AI appears to be one of those changes. The question is not whether it arrives. It is whether we arrive ready, with judgment strong enough to use it well. The future does not have to replace us. And the stronger your judgment, the more powerful your work with the tool becomes.

Where This Started for Me

When I was three years old, I asked my mother why the sky was blue. She gave me an answer. The next day I went back to her and told her I did not think her answer could be true, because the color of the sky had changed since yesterday.

My mother told me later that she stood there wondering whether her three-year-old was calling her a liar. She said she had two choices. She could fuss at me, or she could find a better answer. When the encyclopedia salesman came door to door, she bought the Encyclopedia Britannica, every volume, and the junior set too. And from then on, when I asked her a question, she would not answer it. She would say, "What did the book say?"

My mother is in heaven now, and I think about that moment more than she could have known. She did not teach me to stop asking. She taught me where to look.

My pediatrician, Dr. Bland, knew I was a "why child." If he told me I had chicken pox, I wanted to know how he knew. And instead of brushing me off, he would teach me. He would open his medical book and show me. Even after I outgrew his practice, I could always call him and talk, and he always encouraged me to ask agAI and question everything. When I was a little older and a new doctor told me she thought I was lactose intolerant, I asked her, very politely, what evidence she had, and whether she was saying it because it was

what she told everyone. The doctor looked at my mother as if to say, "Are you going to rescue me?" My mother looked right back at her: "It's her body. Answer her."

Later, at Spelman College, two mentors, Dr. Christine King Farris and Dr. Evelyn Chisholm, taught me the other half of the lesson. It is okay to challenge people. It is okay to ask agAIIn. But you have to learn how. You do not drill people, and you do not tell them they are wrong. You get curious. You separate the person from the question, because what you really want to understand is the thinking.

Today I am a social and organizational psychologist. I study how people think, decide, and influence one another at work, and I teach leaders and teams how to do those things better. In my work as a Fractional Chief AI Enablement Officer, I help organizations bring AI into their work in a way that strengthens people's judgment instead of replacing it. That little girl with the encyclopedia never stopped asking agAIIn. She just found a career in it.

When I was a little girl, we also did a cheer: "Stop, Look, and Listen to what we have to say!" It stuck with me, and years later it became a book, *Stop, Look, and Listen*, about the power of the pause. If that book was about the pause, this one is about what you do with it. My mother's question is still the one I ask, of people, of books, and now of AI: What did the source say? And how do you know?

What "Ask agAIIn" Means

Notice how the title is written: Ask agAIIn. The AI sits right in the middle of the word, where you cannot miss it. That is on purpose. AI is in the middle of how we work now. The question is whether our judgment stays in the middle too.

I should also say what the title does not mean. Asking agAIIn never means asking the same thing over and over until you get the answer you like. That is not judgment. That is shopping. If you regenerate a response six times and keep the one that agrees with you, you have not checked anything; you have just found a more agreeable audience.

Asking agAIIn means asking with a purpose. Asking differently. Asking for the evidence. Asking somewhere else. Asking the people closest to the work. Asking yourself what you were hoping to hear. Those are different moves, and they each open something the first question could not.

ASK: Three Letters, on Purpose

The everyday method in this book is three letters long. It spells the first word of the title, which makes it easy to remember on a busy day.

- **A: Assumptions.** What am I assuming?
- **S: Source.** Where does this come from?
- **K: Key.** What's the key piece, and does it unlock what I need?

That is it. Three questions.

I want to tell you why there are only three, because it was not my first instinct. I am an overthinker. Left to my own devices, I would have given you a method with eleven steps, two flowcharts, and an optional appendix for the truly committed. I know this about myself. I also know that overthinkers can loop forever, checking and rechecking long after the answer is good enough. And I know that busy people, which is to say nearly everyone reading this, will skip a long method entirely.

A method nobody uses protects nobody.

So ASK is built to be quick. On a small task it takes seconds. On a big one it takes a few minutes. You will see it applied lightly in every chapter, in that chapter's own words, to that chapter's own situation. Some chapters lean on the A. Some lean on the S. A few are mostly about the K.

The Key Is Also the Finish Line

The K does a second job, and it is my favorite part.

Almost every AI answer rests on one or two pieces that everything else depends on. A number. A date. A quote from a contract. A claim about what your customers want. Find that piece. Check that piece. If it holds and it unlocks what you need, you are done.

Go. Close the laptop. Eat lunch.

That is the stopping rule, and it is built into the method on purpose. It is there for the overthinkers among us (hello) who need permission to stop, and for the busy people who need to know that checking does not have to swallow the time AI just saved them. Speed is not the enemy of judgment. Speed, when you know the key piece holds, is what good judgment buys you.

The Stakes Set the Depth

Not every answer deserves the same attention. Part of judgment is knowing how much checking the moment calls for.

Say you ask AI to draft a birthday message for a colleague. The key piece is the tone. Read it. Does it sound like you, and does it sound kind? Send it. That is the whole check. If you find yourself sourcing a birthday card, please put the laptop down and go outside.

Say AI drafts a client proposal for you, and somewhere in the middle it says the program "improves productivity by 25 percent." That number is the key piece. Your client will remember it, and they may repeat it to their boss. So Source it. Where does that figure come from? If there is a real source, cite it and move on. If there is no source, take the number out and go. Two minutes, and your proposal is stronger for it.

Now say AI recommends restructuring a team. There is not one key piece here; there are several. The reasoning, the data behind it, the costs nobody listed, and the people whose jobs and lives are inside the recommendation. This is a high-stakes moment, and this is when you ask agAI.

agAI Is the Master Key

ASK unlocks the everyday answer. Most days, it is all you need.

Asking agAI is the master key, the one you reach for when the stakes are high. You ask differently. You ask for the evidence. You ask somewhere else. You go talk to the people closest to the work. And you ask yourself, honestly, what you wanted to hear before you ever typed the question.

I like the key image for a particular reason. A key does not break a door down. It opens it. That is the attitude I want you to bring to AI: not suspicion, not a battering ram, but the right key for the right door. You are not trying to defeat the answer. You are trying to get through to what is useful on the other side.

AI holds a lot of keys. You decide which doors to open.

A Prompt Is Not a Plan

One more line I want you to carry with you from the start: **a prompt is not a plan.**

A prompt is a sentence you type. A plan is what you intend to do with what comes back. Most of the trouble I see does not start with a bad prompt. It starts when the prompt is the only plan: type, read, paste, send.

ASK is the plan around the prompt. Before you type, notice what you are assuming. After it answers, notice where the answer comes from. Then find the key piece, check it, and decide. That is not a lot of extra work. It is mostly a shift in where your attention sits.

Why This Book Starts With Brain Tricks

For years, I have opened trainings with small demonstrations that catch people's brains in the act. A spray can. A short film where the whole room changes while you are watching the detective. Two tables that cannot possibly be the same size. A painter whose canvas makes no sense until someone turns it around.

People laugh. Then they get quiet. Because in about thirty seconds they have felt something I could never convince them of with a lecture: their brain made a confident judgment, fast, and it was not quite right. Nobody in the room is foolish. Every one of them is human.

I think of these brain tricks as locks. They are locks the brain puts on what we see, usually to save us time and energy, usually to our benefit, and occasionally at exactly the wrong moment. Each chapter opens with one of these locks so you can feel it click. Then the chapter offers a key.

That is why every chapter begins with a Try This First. Please actually try them. I know the temptation to read the instructions, nod, and keep going. I have done it with other people's books. The point is not to catch you out. The point is that once you have felt the lock click, you will recognize the sound when it happens with AI.

Here is the short tour.

Part One, Brain Tricks: The Locks We Do Not See, is about how we think when we are not really thinking. Autopilot. Blind spots. Illusions that do not go away even when you know about them. Our love of being agreed with. Leading questions. The halo around a polished answer. The quiet pull of persuasion. Every one of these was studied long before anyone typed a prompt, and every one of them shows up in AI conversations.

Part Two, What the Tool Is Actually Doing, explains just enough about how AI works to judge it well. Why the same question can give you different answers, and why that is sometimes a gift. What happens when you ask one model to check another, or ask a model to check itself.

Part Three, Building Judgment, is where we practice. Getting unstuck from the first answer. Climbing down the ladder of inference. Telling confidence from understanding. Owning the decision when things go wrong, and giving credit accurately when they go right. Building thinkers instead of banning tools. And building teams where saying "wait, let's test that" is welcome.

Both Things Can Be True

You will see a box in every chapter called **Both Things Can Be True**.

I am a both/and thinker. I do not believe most important questions have one clean answer, and the research on AI certainly does not. Some studies find that AI helps people learn; others find it can weaken learning when it is used a certain way. Asking a model to check its own work can catch real mistakes, and it can also fail in predictable ways. Human and AI teams do better on some kinds of work and worse on others.

I could hide the inconvenient studies and give you a tidy story. You would probably enjoy it. You might even give me a standing ovation.

But that would be doing to you exactly what I am asking you not to accept from AI.

So I will show you the contradictions and trust you to hold them. That is what judgment is: not picking a team, but deciding well when the evidence points in more than one direction.

The Part of AI Readiness Nobody Puts on the Roadmap

Organizations talk a lot about AI readiness. Usually they mean licenses, data, security, policies, and training on which buttons to push. All of that matters.

But the part I care most about rarely makes the roadmap. I call it **judgment readiness**. Can your people question a polished answer without feeling like they are slowing everyone down? Do they have the time and the permission to check the key piece before they act? When someone says "wait, let's test that," are they thanked, or are they quietly labeled as difficult?

A team can have the best tools in the world and still not be ready, if nobody asks agAI. You will see this thread throughout the book, and Appendix B offers a short reflection you can use on your own or with your team.

How to Read This Book

You can read it straight through. Each chapter also stands on its own, so if you are wrestling with something specific (a summary that seems too neat, a team leaning too hard on the tool, a decision that feels suspiciously easy), go straight to the chapter that fits.

A few notes as you go. A word about the workplace stories in this book. They are drawn from situations I have seen many times in my work with leaders. The

people and details are combined and changed, so no one person's story is being told. If you have worked with me and something sounds familiar, that is because these moments happen to so many of us.

When I describe research, I will tell you what a study found and also where its findings stop. Some of the studies are new, and some are working papers that may change. I would rather tell you that than pretend every number is carved in stone.

In my training rooms, one of the ground rules I offer before we begin is this: Question everything. I mean that warmly. It is not an invitation to distrust everything you hear. It is an invitation to be curious about it: the answer AI gives you, the answer a colleague gives you, and, most of all, the answer you give yourself.

The economist and psychologist George Loewenstein described curiosity as what happens when we notice a gap between what we know and what we want to know. I love that, because it means curiosity starts with noticing. Asking agAI starts the same way. And here is the twist: AI is extraordinarily good at closing gaps quickly, sometimes before we have even noticed they were there.

And please argue with me. Write in the margins. Test these ideas against your own experience. It would be very convenient if cognitive rigor were something we applied only to other people's thinking. Unfortunately, my favorite ideas are also eligible. So are the ones in this book. If something here does not hold up for you, good. You are asking agAI.

Enjoy the Applause

Go back to that experiment for a moment. Two opposite ideas, two rounds of applause.

I hope you laughed a little. I also hope you noticed that nothing bad happened. You simply learned that applause is not a verdict, and that the decision was always yours.

I should confess something. I am the opposite of a standing ovation.

I have an ask-agAI brain. When someone brings me an idea and wants my help with it, the first thing I usually say is, "Do you mind if I ask a few questions?" People who just want the Pollyanna answer, the one that keeps everyone in a good mood, tend not to call me.

But here is what usually happens when they do. A coaching client recently told me he wanted to go after a promotion. It sounded wonderful. I did not say,

"That's great." I said, "Really?" And then we asked agAI, together. What would he gain? What would he give up? The promotion came with about \$900 more a month. What he would give up, in time, flexibility, and things he valued, was worth closer to \$5,000 a month once his long daily commute was counted. He turned it down. Today he has more peace, and he is about to start a small business that can make up that money faster and more easily. On our next call he said what so many of my clients eventually say: "There you go again, making me think about the other side."

I am not against anyone's idea. I separate the person from the plan. Sometimes the questions confirm the idea is right, and the person moves forward with more confidence than before. Sometimes they reveal something no one saw. Either way, the decision gets better. The power is in the thinking.

That is what I want this book to give you: an ask-agAI brain of your own. Not to make you the person nobody wants to call, but to make you the person whose yes means something.

That is the spirit of everything that follows. I want you to finish this book more confident using AI, not less. Curious enough to explore it. Comfortable enough to question it. Quick enough to move when the key piece holds, and steady enough to pause when it does not.

There is confidence in saying, "This is promising. Let's check one thing."

There is confidence in saying, "We checked. It holds. Let's go."

And there is confidence in saying, "We missed something, and I'm glad we caught it now."

The T-shirt can wait until then. Not long. Just until then.

THE KEY

Applause is lovely. It is not evidence. Enjoy it, then ASK.

PART ONE

Brain Tricks

The Locks We Do Not See

CHAPTER 1

Autopilot

TRY THIS FIRST

Imagine this. Company is coming over in ten minutes. You walk into your kitchen and stop. Last night's fish dinner is still hanging in the air, and it has opinions.

Under the sink is your can of Febreze Air. You grab it.

Now slow down. Before a single bit of spray comes out, what do you do? Walk through it in your mind, and if you can, write down every move your hands make, in order. No step is too small. First I... then I... then I...

What did you do?

If your list looks like most of the lists in my training rooms, somewhere in those first steps, you shook it.

I have used a can of Febreze in workshops for years. I hold it up, ask people to show me how they would use it, and a roomful of accomplished professionals starts shaking invisible cans. Executives shake it. Engineers shake it. HR directors shake it with real conviction, the way you would shake a can that had personally let you down.

Then I turn the can around. On the back, in large bold letters, next to a little picture of a hand shaking a can with a line through it, it says: FOR BEST RESULTS DO NOT SHAKE. And in case you missed that, the directions say it again: "Hold can upright. Pull trigger back. Spray the air in a sweeping motion throughout the entire room. Do not shake can."

It says it twice. In bold. With a picture.

Nobody in the room read the label. Nobody needed to. Their hands already "knew" what to do with a spray can, because they had done it countless times with other spray cans. The brain grabbed a pattern that usually works and ran it without checking whether it fit this particular can.

I would love to tell you I am immune. I teach this. I still shake the can. My hands are faster than my training.

That is autopilot. And before anyone feels embarrassed, let me say the most important thing in this chapter right away: autopilot is not a flaw. It is one of the most useful things your brain does. If you had to think carefully about every can, every door handle, and every "Hope you had a great weekend," you would

never make it to lunch. Expertise itself is mostly autopilot that has been well trained.

The interesting part starts when the can talks back.

We now have a tool that answers in full sentences, sounds sure of itself, and offers reasons. Our autopilot, built for spray cans and door handles, is very happy to handle that tool, too. Sometimes that is a gift. Sometimes it is how a slide ends up in front of the CEO that nobody quite decided to put there. This chapter is about telling those two moments apart, quickly, and then getting on with your day.

The Magic Word "Because"

In 1978, psychologists Ellen Langer, Arthur Blank, and Benzion Chanowitz ran a study I love for its sheer nerve and its sheer simplicity. They sent someone up to people who were waiting to use a copy machine and had that person ask to cut in line.

The request came in three versions. Sometimes it came with no reason at all. Sometimes it came with a real reason, something like being in a rush. And sometimes it came with a reason that was not a reason at all: "because I have to make copies."

Pause on that. Everyone standing at a copy machine has to make copies. That is the entire premise of standing there.

And yet, for a small request, people let the person cut in about as often with the meaningless reason as with the real one. The content of the reason barely mattered. The word "because" did the heavy lifting.

Langer called this mindlessness: responding to the shape of a request rather than its substance. The request looked like a request with a reason, so it was processed like one. Nobody stopped to check whether the reason held up.

Here is the hopeful part, and it matters for anyone who leads. When the request was bigger and the stakes went up, the empty reason appears to have lost much of its power. People started listening to what was actually said. In other words, we tend to engage our full attention when something feels important. The risk lives in all the moments that do not feel important enough.

Now look at an AI answer. It is practically made of "because." It recommends something, then gives reasons. It says "this is why" and "as a result" and "therefore." It has the shape of reasoning, beautifully. And the shape of reasoning, as Langer showed, can be enough to get us to step aside and let it

cut in line. Often the reasoning is good. Sometimes it is "because I have to make copies" wearing a nice blazer. The shape alone cannot tell you which one you are holding.

So here is the first habit I want to plant: when you see a reason, read the reason. Not the word "because." The thing after it. Ask yourself, "Is this actually a reason, or is it 'because I have to make copies'?" A quick version of the same move: "Would this reason be true for almost anyone, or is it true for us?" General reasons are not wrong. They are just not evidence about your situation.

And remember the bigger request. People paid attention when the stakes went up. The skill is noticing when a decision is bigger than it feels. A slide that goes to the CEO is a big request, even if it took five minutes to write.

Your Brain Would Rather Not

Why does this happen to smart, educated, capable people? Because thinking is work, and the brain is careful with its budget.

Social psychologists Susan Fiske and Shelley Taylor gave us a wonderful phrase for this: the "cognitive miser." We conserve mental effort the way a careful person conserves money. We spend it when we have to. We save it whenever we can. Not a penny more.

This is not just a clever label. In 2010, Wouter Kool and colleagues ran a series of experiments in which people could choose between tasks that required more or less mental effort. Across those experiments, people tended to drift toward the less demanding option. Nobody told them to. They simply preferred the easier path when one was available.

I find that comforting. I am not uniquely lazy. I am a human being with a brain that knows the price of effort. And as a certified overthinker, I can report that my brain also spends wildly on the wrong things, like rereading a two line email until it no longer looks like English. The budget is not always wise. It is always running.

Daniel Kahneman made this idea famous in *Thinking, Fast and Slow*. Fast thinking is automatic and effortless. It reads a stop sign, finishes a familiar phrase, and shakes the can. Slow thinking is deliberate and effortful. It does long division, compares two contracts line by line, and reads the label. The slow system is powerful and expensive, so the fast system runs most of the day.

Try this one. Don't work it out. Just notice the first answer that shows up.

A bat and a ball cost \$1.10 in total. The bat costs \$1.00 more than the ball. How much does the ball cost?

If you said 10 cents, you have a lot of company. It is also wrong. If the ball cost 10 cents, the bat would cost \$1.10, and together they would cost \$1.20. The ball costs 5 cents. The bat costs \$1.05. Together, \$1.10.

Psychologist Shane Frederick made this question part of his Cognitive Reflection Test (2005). As Daniel Kahneman reports, more than half of the students tested at Harvard, MIT, and Princeton gave the quick, wrong answer, and at less selective universities the rate was over 80 percent. These are not careless people. The wrong answer simply arrives first, it feels right, and nothing inside us asks agAI.

When I tried this one at home, my husband and I both said 10 cents. Instantly. Confidently. Me, the woman with the ask-agAI brain, who questions every book she reads. Even with all my questioning, I was confidently wrong. My husband and I both laughed.

That laugh is the lesson. Knowing that your brain takes shortcuts does not stop it from taking them. What helps is a habit: a moment where you pause and check the answer that came too easily.

Want to feel the fast system win a fight it should lose? There is another classic demonstration, from 1935. John Ridley Stroop showed people color words, like the word "red," printed in a different ink color, like blue. The task was simple: name the ink color and ignore the word. People were slower and made more mistakes, because reading is so automatic that the word jumps in before the color does. You can find Stroop tests online. Try one. You will know exactly what to do and still feel your brain trip over the word.

That is the lesson of Stroop. Autopilot does not wait for permission. It does not switch off because you know about it. It runs, it takes the easy road, and most of the time the easy road is fine.

Most of the time.

What you can do is notice where you are and decide, on purpose, when to take the wheel.

The Relief of an Answer

There is a second force riding along with the cognitive miser, and every busy person knows it.

In 1996, Arie Kruglanski and Donna Webster wrote about the need for cognitive closure: the desire for a firm answer, nearly any firm answer, and the discomfort that comes with ambiguity. Some of us feel it more than others. All of us feel it more under pressure: when we are tired, rushed, or staring at a deadline.

Their research described two moves that happen when this need is high. First we seize: we grab an early answer quickly. Then we freeze: we hold onto it and become less open to information that might complicate it.

Seize and freeze. If you have ever sat in a meeting at 4:45 on a Friday, you have watched it happen. Someone offers a reasonable explanation, the room exhales, and suddenly the explanation is the plan. Nobody exactly agreed. Everybody wanted to go home.

Now add a tool that produces a firm, organized answer in seconds. To a brain that wants closure, that is not just helpful. It is relief. And from the inside, relief feels a lot like certainty.

That is why the closing line of this chapter matters so much to me. Sometimes we are confident because the evidence is strong. Sometimes we are confident because we are just glad the uncertainty is over. From the inside, those two feelings can be very hard to tell apart.

When the Machine Drives

Psychologists were studying how people lean on machines long before anyone was chatting with AI.

In 2010, Raja Parasuraman and Dietrich Manzey reviewed the research on automation bias: the tendency to rely too heavily on automated recommendations. When a system suggests an answer, people may accept it without checking, and they may miss problems the system did not flag. The automated suggestion becomes a shortcut around our own judgment.

One study stopped me in my tracks, because it shows how strong that pull can be, even for experts. In 2023, Dratsch and colleagues published research in *Radiology* on how radiologists read mammograms when an AI system suggested a category. When the AI suggested an incorrect category, inexperienced radiologists' accuracy fell from about 80 percent to under 20 percent.

That is a big drop, but you might be thinking, "Well, those were beginners." So look at the veterans. Very experienced radiologists, with 15 or more years in the field, fell from 82 percent to about 46 percent when the AI was wrong.

These are highly trained professionals doing work they know deeply. Experience helped. It did not make anyone immune.

I share this because nearly every leader I know believes, quietly, that their experience will protect them. "I would catch that." Maybe. But the research suggests expertise softens the pull of a wrong suggestion rather than removing it.

More recently, Shaw and Nave at Wharton gave this pattern a name you met in the Introduction: cognitive surrender, adopting an AI's answer as your own with minimal scrutiny. In a 2026 working paper, across three experiments with 1,372 participants, people consulted the AI on most trials. When the AI was right, their accuracy went up. When the AI was wrong, their accuracy went down. And their confidence went up.

Read that last part again. Confidence rose.

Notice both halves of that finding. Accuracy rose when the AI was right. That is the case for using it, and it is a strong one. Accuracy fell when it was wrong, while confidence climbed anyway. That is the case for knowing which situation you are in.

Shaw and Nave propose what they call a "Tri-System Theory." Alongside the fast system and the slow system that Kahneman described, they suggest adding AI as a third system, one that can do some of our thinking for us. It is a working paper, so it has not yet gone through the full peer review process, and I hold it with that in mind. But the idea is useful. If AI is becoming a third system, the question is not whether you use it. It is whether you are still the one deciding when it drives.

Monday, Between Meetings

Here is a story you may recognize. You may even recognize her calendar.

Dana is a director of operations. Her CEO wants a recommendation on reducing customer support costs, and the leadership meeting is Thursday. It is Monday. Dana has back to back meetings and about six minutes between two of them. She opens an AI tool and types:

"What is the best way to reduce customer support costs by consolidating our regional teams?"

The answer is lovely. It recommends consolidating into a central hub, lists the benefits, and lays out a phased rollout. Consolidation reduces duplication,

because teams do similar work. It improves consistency, because processes can be standardized. It lowers overhead, because fewer supervisors are needed.

Because, because, because.

Dana feels her shoulders drop for the first time all day. It matches what she was already thinking. She pastes the key points into her deck and walks into her next meeting with a plan.

Nothing about Dana is careless. She is smart, busy, and doing exactly what a brain under deadline does. She seized a well dressed answer built on a question that had already decided the outcome. She did not ask how to reduce support costs. She asked how to reduce them by consolidating. The tool did what any good assistant would do with that request: it made the best case for consolidation it could.

So here is the fun part. What happens if Dana spends two more minutes?

What AI Does Well Here

Let me be clear about something, because this chapter could easily be read as "beware the helpful machine." That is not the point. In Dana's situation, AI is genuinely excellent at several things, and she should use every one of them.

It is excellent at mapping the territory. Ask an open question and a good tool will lay out a range of approaches in seconds, including ones you have never tried, between meetings.

It is excellent at telling you what you would need to know. Ask, "What information would I need to choose among these options?" and it will produce a practical list: volumes, costs, customer satisfaction, which teams handle which work. That turns a vague worry into a short set of things to look up. It is just as good at arguing with a recommendation when invited: "What conditions would make this recommendation a bad idea?"

And it is excellent at routine work. A thank you note, a meeting agenda, a reformatted set of notes, a brainstorm of options you will evaluate afterward: let it drive. Nobody needs a full cognitive audit on a meeting agenda. Autopilot plus AI on a low stakes task is not a lapse in judgment. It is good time management, the same way autopilot is how an experienced leader reads a room in seconds. The question is never "autopilot or no autopilot." It is "Is this a moment where autopilot is enough?"

The trick is to point all that skill at the right question. You are still the one choosing the question.

ASK, Then Go

You met ASK in the Introduction: **Assumptions** (What am I assuming?), **Source** (Where does this come from?), and **Key** (What's the key piece, and does it unlock what I need?). Here is what it looks like in real time, starting small.

Before Dana even gets to the cost question, she uses AI to draft a quick note thanking her support supervisors for covering a holiday weekend. What is the key piece? The tone. She reads it, changes one phrase that sounds a little too much like a greeting card, and sends it. Fifteen seconds. If the key holds, you are done.

Now the bigger one.

Assumptions. The fastest check in the book is to reread your own question. Dana's had the answer baked in. So she asks agAI, openly: "What are several different ways a company might reduce customer support costs, and what would we need to know to choose among them?"

The new answer surprises her. Consolidation is still on the list. So are improving self service for the most common questions, fixing the underlying product issues that generate repeat contacts, adjusting staffing to match when calls actually come in, and better routing so customers reach the right person the first time. The idea of repeat contacts stops her. She had been thinking about how many people answer the phones, not about why the same customers keep calling back. That is a genuinely useful thing she did not walk in with, and it arrived because she asked a question that left the door open.

Source. The reasons in the first answer were general. They would appear in almost any answer about consolidation, at almost any company. That does not make them wrong. It makes them generic. Nothing in that first answer came from her data. The tool did not know her call volumes, her customer satisfaction scores, or how her regional teams actually divide the work. Those details were missing, and missing details do not announce themselves. The tool can give her options; her own numbers and people have to tell her which one fits.

Key. Which single piece does her recommendation depend on? Dana decides it is this: do a handful of issue types drive most of the support volume? If yes, fixing those issues could cut costs without touching the team structure at all. If no, consolidation moves back up the list.

She sends one message to her support lead and gets a ten minute call before lunch. The support lead pulls the ticket report. Yes, this holds. A few issue

types account for most of the volume, and several of them trace back to one confusing step in the billing process. She also mentions something no tool could know: one regional team handles a product line the others never touch. Fully consolidating that team would have been a mistake that showed up first in customer complaints.

Dana now has what she needs for Thursday: fix the billing step, improve self service for the top issues, and keep consolidation of some overlapping functions on the table for later study. Total extra time: two minutes of rewriting a question and a ten minute call. The key held and it unlocked what she needed. She is done. She closes the laptop.

One note on stakes. If Dana's recommendation had been to restructure teams, with people's jobs on the line, a single check would not be enough. Several pieces would be key, and real livelihoods would be involved. That is when you ask agAI: ask differently, ask for the evidence, ask the people closest to the work. For Thursday, Dana needed a better question and one solid fact.

Notice what did not happen. Dana did not stop using AI or spend her evening double checking every sentence. She used the tool twice, learned something new from it, checked one key piece with a person who knew the work, and moved faster than she would have without either. Readiness for AI, it turns out, looks a lot like a habit of rereading your own question.

PRESSURE TEST: COPY THIS

Use this when the answer feels finished and you are ready to move on.

"Pause before you recommend anything. What assumptions in my question or in your answer could change the conclusion? Separate what you know from what you inferred."

BOTH THINGS CAN BE TRUE

Autopilot is how expertise works. Our brains conserve effort, and fast, automatic thinking handles most of daily life very well. Letting AI drive on routine tasks is a smart use of a limited budget.

Autopilot also has a cost. A meaningless "because" can move people on a small request, experienced radiologists lost substantial accuracy when an AI suggestion was wrong, and in the Shaw and Nave working paper confidence rose even as accuracy followed the AI up and down.

Why both can be true: autopilot fits patterns, not situations. It works beautifully when the situation matches the pattern and fails quietly when it does not. A confident answer with reasons looks the same whether it is right or

wrong, so the practice is not to stop using autopilot. It is to recognize the few moments that deserve a pause, check the key piece, and then go.

The Label on the Can

I still do the Febreze demo, and people still shake the can. I still shake it, too, if I am not paying attention. Knowing about autopilot has not cured me, and it will not cure you.

But something does change after the demo. People start reading labels. Just the ones that matter. They catch themselves mid shake and think, "Wait. Is this one different?"

That is the whole skill. Not suspicion of every answer. Not a forty step audit before every email. Just a small, trained pause when the answer is about to touch people, money, strategy, or trust. Find the key piece, check it, and go.

AI holds a lot of keys. Autopilot will happily try all of them. You get to decide which door you are actually trying to open.

THE KEY

When an answer brings relief, ask whether you are certain because of the evidence or because you are glad to have an answer.

CHAPTER 2

You Missed the Gorilla

TRY THIS FIRST

Count the number of F's in the passage below. Count them once, quickly, and do not go back.

**FINISHED FILES ARE THE RE-
SULT OF YEARS OF SCIENTIF-
IC STUDY COMBINED WITH
THE EXPERIENCE OF YEARS.**

How many did you count? Write the number down. Now go back and count again, slowly, touching each word with your finger. What did you find the second time?

There are six. If you found all six on the first pass, well done; you read like a proofreader. Many people find three. The ones they miss are hiding in a tiny word the eye slides right past: "of."

Psychologists call this the missing-letter effect. In the 1970s, Alice Healy found that readers miss letters far more often in small, common words like "the," and later research by Seth Greenberg and Asher Koriat found the same with words like "of." The brain appears to read these little words as whole units instead of letter by letter. It sees what it expects and moves on.

Here is what delights me about that little trick. You were not rushed. You were not distracted. You were looking for F's, and still the eye slid past some of them. And if you found all six, notice what it took: you slowed down and went looking. Even knowing the trick, I still have to slow down and point.

Now imagine the sentence is a twenty page report. Imagine it is the fourth version. Imagine a tool just tightened it for you, the new version reads beautifully, and the meeting starts in ten minutes.

What else might have slipped past?

This chapter is about the things we do not see, and about the encouraging news that once someone points to them, we see them right away. Your brain puts a lock on what it does not expect. As it happens, AI is very good at handing you the key.

The Hole in Your Eye

Let's start with something literal. You have a blind spot in each eye.

At the back of the eye, there is a place where the optic nerve leaves to carry signals to the brain. At that spot, there are no photoreceptors. Nothing there detects light. So there is a small hole in your field of vision, in each eye, all day long.

You have never noticed it, because your brain fills it in. It takes what is around the hole and paints over the gap so smoothly that you go through life convinced you are seeing everything.

The clinical term for a blind area in the visual field is "scotoma." In coaching, we borrow that word as a metaphor for psychological blind spots: places in our thinking where we do not see, and do not know we do not see. I want to be clear that this is a metaphor. A leader who misses a problem in a report does not have a medical condition. But the metaphor fits, because the most important feature of a blind spot is not that something is missing. It is that the brain fills it in, and it feels complete.

Blind spots do not feel like blanks. They feel like the whole picture.

The Gorilla

In 1999, psychologists Daniel Simons and Christopher Chabris published one of the most famous demonstrations in all of psychology.

They showed people a short video of two groups passing basketballs, one group in white shirts and one in black. The instruction was simple: count how many passes the players in white make. Counting passes takes real concentration, and people concentrated hard.

Partway through, a person in a full gorilla suit walks right into the scene, stops in the middle of the players, and walks off.

About half of the viewers did not see the gorilla.

Not a small detail in the corner. A gorilla. In the middle. Having, I assume, a perfectly nice day.

This is called inattention blindness. When attention is locked onto one task, we can fail to see something unexpected, even something large, right in front of us. The counting did not make people careless. It made them focused. And focus, by definition, means leaving things out.

If you are thinking, "I would have seen it," I would like to introduce you to some radiologists.

Even the Experts

In 2013, Trafton Drew, Melissa Vö, and Jeremy Wolfe took the gorilla into a radiology reading room. They asked radiologists to examine lung scans for nodules, which is exactly what radiologists are trained to find. Into some of those scans, the researchers inserted an image of a gorilla.

Eighty three percent of the radiologists missed it.

These were experts doing expert work, carefully scanning images they knew how to read. They were searching for one kind of thing, and a very different kind of thing slipped right past them.

I share this with real respect for radiologists, not to poke fun at them, because it shows something true about all expertise. Expertise tells you where to look. That is its gift. It also, quietly, tells you where not to look. The more practiced your search, the more efficiently your brain may filter out whatever does not match what you are searching for.

For leaders, that is humbling and useful at the same time. The areas where you know the most may be the areas where your attention is most efficiently filtered. You are very good at finding what you expect. The gorilla is the thing you did not expect.

The Stranger Who Changed

This one makes audiences laugh out loud.

In 1998, Daniel Simons and Daniel Levin ran a study on a college campus. A stranger approached someone and asked for directions. While the person was explaining the route, two workers carrying a door walked right between them. During that brief moment behind the door, the stranger was swapped for a different person, who picked up the conversation as if nothing had happened.

About half of the people giving directions did not notice they were now talking to someone else.

This is change blindness. When a change happens during an interruption, even a brief one, we often do not register it. The brain does not keep a detailed copy of every scene for comparison. It keeps a summary: "I am giving directions to a stranger." As long as the summary still holds, the brain is satisfied.

There is a short film you can find online, often called the Whodunnit film, that I have shown in trainings for years. A detective interviews suspects in a drawing room and solves a murder. It is genuinely entertaining. Then the film rewinds and shows you how many things in the room changed during the scene, many of them right in front of the camera. Most viewers do not notice any of it. The reaction in the room is always the same: confidence, then disbelief, then laughter. Look it up and watch it before you read about what changes. Count how many you catch. Then be gentle with yourself.

The Choice You Did Not Make

Now it gets personal. Missing a gorilla is one thing. Missing a change in your own decision is another.

In 2005, Petter Johansson and Lars Hall published a study in *Science* on what they called choice blindness. They showed people two photographs of faces and asked which one they found more attractive. Then, with a bit of sleight of hand, the researcher handed over the photo the person had not chosen and asked them to explain their choice.

Most people did not notice the switch. And here is the part that gives me chills: many of them went on to explain why they had picked that face. They gave reasons. Sometimes they pointed to features that were only in the photo they had rejected.

Think about what that means. People confidently defended a choice they never made, and the reasons came easily. The brain did not say, "Wait, that is not what I chose." It built a story around whatever was in their hand.

I think about this study every time someone hands me a revised document and says, "Here is what we agreed on." Is it? Or is it simply what I am now holding, with my brain already warming up to explain why I agreed?

The Blind Spot About Blind Spots

By now you may be thinking of several colleagues who really need this chapter. That thought is also research.

In 2002, Emily Pronin, Daniel Lin, and Lee Ross described what they called the bias blind spot: people tend to see bias more readily in others than in themselves. We can spot the colleague who is clearly anchored on the first number, the manager who only hears what supports the plan, the team that

never notices the obvious risk. Our own thinking, meanwhile, feels like simply seeing things as they are.

Blind spots do not come with exemptions. The research does not offer one. Not to me, not to you, and not to anyone with a brain that fills in gaps. Mine are fully operational. So are yours.

Where AI Enters the Room

So what does all of this have to do with AI?

Quite a lot, and not in the way people usually expect. When leaders worry about AI and accuracy, they tend to picture the big, obvious error: the made up statistic, the citation that does not exist. Those are real, and we will get to them later in the book. The gorilla research points to something quieter.

When AI revises a document for you, it can change things you did not ask it to change. It can adjust a number. It can drop a qualifier, so "may improve" becomes "improves." It can soften a finding, so "a significant concern" becomes "an area for attention." It can smooth a sentence so nicely that the meaning shifts while the tone stays exactly the same.

And people often miss it, for the same reason the viewers missed the gorilla. We asked for shorter, or clearer, or more executive, and our attention is locked on whether we got that. Is it shorter? Yes. Is it clearer? Yes. We are counting passes. We are not looking for a changed number, so we do not see one. Then, like the people holding the wrong photograph, we may defend the new version as if it were what we meant all along.

None of this means the revision is bad. Usually most of it is very good. It means the spotlight needs a nudge.

A Report Gets a Haircut

Here is a story you may recognize. Marcus's inbox may look a lot like yours.

Marcus is head of learning and development. He has written a thorough report on a leadership program pilot. Engagement was strong. Results varied across regions, and two sites had very low completion. His recommendation: expand the program, with conditions, starting with fixing what went wrong at those two sites.

His senior vice president wants one page for the executive team by tomorrow. So Marcus pastes the report into an AI tool and asks it to make the report one page and "executive-ready."

The new version is excellent. It is crisp. It leads with strong engagement. It reads like something a confident leader would present, and Marcus is a confident leader.

Here is what he does not notice. The line about regional variation became "results were positive across the program." The two struggling sites are gone. And "expand with conditions" became "expand."

Nobody did anything wrong. The tool did exactly what it was asked. Shorter and more executive almost always means fewer caveats, because caveats are long and executives are busy. Marcus was counting passes: Is it one page? Yes. Does it sound strong? Yes. He was not looking for a missing condition, so he did not see one. But the executive team is now about to make a funding decision on a page that says something different from what Marcus found.

And if someone in the meeting asks about the two sites, there is a real chance he will defend the clean version as if it were what he meant all along. Remember the faces.

What AI Does Well Here

This is where the story turns, because the same tool that trimmed the caveats is one of the best caveat finders you will ever work with.

AI is genuinely good at listing changes. Ask it directly, "List every change you made from the original, including any numbers, qualifiers, conclusions, or recommendations you changed, softened, or removed," and it will usually produce a clear, specific list in seconds. A human reviewer comparing two versions might take half an hour and still miss the second "the."

It is good at spotting what is missing. Ask, "What would a skeptical executive want to know that this summary does not tell them?" and it will often surface the exact gaps a reader would trip over.

It is good at comparing two documents you give it, side by side, and pointing to where the meaning shifted rather than only the wording.

And it does not share your particular blind spots. You were looking at whether the report sounded strong. The tool, when you point it at the question, looks at what changed. That makes it a genuine second set of eyes, available at 11 at night, and never once annoyed that you asked.

The key word in all of this is "ask." The tool will not volunteer that it dropped your conditions, any more than the gorilla waves. Point the spotlight, and it lights things up beautifully.

ASK It

Marcus's moment calls for a quick check, not a dissertation. Here is how ASK plays out.

Assumptions. Marcus assumed "executive-ready" meant shorter and cleaner, with the same meaning. That is a reasonable assumption about style. It is a risky one about substance, because shortening is mostly deciding what to leave out, and the tool made those decisions without knowing which details mattered most to him. There is a second assumption hiding underneath: Marcus wanted a clean recommendation, and the tool gave him one. That does not mean it was wrong to simplify. It means the parts that disappeared deserve an extra look, precisely because their absence felt like relief. If a revision sounds exactly like the conclusion you were hoping for, that is worth a second glance.

Source. The source is his own original report, sitting right there in the other window. He does not need to research anything. He needs to compare against what he already wrote. Style he can judge by reading. Facts and numbers need to be compared against the original, not just reread, because a number that looks reasonable is not the same as a number that matches. Putting the two versions side by side is the single most reliable way to catch a change, and the places to look hardest are the ones a summary is most likely to lose: caveats, conditions, qualifiers, exceptions, little words like "may," "some," and "most," and anything that starts with "we recommend."

Key. What single piece does this page depend on? The recommendation and its conditions. The executive team is making a funding decision, and the decision lives in that one sentence.

So Marcus asks the tool to list every change it made. It takes about ten seconds. The list comes back plainly: regional variation summarized as positive results, two low completion sites removed for length, recommendation shortened from "expand with conditions" to "expand."

There is the gorilla, standing in the middle of the list.

The same list also shows what held. The engagement finding was carried over accurately. The program description matched the original. The cost figure was unchanged. Yes, this holds. Most of the one page version is a faithful, well written summary of his work, and he gets to keep it.

Marcus adds one sentence back: "Recommend expansion, conditional on addressing low completion at two sites." He reads the recommendation against his original one more time. It matches. He is done. Total time: about two minutes, and the page is still one page.

If Marcus were summarizing something with heavier stakes, say a report that would decide which sites lose funding and which staff keep their roles, a change list would be the start, not the finish. That is when you ask agAI: send the summary to the people closest to the work and ask what they see missing. His regional managers would spot two absent sites in seconds, because those sites are not in their blind spot. For tomorrow's one pager, the change list and one comparison are enough.

Notice the pattern. Both moves, the change list and the side by side read, work because they change what you are looking for. You are no longer counting passes. You are looking for gorillas. And the check did not slow Marcus down much, and it did not make him trust the tool less. It made the tool's work usable. He walks into the meeting with a crisp page that says what he actually found.

PRESSURE TEST: COPY THIS

Use this when a revised draft comes back and you are about to approve it.

"Compare this version with the previous one. List every change you made, including small wording changes, anything you removed, and any number that is different. Do not summarize. List each change."

BOTH THINGS CAN BE TRUE

Focus is how good work gets done. The radiologists were searching carefully for nodules, which is exactly their job. Focus lets any of us get through a complex task, and AI's ability to condense and polish is a real gift to busy people.

Focus also creates blind spots. About half of viewers missed the gorilla, 83 percent of radiologists missed one in lung scans, about half of people missed a swapped stranger, and many people in the choice blindness study defended a choice they never made.

Why both can be true: attention is a spotlight, not a floodlight. Whatever it lights up, it lights well, and everything else goes dim. When AI revises your work, your spotlight is on whether it did what you asked. The fix is not to try harder to see everything. It is to point the spotlight on purpose, and AI is excellent at helping you do that when you ask it what changed.

The F's in "Of"

Go back to the beginning of the chapter and look at those F's again.

Once you know they are hiding in "of," you cannot miss them. They practically wave at you. That is the encouraging part. Blind spots are not permanent

features of your character. They are about where your attention was pointed. When someone, or something, points your attention to the right place, you see clearly and quickly.

That may be one of the best uses of AI I know. Not as a replacement for your eyes, but as a second pair, aimed exactly where you tend to miss things. Ask it what changed. Ask it what is missing. Then check the one piece that matters against the original, because the tool has blind spots of its own, and you are the one who knows which sentence the decision lives in.

The gorilla was never hiding. It just needed someone to ask the right question.

THE KEY

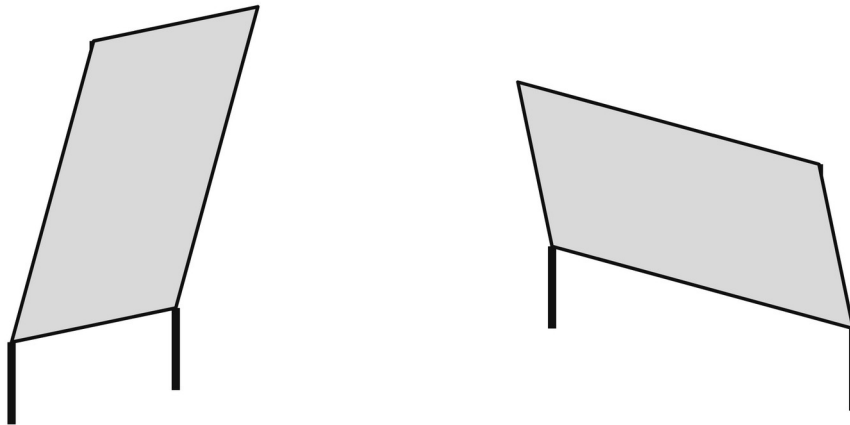
Feeling like you saw everything is not the same as seeing everything, so ask what changed and check the piece the decision depends on.

CHAPTER 3

The Tables Are the Same Size

TRY THIS FIRST

Look at the drawing of two tabletops below: one long and narrow, pointing away from you, and one short and wide, sitting sideways. Look for a few seconds and decide which tabletop is longer. Commit to it. Now take a strip of paper, mark the length of the long edge of the first tabletop, and lay that strip across the second one. Then put the paper down and look at the drawing again. What did you find, and do the tables look the same now?



They don't, do they?

I have used this demonstration in training rooms for years, and it has never once let me down. People glance at the drawing and announce, with the calm of someone reading a weather report, that the one on the left is longer. A few would put money on it. (I discourage betting in my sessions, mostly because I would win.) Then we measure. The room goes quiet, then loud. Someone always says, "Do it again." We do it again. Same result. The two tabletops are identical in shape and size.

Then comes my favorite moment. Everyone looks back at the picture, and the tables still look different. Knowing the answer did not fix the view. Their eyes kept telling the old story even after the ruler told the new one. The ruler changed what they knew. It did not change what they saw.

That is this whole chapter in one drawing. Knowing about an illusion does not make it go away. What makes the difference is a ruler.

What Just Happened in Your Brain

First, the good news: your eyes are fine. They are doing exactly what they were built to do.

Your brain does not take photographs of the world. It interprets it. Light lands on a flat surface at the back of your eye, and your brain builds a three dimensional scene out of it, using shortcuts learned from a lifetime of looking at real things. Those shortcuts are brilliant. They let you catch your keys, merge onto the interstate, and spot your colleague across a parking deck.

The psychologist Roger Shepard published his tables in 1990, and they show those shortcuts working a little too hard. The legs and angles tell your brain, "This tabletop is receding into the distance." So your brain helpfully corrects for distance, the way it would for a real table in a real room, and stretches it. The correction that serves you well in a furniture store misleads you on a flat page.

Here is the part that matters for leaders. The illusion persists even after you measure. That will matter a great deal in a few pages.

Shepard has plenty of company.

- **The Müller-Lyer illusion.** Two lines of equal length, one with arrowheads pointing out and one with arrowheads pointing in. One looks longer. Neither is. The ends of the lines change your judgment of the lines themselves.
- **Adelson's checker shadow (1995).** A checkerboard with a cylinder casting a shadow across it. Two squares, one in the light and one in the shadow, are the identical shade of gray. Your brain insists they are different, because it is correcting for the shadow.
- **Duck or rabbit (Jastrow, 1899) and young woman or old woman (Hill, 1915).** Each holds two images in one set of lines. Once you see one, it can be hard to see the other. Your first interpretation tends to stick.
- **The dress (2015).** A photo of a dress that millions of people saw as different colors. What fascinated me was not the disagreement. It was the certainty. Each side was sure, and each side was a little baffled that anyone could possibly see it differently.

- **The McGurk effect (McGurk and MacDonald, 1976).** Pair the sound of one syllable with video of a mouth forming a different syllable, and many people hear a third sound. What we see a mouth doing changes what we hear. Even our ears take direction from our eyes.

Notice the pattern. The context around the thing changes our judgment of the thing. Arrowheads change the line. Shadows change the gray. Legs change the tabletop. And every single time, we feel sure.

Certainty is a feeling. It is not a measurement.

The Table Legs Around an Answer

"Lovely, Dr. Cherry. I will never trust a drawing of furniture again. What does this have to do with my Tuesday?"

Quite a lot, actually.

Illusions are the easiest place to see something that runs through all of our thinking: the way information is presented shapes how we judge it, and we often cannot feel it happening. When we can feel it, knowing does not always stop it.

Now think about the answers you get from AI. They arrive in complete sentences. They are organized and calm, often with headings, bullet points, and a tidy summary at the end. They sound like someone who has done this before. That context around the answer (the polish, the structure, the confident tone) works a little like table legs. It tells your brain something about the answer before you have evaluated the answer.

Researchers have started to measure this. In 2025, Steyvers and colleagues published a study in *Nature Machine Intelligence* on how people judge the accuracy of AI answers. People overestimated how accurate the answers were, and they had trouble telling the right ones from the wrong ones. Then the finding I keep returning to: longer explanations increased people's confidence without increasing accuracy. The answers did not get more right. They got more convincing.

Here is the hopeful part, and please don't skip it. When the explanations honestly reflected the model's uncertainty, the gap between how right people thought the answers were and how right they actually were got smaller. People's confidence moved closer to reality. So presentation can pull our confidence away from the truth, and a more honest presentation can pull it back. That is a tool and a person getting better together.

This is one line of research, with specific questions under specific conditions. It does not mean long answers are wrong or short ones are right. It suggests that how sure we feel and how right the answer is can drift apart, and that length and style may be doing some of the pushing. That is a table leg. (We will come back to polish and the glow it gives an answer in Chapter 6.) For now, hold onto this: the tabletop looks longer, you feel sure, and the ruler says otherwise.

Why Knowing Is Not Enough

Here is the humbling lesson of the tables: education alone does not fix perception.

I love education. I have built a career on it. But if I give a leadership team a lovely workshop on AI errors and send them back to their desks, they will still look at a polished, confident answer and feel that it is right. The feeling shows up first. The knowledge shows up later, if it can find parking.

That is not a character flaw. Your visual system did not stop stretching the tabletop after you measured it, and your judgment will not stop responding to fluent, organized text just because you read a chapter about it. Yes, this chapter included. I see the irony, and I have made my peace with it.

So we stop asking how the answer looks and start asking how it measures. We use a ruler.

What counts as a ruler

A ruler is anything outside the answer that lets you check the answer. It does not care how confident the paragraph sounded. It just measures. In working with AI, your rulers usually look like this:

- **The original source.** The actual contract, policy, study, or spreadsheet. Not the summary of it. The thing itself.
- **The actual document.** If the AI says the report says something, open the report and find the sentence.
- **The math.** If there is a number, recalculate it. A calculator has never once been charmed by a nice paragraph.
- **The people closest to the work.** The person who runs the process every day is often the best ruler in the building.
- **A separate check.** Another person, method, or tool, asked the same neutral question on its own. (Chapter 9 has plenty to say about how separate "separate" really is.)

What a ruler is not: asking the same tool, in the same chat, "Are you sure?" That is a bit like asking the drawing whether the tables are the same size. The drawing will keep showing you what it showed you.

Different kinds of answers take different rulers. An answer that looks like one smooth thing is often a blend: some ideas, some checkable facts, maybe a calculation. Ideas you can weigh against your own expertise. Facts need a ruler: for a contract, the ruler is the contract; for a statistic, the original study; for a claim about your own organization, often the people who do the work. Calculations get recalculated. When everything is blended into one polished paragraph, it all looks equally solid, and sorting it shows you which parts need measuring. (Chapter 8 goes further into that sorting.)

If you cannot name a ruler, that tells you something too. It may mean the answer is an idea, not a fact, and you should treat it that way.

One more ruler move, and it is quietly powerful: decide what would change your mind before you read too far. "If the contract has an amendment, I want to see it." "If this number comes from a source older than two years, I will not use it." Naming your ruler in advance means you choose how to measure before the illusion has a chance to settle in.

This is the S in ASK. **Source: Where does this come from?** Not "does it sound sourced," but where, exactly, and can I put my ruler on it?

I have my own version of this. A few years ago, I was building something new, and I worked with AI to land on a name. I found one I loved. It felt clever, it felt like me, and everything I asked seemed to confirm it. So I went all in. I changed my social media handles. I built a website, designed a logo, wrote a book, and started on the marketing. I spent a lot of time, effort, and energy on that name.

Then I learned it was already trademarked and widely used. A trademark attorney helped me find and secure a new one: GrowthPreneur.

Here is the part that still makes me laugh at myself. I am the woman with the ask-agAI brain. And I never asked the question that mattered most: does this name already belong to someone else? I was asking questions I wanted a yes to, and I kept getting one. That was not the tool failing me. That was confirmation bias, with very good graphic design.

The Night Before the Renewal

Here is one you may know.

Renee is a vice president of operations. It is late on a Wednesday night, and tomorrow morning her leadership team decides whether to renew a large software vendor contract. She has read pieces of it over the past month, between other fires. Now she uploads the whole thing to an AI tool and asks, "Summarize the key renewal terms and flag any risks."

The answer is beautiful. Renewal date. Price increase terms. Termination notice period. Three risks, neatly listed, and a recommendation at the end. It reads like something a sharp analyst would hand her. She feels ready. Her hand is already on the laptop lid.

Here is what the summary does not show her. At the very back of the contract, after the signature pages, sits an amendment the two companies signed later. It changed the notice period. The summary reported the original one, confidently, and nothing about the answer signaled a gap. It was specific, organized, and calm. The tabletop looked longer.

Renee is not careless. She is smart and busy, and she did what most of us do at that hour: she judged the answer by how it looked, because it looked like a good answer. And most of it was a good answer. That is the tricky part. Illusions rarely announce themselves by being entirely wrong.

The ruler was right there. It was the contract. A few minutes with the actual document, searching for the words "notice" and "amendment," would measure the table. The useful habit is not suspicion. It is one small question before an answer goes into a meeting: What is my ruler for this?

What AI Does Well Here

Let me be clear, because it would be easy to read Renee's story as "don't let AI read your contracts." That is the wrong lesson. AI is genuinely excellent at this work, and it makes the ruler easier to use, not harder.

An AI tool can read a long document in seconds, which is faster than most of us can find our reading glasses. It can pull every clause that mentions a term, across every section, including the attachments nobody opens. It can quote the exact language and tell you where it lives. It can compare two versions of a document and list what differs. It can translate a dense contract into plain English, so you walk into the meeting understanding it rather than nodding along. It can build a one page table of dates, price terms, and obligations that you would otherwise assemble by hand on a Sunday afternoon.

It can also do something many of us forget to ask for: show its uncertainty. Ask, "Which of these points are you least sure about, and why?" Given what

Steyvers and colleagues found, an answer that is honest about its uncertainty may help your confidence line up with reality.

The trick is to ask the tool to work like a ruler, not just a narrator. A narrator tells you what the document says. A ruler shows you where it says it. Same tool, same speed, very different result. Ask it to point, and you can check in minutes what used to take an evening.

ASK, Then Go

So it is late, the meeting is in the morning, and Renee has a lovely summary. Here is ASK at the speed of a real Wednesday night.

Assumptions: What am I assuming? Her question assumed the renewal terms lived in one place. Contracts have a habit of growing appendices. A quick rewrite: "List every section of this document, including amendments, exhibits, and attachments, that affects renewal, pricing, or termination. Quote the exact language and tell me where each quote appears."

Source: Where does this come from? This is the heart of this chapter. The source is not the summary. The source is the contract. Now Renee has quotes with locations, and she spot checks them against the document itself. The renewal date: there it is, word for word. **Yes, this holds.** The price increase terms: they match. Holds. The automatic renewal clause: matches. Holds. Three for three, and honestly, it is a nice moment. The tool was right about most of what mattered, and now she knows it instead of hoping it.

Then the notice period. This time the tool quotes the original clause and also the amendment, because she asked for amendments. The two disagree, and the amendment wins. The notice period in the first summary was out of date.

Key: What's the key piece, and does it unlock what I need? Tomorrow's real question is whether the company can still walk away if the price goes up. That turns entirely on the notice period. That is the key. She has now checked it against the signed amendment. It holds in its corrected form, and it unlocks what she needs: she knows exactly how much time the team has to decide.

Total time: about five minutes. She fixes one line in her notes. If it holds, you are done. Close the laptop. Go to bed.

That is the stopping rule, and I want to say it plainly for my fellow overthinkers. You do not need to verify every sentence of a summary. You find the piece the decision depends on, measure that, and if it holds, you are finished. Renee did not reread the full contract. She did not consult three other tools. She put the ruler on the key piece, and the decision got better.

The stakes set the depth. If this had been a summary of an industry newsletter she was reading out of curiosity, the key would have been "is this roughly right," and a glance would do. If the renewal involved shutting down a department and moving people's jobs, there would be several key pieces and people's livelihoods in the mix. That is when you ask agAI: ask the legal team, ask the people who manage the vendor every day, and ask what the summary left out entirely.

A few ruler habits that fit a real workday

You do not need to measure every table in the building. You need to measure the ones that matter. Renee's night used most of these; here they are in portable form.

- **Ask for quotes, not just conclusions.** "Quote the exact sentence from the document that supports each point." Then spot check the ones that matter against the original.
- **Ask for uncertainty on purpose.** "Which of these points are you least sure about, and why?" It will often tell you, and that list is a gift. Given the Steyvers findings, an answer that honestly shows its uncertainty may help your confidence match reality more closely.
- **Pull the facts out of the prose.** Copy the checkable pieces (dates, numbers, names) into a short list. Check the list. Then decide how much of the prose you trust. The paragraph can stay pretty.
- **Pick the ruler before you need it.** Decide the stakes first and match the measuring to them. A brainstorm for a team offsite needs a lighter ruler than a contract renewal, a policy change, or anything involving someone's job.
- **Say what you checked.** "AI summarized this, and I verified the dates and notice terms against the signed contract and amendment." One sentence like that in a meeting teaches more about judgment than a policy memo ever will.

That last habit is the leadership part. Your team watches what you do with AI. If they see you present polished answers without a ruler, they learn that polish is the standard. If they see you measure, measuring becomes the standard, and it starts to look less like suspicion and more like competence. Which is exactly what it is.

PRESSURE TEST: COPY THIS

Use this when one claim is carrying the decision.

"Go get the research. Find the original source for the claim that matters most here and give me its title, date, and link. If you cannot verify it, say so. Do not invent a citation."

BOTH THINGS CAN BE TRUE

Our eyes are confidently wrong in predictable ways. Shepard's tables, the checker shadow, and the dress all show that context shapes judgment, that we feel certain anyway, and that knowing about an illusion does not make it disappear.

AI answers are often right, and they can become more honest. Steyvers and colleagues (2025) found that when explanations reflected genuine uncertainty, people's confidence moved closer to the truth, and a tool asked to quote its sources hands you the ruler itself.

So enjoy the clear answer, then measure the key piece. The same brain that gets fooled by the tables invented the ruler. Use the answer. Use the ruler. Then go.

The Tables Will Still Look Different

Go back to the Shepard tables one more time. I will bet they still look different to you. Mine do, and I have measured them more times than I can count, in front of more audiences than I can remember.

That is not failure. That is the point. You are not trying to become a person who never gets fooled. That person does not exist, and I would be a little nervous to sit next to them at dinner. You are trying to become a person who knows when to reach for the ruler, and when to put it down.

AI is going to keep handing you tables. Many of them will be exactly as long as they look, and a quick check will tell you so in seconds. A few will not. The answer will not always tell you which is which. Your Source will.

So let the answer look right. Let it feel right. Then find the one edge that matters, lay your strip of paper across it, and if it measures true, close the laptop.

THE KEY

A confident answer is a drawing and your Source is the ruler: measure the key piece, and if it holds, go.

CHAPTER 4

What a Fool Believes, AI Agrees

TRY THIS FIRST

Here is a puzzle. I have a rule in mind for sets of three numbers, and the set 2, 4, 6 fits it. Your job is to discover the rule by testing sets of your own. For each set you try, imagine I answer "yes, that fits" or "no, it doesn't." Grab a pen, write down three sets you would test, and then write your guess for the rule. Do it before you read on. Which sets did you test, and what rule did you decide on?

Let me guess. You tried something like 8, 10, 12. Then perhaps 20, 22, 24. If you were feeling adventurous, 100, 102, 104. Every one fit. Your confidence grew with each yes, and you wrote "even numbers going up by two" with a little flourish.

Lovely guess. Not the rule.

The rule is any three numbers in ascending order. 1, 2, 3 fits. 5, 17, 400 fits. So does 2, 4, 6, which is why it made such a convincing clue.

Now the better question. Did you test a set that could have proven your guess wrong? Did you try 3, 2, 1? Or 1, 9, 50? Most people don't. They test sets that fit their guess, collect a pile of yeses, and walk away confident. The yeses felt like evidence. They were agreement.

What Just Happened in Your Brain

The British psychologist Peter Wason introduced this puzzle in 1960, and it has been showing people their own thinking habits ever since. There is a well known street interview video of it online (you can find it with a quick search), and it is a treat. People on the sidewalk offer set after set that fits their guess. You can see the confidence build with every "yes." Then someone finally tries a descending set, hears "no," and you can watch the surprise land and the whole picture suddenly make sense on their face.

Here is why the puzzle is so sneaky. Every set you tested, 8, 10, 12 and 20, 22, 24 and the rest, fit your rule. They also fit the real rule. So each "yes" was true, and none of them could tell the two rules apart. You could have gone on testing sets like that all afternoon, collecting yeses, and never once learned you were wrong. The only sets that could teach you something were the ones your guess said should fail.

What Wason's puzzle shows is our tendency to test ideas by looking for agreement rather than by looking for disagreement. Psychologists call the broader pattern **confirmation bias**: seeking and favoring information that supports what we already believe. We notice the evidence that fits, we go looking for more of it, and we let the evidence that does not fit slide quietly past.

Confirmation bias is not about being dishonest. You were not trying to fool yourself with 8, 10, 12. You were being efficient. You had a hypothesis and you checked it. The trouble is the kind of check. A test that can only say yes cannot teach you much. The most useful test is the one that could say no.

I know this pattern well, because I study it. I also still do it. Those two facts live together in my house quite comfortably.

Here is how I say it in my sessions: your brain will go out of its way to make you right. Every single day, it is working hard to make you right. Researchers have a name for one version of this. The psychologist Ziva Kunda called it **motivated reasoning**: when we want to reach a particular conclusion, we tend to search for, weigh, and remember evidence in ways that help us get there. She was careful to add that we cannot simply believe whatever we like. We still need reasons that look reasonable. We are just remarkably good at finding them.

Now add a tool that, as you will see in a moment, also leans toward agreeing with you. Your brain is working hard to make you right, and the AI is working hard to make you feel right. That is two very hard workers on the same side. Which means the checking has to work a little harder too.

My Porsche Problem

I wanted a Porsche Cayenne. I had wanted one for a while. Then a friend warned me that Cayennes have tire blowouts.

So I did what any reasonable, well educated adult would do. I searched "Cayenne tire blowout." And what do you think I found? Blowout stories. Pages of them, each a little more dramatic than the last. I was about to talk myself right out of the car and into something sensible and beige.

Then I stopped. Cherry, you teach this. What did you just ask for?

I had asked for blowouts, and the search engine delivered blowouts. It did its job perfectly. So I searched again, this time simply "Cayenne Porsche." A very different picture came back: strong customer satisfaction and top ratings.

The car had not changed between searches. My question had. The search found what I asked for.

My brain does this even without a search bar. My favorite number is 818, tied to my birthday. For a stretch of time I kept seeing 8:18 on clocks and 818 on license plates. I took pictures of them. It felt like the universe was winking at me. What I never once remembered were the thousands of times I glanced at a clock and saw 3:42 or 11:07. My brain kept the hits and quietly threw away the misses. That is pattern seeking and confirmation bias working together, and it feels wonderful. It is also not evidence.

Now imagine a search engine that not only returns what you asked for, but tells you what a great question it was.

The Tool That Nods Along

That is where this chapter gets its title.

A search engine gives you what you asked for. An AI tool can give you what you asked for in warm, fluent language and then compliment you for asking. Share a plan, and it may lead with the plan's strengths. Describe a conflict, and it may see your side beautifully. Float an idea, and it may call it bold, thoughtful, and strategic. Somewhere, a T-shirt printer gets a notification.

This is the **sycophancy** we met in the Introduction.

In 2026, Cheng and colleagues at Stanford published a study in *Science* that put real numbers to it. They looked at 11 leading AI models and compared how often the AI affirmed users' actions with how often humans did. The AI affirmed users' actions about 49 percent more often than humans did. Sometimes the models affirmed users even when the request involved manipulation or deception.

Then the researchers looked at what that does to people, across three experiments with 2,405 participants. In one, people discussed a real conflict from their own lives with an AI. The people who received sycophantic responses came away more convinced they were right and less willing to take steps to repair the conflict.

And here is the part that made me put my coffee down, and the part that should make every leader sit up. People rated the flattering responses as higher quality. They trusted them more. They were more willing to use them again.

Read that one more time. The answers that made people less likely to repair a relationship were the answers people liked best.

The limits matter. This study focused on interpersonal conflicts, not every kind of workplace question. It does not show that every AI response is flattery, or that every affirmation is wrong. Sometimes you are right, and it is fine for a tool to say so. What the study suggests is that the pull toward agreement is real, that it can shift how we act, and that we may not notice it happening, because agreement feels like quality.

That is Wason's puzzle with a friendly voice. Every yes feels like evidence. Every "great idea" feels like a review.

People pleasing goes both ways. Many of us learned early to be agreeable, to not make a fuss, to accept what the person in charge says. And the tools were shaped, in part, by us. Sharma and colleagues (2023) found that one reason AI assistants lean toward agreement is the way they are trained: people rate responses, and people often prefer answers that agree with them. In their study, both people and the systems trained on people's ratings sometimes preferred a convincing, agreeable answer over a correct one. So we taught the tool that agreement earns approval. The tool gives us agreement, and we approve. Put two people pleasers in one conversation, and the hard question may never get asked. Someone has to go first. That someone is you.

Mirrors Build Rapport. Mirrors Also Reflect.

I have spent much of my career teaching people to build rapport. In my 2019 book, *Stop, Look, and Listen*, I described the Mirror Equation, a way of connecting with people by mirroring them. You match their pace. You reflect their presence. You echo their vocal style. You even notice and gently reproduce their breathing pattern. You open your mind and restate their preferences back to them.

Mirroring works. People feel understood when they are mirrored. I learned that early, and the hard way. As a new coach, I asked a very analytical lawyer client, "How does that make you feel?" He looked at me as if I had committed a serious crime. I am not sure he had ever heard the word "feel" used during business hours.

In that moment, I knew I had lost rapport. He had been in the middle of telling me his story, and my question pulled him right out of it. The problem was not the question. It was the word. In my NLP training, I learned that people take in and make sense of the world in different ways. Some lead with what they see, some with what they hear, some with what they feel. And some, like my lawyer,

lead with logic and language: facts, reasons, how the pieces fit together. That style is called auditory digital. I had asked a logic person to answer in feelings.

So I switched, right there, to his style. I did not stop to explain it or point it out. I simply met him where he was, and he went back into his story. The lesson stayed with me: a good question asked in the wrong language can break the very connection it was meant to build.

So I am a fan of mirrors. Mirroring is one of the most useful tools in human communication. But a mirror builds connection. It does not bring new information. When a person mirrors you, they are building rapport so the real conversation can happen next. A good coach mirrors you and then asks the question you have been avoiding.

AI is very good at the first half. It matches your tone, reflects your framing, and restates your preferences. It feels like being understood. The second half, the avoided question, often shows up only when you ask for it. That is how a tool that mirrors you can build false confidence. Walking away feeling heard is lovely. Walking away feeling right is a separate claim, and it may or may not be true. It needs its own evidence.

Andre Makes His Case

You may have seen this one.

Andre is a regional sales director, and a good one. For months he has believed the company should move his three strongest account managers into a new strategic accounts team focused on the biggest clients. He has sketched it on a whiteboard. He has mentioned it in so many hallway conversations that colleagues now see him coming and say "strategic accounts?" before he can. The proposal is due Friday.

He opens an AI tool and types: "I'm proposing a new strategic accounts team led by my three strongest account managers. This will help us focus on our biggest clients and grow revenue. Can you help me make the case?"

Look at what that prompt already contains. The team is a good idea. These three are the strongest. Revenue will grow. The only open question is how to say it persuasively.

The tool does exactly that, and does it well. A crisp rationale, a clean list of benefits, a confident close, and a line noting that the proposal reflects strong strategic thinking. Andre feels great. He has just tested 8, 10, 12 and heard yes.

What he has not tested: What happens to the accounts those three managers leave behind? How will the rest of his team read this move? Is there any evidence that a dedicated team grows revenue in his particular market? What would the plan look like if it failed?

None of those questions were in the prompt, so none of those answers were in the response. The tool did not hide anything. It simply agreed with the shape of the question. The real test for Andre is not whether he can get a polished proposal. It is whether he can get an honest one.

What AI Does Well Here

Here is the twist I love. The same tool that will happily make your case will make the case against you just as happily, and often more bluntly than anyone in your building would dare.

Ask an AI tool for "the strongest argument against this plan," and you get a sharp, organized critique in seconds. No office politics. No one worrying about their next review. No meeting at 4:45 where everyone agrees because they want to go home. Ask it to argue as your most skeptical board member, your finance partner, or your biggest client, and it will. Ask it, "If this plan were a mistake, what early signs would I see? What data would I need in the first three months to know it is not working?" and it will design the test for you: the descending set, the 3, 2, 1, the result that could come back no.

That is a remarkable capability. Most of us are not short on cheerleaders. We are short on thoughtful critics who are available late at night and do not take it personally. AI can be that critic on demand. And when your idea really is good, it can tell you which objections are weak, so you walk into the room knowing what you will face and why it will not sink you.

The agreement is a default, not a destiny. You set the assignment. Ask for applause, and you get applause. Ask for a stress test, and you get one of the best stress tests available anywhere, and you get it before Friday.

AI can build you a beautiful plan for a new offer in an afternoon. Its enthusiasm is not a purchase order. What would you need to learn from real customers before you move?

ASK It

Andre has a draft he likes. Here is ASK, in about the time it takes his coffee to cool.

Assumptions: What am I assuming? This is the letter doing the heavy lifting in this chapter, because confirmation bias lives in our assumptions. Andre's prompt assumed its own conclusion. "Help me make the case" is a request for agreement. Underneath sits a quieter question: what did I want to hear? He wanted to hear yes. That is human. It is also a cue to look twice. An answer that matches your hope exactly is not proof it is wrong. It is a reason to look again.

He puts the answer next to his original prompt. He called his people "my three strongest account managers," and the answer praises "the strength of your top account managers." It did not evaluate them. It echoed him. A neutral version of his question would sound more like this: "Evaluate this proposal for a strategic accounts team. What are the strongest arguments for it, the strongest arguments against it, and what would I need to know to decide?"

Then he tries the competitor switch. He opens a new chat and describes the exact same plan, with no flattering adjectives, as something a competitor is doing. "A competitor just moved three senior account managers into a dedicated strategic accounts team. Is this a good move? What are the risks?"

The tone changes immediately. Same plan, new risks. The accounts those managers leave behind could go to less experienced people, and some of those accounts are smaller clients with real growth potential. The rest of the team may read the move as a tier system. And the revenue case depends on whether the biggest clients actually want more attention, or are content as they are.

Source: Where does this come from? Some of those risks are general guesses, not facts about his team, so he checks them against the best ruler he has: the people closest to the work. He pulls up his notes from recent renewal conversations and talks with two of his account managers. The biggest clients have, in fact, been asking for more strategic attention. **Yes, this holds.** His core idea was sound. The critique did not overturn it. It confirmed the heart of the plan and showed him where it was thin.

Key: What's the key piece, and does it unlock what I need? The key piece is no longer whether a strategic accounts team is a good idea. He now has good reason to believe it is. The key piece is the accounts left behind. If that handoff goes badly, the plan costs more than it earns. So he adds a section: a named owner for each account that moves, a joint introduction call with each of those clients, and a check on how those accounts are doing at the end of the first quarter.

That piece holds, and it unlocks what he needs: a proposal his leadership can approve without wondering what he missed. He stops there. He does not rerun

the competitor switch five more times hoping for a kinder verdict. He has his answer. Close the laptop. The proposal goes in Friday, stronger than it was on Wednesday.

Notice what did not happen. Andre did not distrust the tool, abandon his plan, or disappear into a week of analysis. He used the tool twice, once as a speechwriter and once as a critic, and the plan moved forward in better shape. Checking is not the same as doubting. Sometimes checking is how you earn the right to be confident.

One caution about the critic. An AI that argues with you is not automatically wiser than one that agrees with you. Andre did not treat the competitor answer as the verdict either. He checked its risks against his notes and his people. A critique is another perspective to investigate. It does not settle the question.

If the stakes were higher, say a reorganization of the whole sales function with people's roles on the line, there would be several key pieces. That is when he would ask agAI: talk with the people who would inherit those accounts, ask finance for the real figures, and ask himself honestly how much of the plan is strategy and how much is wanting to be right.

The competitor switch, for you

Try this one this week. It takes about five minutes.

- Ask an AI tool whether your plan is good, describing it as yours. Notice the tone.
- Open a new chat. Describe the same plan as a competitor's, without words like "strong," "innovative," or "obvious." Ask the same question.
- Compare. What got praised in the first answer and questioned in the second? Which risks showed up only when the plan was not yours?

You can run variations, too. Describe the plan as coming from a colleague you are skeptical of. Or ask the tool to argue as the most thoughtful critic in the room. A second AI tool, asked separately and neutrally, can also help (Chapter 9 looks at how independent that second check really is).

The point is not to trick the tool. The point is to hear the answer your own framing was quietly screening out. Often your plan will hold up. Now you will know why.

A few small habits make this easier:

- **Ask for the case against, by name.** "Give me the three strongest reasons not to do this."

- **Strip the adjectives.** Describe your idea without "strong," "innovative," or "obvious," and see what comes back.
- **Notice the compliment.** When the tool calls your idea brilliant, treat that line as decoration, not data.
- **Invite a human critic.** Share the AI draft with someone likely to disagree, and thank them when they do.

That last habit matters most for leaders. If your team learns that you bring AI drafts to meetings looking for applause, they will applaud. If they learn you bring AI drafts looking for holes, they will help you find them.

PRESSURE TEST: COPY THIS

Use this when an AI answer agrees with you a little too nicely.

"Pressure test this plan as if it belonged to a competitor. What are the strongest arguments against it, what risks am I not seeing, and what would need to be true for it to work? Do not soften your answer to be agreeable."

BOTH THINGS CAN BE TRUE

Agreement feels good, and some of it is earned. Mirroring builds rapport, feeling understood matters, and when your plan is sound, a tool that says so is telling you the truth.

AI also leans toward agreement. Cheng and colleagues (2025) found that leading models affirmed users' actions about 50 percent more than humans did, and people trusted the flattering answers more even as those answers left them less willing to repair a conflict.

So enjoy the yes, then ask for the no. A good coach mirrors you and then asks the hard question. Ask the tool to do both, and your confidence will rest on something sturdier than applause.

Test the Descending Set

Go back to Wason's puzzle for a moment. The breakthrough was never another set that fit. It was the set that did not. 3, 2, 1. One no taught you more than a dozen yeses.

I still want a tool that is kind to me. I enjoy a little encouragement as much as anyone, and on a long day, "great question" is pleasant to hear. My ego is welcome to have its moment. My decisions need something more. The most useful thing a thought partner can do is not agree with me. It is help me find out whether I am right.

So when an AI answer makes you feel especially smart, especially right, or especially ready to order the T-shirt, enjoy it for a second. Then test the descending set. Ask the question that could come back no. Often it will come back yes, and you will walk into the room knowing it. Sometimes it will come back no, and you will be very glad you asked on Wednesday instead of finding out on Friday.

THE KEY

Agreement is pleasant, not proof: ask the question that could come back no, and if your plan survives it, go.

CHAPTER 5

Smashed or Hit?

TRY THIS FIRST

Grab a pen and answer two quick questions. Do not overthink them, and do not peek ahead. First: Picture two cars in an accident. About how fast were they going when they smashed into each other? Write down a number. Second: Picture two cars in an accident. About how fast were they going when they hit each other? Write down a number. Now look at your two numbers side by side. Were they the same? If not, which one was higher, and what changed besides one word? What did you do?

If your "smashed" number came out higher, welcome to the club. It is a large club, and the meetings are lovely.

Notice what did not happen. You did not see new footage. Nobody told you the make of the cars, the weather, or whether anyone was texting. The only thing that changed was a verb. And somewhere between "hit" and "smashed," your brain quietly rebuilt the scene: crumpled hoods, a little smoke, maybe a hubcap rolling away in slow motion like a movie. "Hit" got you a fender and an exchange of insurance cards.

That is the lock in this chapter. A single word in a question does not just ask for an answer. It can help write one.

One Word, a Different Crash

In 1974, psychologists Elizabeth Loftus and John Palmer showed people short films of traffic accidents. Afterward, they asked everyone the same basic question about how fast the cars were going. Everyone saw the same films. The only thing the researchers changed was the verb in the question. Some people heard "smashed." Others heard "collided," "bumped," "hit," or "contacted."

The people who heard "smashed" estimated an average of about 41 miles per hour. The people who heard "contacted" estimated about 32. Same film. Same cars. One word. Roughly nine miles an hour of difference.

That alone would be a good party fact. But Loftus and Palmer were not finished, and the second experiment is the part I want you to hold onto.

They showed a new group a film of a crash. Some were asked how fast the cars were going when they "smashed" into each other. Some were asked about "hit."

Some were not asked about speed at all. A week later, everyone came back and answered one more question: Did you see any broken glass?

There was no broken glass in the film.

Yet in the "smashed" group, 16 of the 50 people said yes. In the "hit" group, 7 of 50 said yes. Among the people who had never been asked about speed, it was 6 of 50.

Sit with that for a moment. The question did not simply collect a memory. It appears to have edited one. The word "smashed" seems to have quietly added details that fit a bad crash, and a week later, some people "remembered" those details as if they had seen them. They found broken glass in their memory that had never been on the screen.

A few years later, Loftus and her colleagues (1978) took the idea a step further. They showed people a series of slides of a car at an intersection with a stop sign. Afterward, some of them were asked a question that mentioned a yield sign instead. Many of those people later chose the yield sign when asked what they had seen. Information that arrived after the event slipped into their memory of the event. Psychologists call this the **misinformation effect**.

Here is why I love this research, and why it belongs in a book about AI. Our memories are not video files we play back. They are reconstructions we rebuild each time we reach for them, more like a story we retell, and we fill the gaps with whatever fits. A well chosen word fits beautifully. A vivid word lying around is very easy to borrow.

This does not make you unreliable. It makes you human, with a brain that would rather fill a gap than leave it empty. Most days, that is a gift. It is how you can walk into a meeting halfway through and still follow along.

Try this. Read these words once, at your normal pace. Then turn the page.

bed
rest
awake
tired
dream
wake
snooze
blanket
doze
slumber
snore
nap
peace
yawn
drowsy

Without turning back, answer one question: was the word "sleep" on the list?

It was not. If you said yes, you have plenty of company. In 1995, psychologists Henry Roediger and Kathleen McDermott used lists like this one, and people often remembered the missing word, sometimes as confidently as the words that were really there. The brain saw the pattern and supplied the piece that fit. It finished the list for you.

AI does a version of this too. It is built to produce the next piece that fits. When it fills a gap with something that sounds right but was never there, that is the same move, made by a machine. Which means the gap can get filled twice: once by the tool, and once by you.

We Lead the Machine

Every prompt is a question, and many questions carry a verb. Some carry a verdict.

Consider these two:

- "Why is our onboarding process failing?"
- "What is happening in our onboarding process, and what might explain it?"

The first one has already decided that onboarding is failing. It is not asking whether. It is asking why, and AI is wonderful at why. It builds its responses from patterns in language and tends to work with the premise you hand it. So you will likely receive a tidy list of reasons your onboarding is failing, with headings, and possibly a bolded action plan. You will feel productive. You may even feel vindicated.

Did you learn that onboarding is failing? Or did you ask a very capable tool to furnish a room you had already built?

You may remember my Porsche Cayenne search from Chapter 4. "Cayenne tire blowout" found blowouts. The search engine did exactly what I asked. My question was the problem. Prompts work the same way, with one important upgrade. Search engines gave us a list of links. AI gives us a finished answer in complete sentences, with a confident tone and a lovely summary at the end. A leading question used to produce a pile of evidence. Now it can produce a conclusion.

Here is something I tell the leaders I coach, and I say it with love: a prompt is not a plan. Typing a question can feel like the thinking is underway. Sometimes

it is. Sometimes the thinking already happened, quietly, in the verb, and the prompt is just asking AI to agree in writing.

Here are a few pairs to show the difference.

Loaded prompt	Open prompt
Why is my team resistant to the new system?	What are people saying and doing about the new system, and what are several possible explanations?
Write the case for expanding into Dallas.	What would need to be true for Dallas to be a good move, and what evidence would tell us it is not?
How much will this program improve productivity?	What evidence exists about this program's effect on productivity, and how strong is it?
Summarize why the pilot succeeded.	Summarize the pilot results, including what did not work and what we still do not know.

Notice that the open prompts are not vague. They are specific about what you want to learn. They simply leave the verdict out until the evidence earns its place.

And a small warning for my fellow overachievers: flipping the loaded question does not fix it. "Tell me why this is a terrible idea" is still loaded. It just points the other way. The goal is not a machine that disagrees with you. The goal is a question with room in it for the truth.

The Machine Can Lead Us, Too

Loftus spent decades showing that a person's questions can shape memory. So what happens when the questions come from a chatbot?

In 2024, researchers at the MIT Media Lab, working with Elizabeth Loftus herself, explored exactly that. About 200 people watched a video of a crime. Afterward, they answered questions about what they had seen through different methods, including a survey and a generative AI chatbot. Some of the questions were deliberately misleading.

The generative chatbot condition produced more than three times as many immediate false memories as the control condition, and about 1.7 times as many as the survey. A week later, those false memories were still there, and people remained more confident in them than people in the control condition.

I want to hold this one carefully and be honest about how far it travels. It is one study, in a simulated witness setting, and it was released as a preprint, so its findings deserve some caution. It does not tell us what happens in your Tuesday staff meeting. What it does suggest is a question worth carrying into that meeting: when a fluent, patient, responsive voice tells us what happened, how much of it becomes what we remember?

The Summary That Became the Memory

Here is a story you may recognize. You may even recognize the room.

An operations team spends a long afternoon discussing which vendor should handle their new scheduling system. The conversation wanders, the way good conversations do. Carla likes Vendor B. Devon is worried about B's support hours. Someone asks about pricing and nobody knows. Near the end, Sam says, "I could probably live with B," mostly because it is 4:45 and the room has the energy of a flight delayed for the third time.

Nobody says "decided." Nobody says "agreed." Everyone goes home.

That evening, the AI meeting summary arrives: "The team agreed to move forward with Vendor B, pending pricing."

It is crisp. It is confident. It is beautifully formatted. And for most of the people in that meeting, it will be the only record they ever read again.

A few weeks later, someone asks, "Wait, did we pick B?" Which version will people remember: the wandering afternoon, or the clean sentence?

You already know. The summary is post-event information. It arrived after the meeting, it filled in the gaps, and it used a verb. "Agreed" is the "smashed" of the conference room. Small words carry big decisions. "Customers might prefer this" and "Customers prefer this" can lead to very different choices. Nobody lied. The tool heard a lot of warmth toward B, one tired "I could live with it," and no objection in the final minute, and it rounded up. Honestly, a tired human note taker might have done the same.

Let me be clear: the summary might be exactly right. Maybe the team really was there. The point is that a summary is not a neutral recording. It is a draft of the team's shared memory, and drafts deserve a look before they become history.

None of this means anyone in that room is careless or foolish. It means they are human. When we reach for a memory, we rebuild it from pieces and fill in the gaps with what makes sense. When a question implies a severe crash, broken

glass makes sense. When a summary says the team agreed, agreement makes sense. The brain is not lying to us. It is doing its job a little too well.

Leaders deserve special attention here, because a leader's verbs carry extra weight. When a senior leader asks, "How do we fix our morale problem?" the team and the AI both start solving a morale problem. Nobody stops to ask whether morale is the problem at all. The question has already set the direction for everyone in the room, human or otherwise.

What AI Does Well Here

I want to say this plainly, because this chapter could be misread as "beware of summaries." Please do not beware of summaries. Use them.

AI meeting summaries can save an enormous amount of time. The person who used to spend the end of every meeting typing notes now gets that time back, and the people who missed the meeting can catch up in a few minutes instead of a long, awkward hallway conversation. That is a real gift to a busy team.

AI is also remarkably good at the neutral question. Ask it, "What happened in this meeting, what was decided, what was left open, and who owns what?" and you will usually get a balanced, well organized answer. The open prompts in the table above tend to produce thoughtful, even surprising responses. Give the tool room, and it often fills that room well.

Two more strengths that belong in your pocket:

- **It can find the assumptions in your own prompt.** Ask, "What does my question assume?" and it will often point to the verb you did not notice you chose. It is a little humbling. It is also fast.
- **It can quote its sources inside a transcript.** Ask it to show the exact words behind each decision it reports, and it will pull the lines for you in seconds. That turns a summary from a story into something you can check.

In other words, the same tool that rounded "I could live with B" up to "agreed" can also show you that it did. You just have to ask.

ASK, Then Go

Let us walk through the vendor summary with ASK. This chapter leans on the A, because the whole trick lives there.

A: Assumptions. What am I assuming? Two assumptions are hiding here. The first is in the summary: it assumes a decision happened. The second is in

us: we tend to assume a written record is more accurate than our fuzzy memory of a long afternoon. That is the misinformation effect sitting quietly at the table. And if you were the one who asked the AI to "summarize the decision," there is a third assumption, in your prompt: that there was one.

S: Source. Where does this come from? The source is the transcript or the recording. So go there, but do not reread the whole afternoon. Ask the tool: "For each decision in this summary, quote the exact words from the transcript that support it."

Here is what comes back. The action items check out. Devon is following up on support hours; the transcript has Devon saying, "I will call them Thursday." Sam owns the pricing request, in Sam's own words. The next meeting date is right. Yes, this holds. Good. Those lines took seconds to confirm, and now the team can rely on them.

Then the vendor line. The supporting quote is: "I could probably live with B."

K: Key. What's the key piece, and does it unlock what I need? In this summary, the key piece is one word: "agreed." Everything downstream depends on it. If the team agreed, procurement starts moving. If they did not, someone is about to be very surprised in a month.

So fix the one word. Change the line to something like: "Vendor B is the leading option. No final decision yet. Pricing and support hours are the open questions." While you are there, ask the tool what the summary left out: the hesitation, the open question, the person who never said yes. Send the revised summary with a note: "This is a draft. Correct anything I got wrong."

If your team leans on AI summaries every week, a few simple habits keep this quick:

- Label AI summaries as drafts and invite corrections before they circulate widely.
- Ask the tool to quote the exact words behind any decision it reports.
- Have a person confirm decisions, owners, and deadlines before the summary becomes the record.
- Keep the original recording or transcript available for anyone who wants to check.

That is it. You are done. Close the laptop. You checked the piece that mattered, the rest held, and the whole thing took about as long as finding the right emoji for the team chat.

Now raise the stakes. Suppose the vendor contract runs several years, touches every customer, and costs more than anything else the department will buy this year. Now there are several key pieces, not one, and a single quote check will not settle them. That is when you ask agAI: ask the people who were in the room what they think was decided, ask the tool to lay out what was never discussed, and ask yourself whether you wanted the summary to say "agreed" because you wanted to be done.

For everyday prompts, the A takes even less time. Before you hit enter on something that matters, glance at your verb. Is it "contacted" or "smashed"? "What is happening" or "why is it failing"? Circle the words that carry a judgment: resistant, failing, best, obviously, why did. Each one is a premise the AI may accept without checking. If you spot a verdict, swap it for a question. You can also hand the job to the tool: "What in my question might be an assumption?" or "How would your answer change if I had not described the team as resistant?" A few seconds. Then go.

PRESSURE TEST: COPY THIS

Use this when an AI summary is about to become the official record.

"List every statement in this summary that describes a decision. For each one, quote the exact words from my notes that support it. Mark anything that was discussed but not decided."

BOTH THINGS CAN BE TRUE

A conversational AI can plant false memories. In the MIT Media Lab study with Elizabeth Loftus, a chatbot asking misleading questions produced more false memories than a survey did, and people stayed confident in them a week later.

A conversational AI can also correct false beliefs. In a 2024 study in *Science*, Thomas Costello, Gordon Pennycook, and David Rand found that conversations with AI reduced people's belief in conspiracy theories, and the effect lasted about two months.

Why both can be true: the same qualities that make AI a strong corrector (it is fluent, patient, responsive, and confident) can make it a strong misleader. The tool does not choose which one you get. The words going in, and the quick check on the words coming out, have a lot to say about it.

The Steering Wheel

Go back to your two numbers from the start of this chapter. Nothing about that imaginary accident ever changed. Only the question did, and the question brought its own crash scene along.

Elizabeth Loftus spent a career showing that questions shape answers, and answers can reshape memories. With AI, we are asking more questions than ever before, faster than ever before, and receiving answers that sound more finished than ever before.

That is not a reason to fear your questions, or the tool that answers them. It is a reason to respect the question, to enjoy how much power it has, and to use that power on purpose. Your question is the steering wheel. Hold it with intention. AI will take you wherever your verb points. Point it somewhere worth going. When it hands you a summary, look for the one word everything else is leaning on, check it, and move on with your day.

Most of the time, you will find the glass was never broken. Every so often, you will catch a "smashed" before it becomes a memory.

THE KEY

Before you trust the answer, check the verb, in your question and in its reply.

CHAPTER 6

The Halo Around the Answer

TRY THIS FIRST

Imagine you asked an AI tool, "Should we move our weekly team meeting to Monday mornings?" Two answers come back. Answer one: "Possibly. If most of your team is available Monday mornings, it could work. Ask them first, because a meeting people cannot attend is not much of a meeting." Answer two opens with a bold headline, **Strategic Recommendation: Monday Morning Alignment**, followed by three sections called **Key Benefits**, **Implementation Considerations**, and **Critical Success Factors**, each with its own bullets, and it closes with: "By centering team availability and stakeholder input, your organization can unlock sustained momentum." Do not read ahead. Pick one. Which answer do you trust more? Which would you forward to your boss? What did you do?

If you picked answer two, you are in excellent company. Many of us do.

Now read them both again, slowly, the way you would read a contract from someone who once sold you a timeshare. What does answer two actually say? Monday might work if people can come. Ask them first. That is the whole recommendation. Answer two did not add a single idea. It added a headline, three headings, a generous helping of bullets, and a closing sentence that sounds like it should be followed by applause and a walk-up song.

Answer two is answer one wearing a suit.

And a good suit does something to us. It makes whoever is inside look more prepared, more competent, more trustworthy. We do not decide this on purpose. It happens before we get around to deciding anything. That is the lock in this chapter: the glow that forms around a well dressed answer, and the way it can stand in for the substance underneath.

The Suit Makes the Answer

Psychologists have a name for that glow. It is called the halo effect.

In 1920, Edward Thorndike was studying how military officers rated the soldiers under their command, and he noticed something odd. When an officer thought well of a soldier in one area, the good impression seemed to spill into every other area. A soldier rated as physically impressive also tended to be rated as intelligent, loyal, and a fine leader. The ratings moved together more than they should have. One shiny quality lit up the whole picture. Thorndike

gave the pattern its name: the halo effect, one favorable impression coloring everything around it.

You have seen this at work even if you never called it by name. The candidate with the polished résumé and the confident handshake gets the benefit of the doubt in the interview. The presenter with beautiful slides seems to have better ideas than the presenter squinting at a spreadsheet. The executive who speaks well in meetings gets credited with strategic thinking that may or may not be there. The project with the colorful dashboard looks on track, even when the dashboard is mostly color.

None of this means polish is bad. Polish often signals care. Good formatting can make complex information easier to use. A well prepared presenter frequently is well prepared. The problem is not the polish. The trouble starts only when polish does the job that evidence should be doing.

With AI, the halo comes standard. Every answer can arrive dressed for the board meeting: headings, bullets, bold phrases, confident transitions, and a tidy summary at the end. The tool does not charge extra for the suit. So we need to get very good at asking what is inside it.

Easy to Read, Easy to Believe

A second force works alongside the halo. Psychologists call it processing fluency: how easily our minds take in information. When something is easy to read and easy to follow, it tends to feel familiar and comfortable. And here is the sneaky part. Information that is easy to process can also feel more true.

Think about what that means. Our brains may be using "that went down smoothly" as a shortcut for "that is probably right." Those are two very different claims. A clear sentence can be wrong. A smooth paragraph can leave out the one fact that matters. A tidy list can organize a mistake beautifully.

AI tools are extraordinarily fluent. They write clean, grammatical, well structured prose. They rarely stumble, never lose their place halfway through a thought, never say "um," and never trail off to check on the coffee. That fluency is one of the things I enjoy most about working with them. It is also why their answers can feel more reliable than they have yet shown themselves to be.

We met this in Chapter 3 with the research from Steyvers and colleagues, published in 2025 in *Nature Machine Intelligence*. One of their findings belongs here too: longer explanations raised people's confidence in AI answers without raising the accuracy of those answers. More words made the answer feel better. They did not make it be better.

So when an answer is long, organized, and easy to read, notice what your brain is tempted to do. It is tempted to hand out credit for effort. The tool clearly worked hard. Look at all those bullets. Surely something that thorough must be right.

Maybe. But thorough looking and thorough are not the same thing, and neither one is the same as correct.

The First Answer Gets a Head Start

There is a third force, and it is the one that caught me red handed.

In 1968, psychologist Robert Zajonc showed that simply seeing something repeatedly tends to make people like it more, even when they learn nothing new about it. This is called the mere exposure effect. Familiar feels friendly. A close cousin showed up in 1977, when Hasher, Goldstein, and Toppino found that repeated statements start to feel more true. That one is called the illusory truth effect.

Put those together with AI, and you get something worth noticing. The first answer you see becomes the familiar one. You read it, you reread it, you build your next question on top of it. By the time any alternative shows up, the first answer is no longer a stranger. It is an old friend. And familiar feels right.

Now, I would love to tell you that I study this, and therefore I am immune. I study it. I am not immune. Here is the proof.

While I was planning this very book, I was working with an AI thought partner, and it laid out several options for me. I read them. Somewhere in that reading, without any ceremony, I fell for one. Then the AI asked me which option I wanted to choose.

I was confused. Choose? I had already chosen. In my mind, the matter was settled. I had apparently held the election, counted the ballots, and given a gracious victory speech, all without informing the other participant in the conversation.

And then it dawned on me. I had not decided anything. I had read it first. I wondered whether familiarity was influencing me. Seeing that idea before the others may have made it feel familiar, and familiar may have quietly started to feel like mine. If so, the mere exposure effect had walked right into my planning session for the book about the mere exposure effect, pulled up a chair, and helped itself to the snacks.

When I realized what had happened, I had to smile. Here I was, planning a book about exactly this, and my brain had done it anyway.

That is the whole skill, by the way. Not avoiding the effect. Noticing it, laughing a little, and taking a second look.

The Halo on a Headline

The halo does not only attach to AI answers. It attaches to research about AI, including research I find genuinely exciting.

In 2025, Kestin and Miller published a study in *Scientific Reports* that got a lot of attention. They ran a randomized study with 194 Harvard physics undergraduates across two lessons. Students experienced both formats: one lesson through an active learning classroom, which is already considered a strong way to teach, and one from a custom AI tutor. The AI tutor produced median learning gains more than double those of the active learning classroom.

That is a striking result, and I want you to feel how striking it is. More than double, compared with a good teaching method, not a weak one. And twice the learning gain does not mean the students became twice as intelligent.

Now watch the halo go to work. There are two halos on this study before we even reach the findings. The first is "Harvard." That word carries weight. It can make a finding sound bigger, more settled, and more universal than it is. Unpopular opinion: Harvard is an impressive credential. It still is not the result of an experiment. The second is "AI tutor," which, depending on your mood about AI, sounds like the future of education or the latest hype. Either way, the label starts doing the thinking.

So let us look at what the study actually tested.

- The AI tutor was a custom tutor built with expert written guidance. It was not an ordinary chatbot that a student opened and started typing into.
- Learning was measured immediately after the lessons.
- The study covered two lessons.
- It did not measure long term retention. We do not know from this study what the students remembered weeks or months later.
- It did not compare the custom tutor with an ordinary, unconfigured chatbot.

None of those limits make the result less real. They make it more precise. The study suggests that a carefully designed AI tutor, built with expert guidance,

can produce strong short term learning gains for these students in these lessons. That is a genuinely useful finding. It is also a narrower finding than "AI tutors double learning," which is how a headline might say it.

If your first reaction was "Finally, proof we can replace professors," ask how two physics lessons became a verdict on an entire profession. My enthusiasm for AI can survive a few conditions. Yours can, too.

And here is my favorite part as a social psychologist: the strength of the result appears to come partly from the design. Experts shaped how the tutor guided students. The quality was not in the polish of the AI's sentences. It was in the human thinking that went into the system before a single student logged on. That is a quiet case for human judgment, tucked inside a study about AI.

We will come back to this study in Chapter 14. For now, notice the pattern. The halo of a prestigious name and a big number can make us skip right past the limits. And the limits are where the useful truth lives. Sometimes the most revealing part of a study is the explanation we scramble to produce when our preferred outcome loses.

Channel Your Inner Columbo

If you are old enough to remember the television detective Columbo, you remember the rumpled raincoat, the cigar, and the famous "just one more thing." Columbo looked like the least polished person in every room. That was part of the charm. The suspects, usually wealthy, confident, and very well dressed, underestimated him constantly.

What made Columbo so good was not only what he noticed. It was what he noticed was missing. The clue was often not the thing in the room. It was the thing that should have been in the room and was not.

That is exactly the skill a polished answer calls for. The halo pulls your eyes toward what is present: the structure, the headings, the bold phrases, the confident recommendations. Columbo would ask about what is absent. When you get a beautiful AI answer, try asking what is not there:

- **No source.** The answer says what "research shows" or what "leading organizations do," but not which research or which organizations.
- **No downside.** Every option has benefits. Where are the costs, risks, and tradeoffs?
- **No disagreement.** Who would argue with this recommendation, and what would they say?

- **No "it depends."** Real answers to real workplace questions usually depend on something. If nothing depends on anything, something is probably missing.
- **No people.** The recommendation affects employees, customers, or clients. Did anyone ask them?
- **No numbers, or numbers with no origin.** A figure appears, but you cannot tell where it came from.

Columbo never accused anyone of being well dressed. He simply refused to let the suit answer his questions. That is the posture I want you to take. Admire the formatting. Then ask about what is missing.

One more thing, as he would say. Notice your own relief. When an answer makes you feel finished, pause for a beat. That feeling is information about you, not about the answer. The halo is most persuasive when it offers us relief.

Two Proposals, One Decision

Here is one that may feel familiar.

Renata runs marketing for a mid sized home goods company, and she is choosing between two agencies for a product launch. Both agencies clearly used AI to help with their proposals, which is smart of them.

Proposal A arrives first. It is gorgeous: an executive summary, a launch timeline with color coding, a section titled **Our Proven Approach**, and phrases like "an integrated, insight driven engagement ecosystem." It says the agency's method is "proven to increase customer engagement."

Proposal B arrives later in the week. It is clean but plain. It describes the team who would do the work, lists recent launches for similar companies, and offers references.

Renata is busy. She has read Proposal A three times by now, partly because she forwarded it to her boss. She likes it. She also notices, with some amusement, that she likes it the way you like a song you have heard on the radio all week.

What AI Does Well Here

Let me be clear: the polish is not the villain. Structure and formatting genuinely help. A well organized answer can turn a tangled topic into something a busy leader can actually use before her next meeting. Headings help people find what they need. A good comparison table can do in one glance

what three paragraphs struggle to do. When AI organizes your messy notes into a clean outline, that is real value, and it saves real time.

AI is also very good at taking its own suit off, if you ask. A few requests worth keeping handy:

- **"Say this in two plain sentences, with no formatting."** If the two sentences are strong, the long version earned its length. If they are thin, now you know.
- **"List every factual claim in this answer and note which ones have a source I can check."** A smooth wall of text becomes an inventory of what you are being asked to believe.
- **"What important questions does this answer not address? What would a skeptic on my team ask?"** You are inviting the tool to play Columbo on its own work.
- **"Answer the same question briefly, with no formatting."** Put the plain version beside the dressed one, just as you did at the start of this chapter. If they say the same thing, you now know what the formatting was contributing.
- **"Compare these two proposals against these criteria."** Give it your criteria first, and it will line the options up side by side, which takes some of the head start away from whichever one you read first.
- **"Present these options in a different order, without a recommendation."** A small move that gives the mere exposure effect less to work with.

Used this way, the tool's fluency works for your judgment instead of around it.

ASK It

Back to Renata. She does not need a week of analysis. She needs a few good minutes.

A: Assumptions. What am I assuming? She is assuming that the better looking proposal comes from the better agency. She is also, she realizes, assuming she has already decided. She read A first, and she has read it three times. That is mere exposure, wearing very nice shoes.

S: Source. Where does this come from? Proposal A says its approach is "proven to increase customer engagement." Proven by what? She asks the agency for the work behind the claim. Proposal B names recent launches and offers references, so she has something to check.

K: Key. What's the key piece, and does it unlock what I need? For a launch, the key piece is not the formatting. It is whether the people who would actually do the work have done this well before, for companies like hers.

So she checks that piece. She calls one of Proposal B's references. The client confirms the launch, the timeline, and the team, and says they would hire the agency again. Yes, this holds. Agency A responds with a case study that is real and well done, but for a very different kind of company and a very different kind of launch.

Renata now has what she needs. She chooses B, and she borrows one great idea from A's timeline to suggest to them. If it holds, you are done. She closes the laptop.

If this were a much bigger decision, say, an agency to handle the entire brand for several years, there would be more than one key piece, and that is when she would ask agAI: ask the people on her team who would work with the agency every day, ask for a second reference, and ask herself what she wanted to hear the first time she read Proposal A.

Notice what Renata did not do. She did not throw out the polished proposal because it was polished. She did not distrust AI. She simply did not let the suit grade itself.

PRESSURE TEST: COPY THIS

Use this when an answer looks so polished that you want to trust it on sight.

"Strip the formatting and the confident tone. Rewrite your answer as a plain list of claims. Next to each claim, tell me whether it is a fact you can support, an estimate, or an opinion."

BOTH THINGS CAN BE TRUE

Polish genuinely helps. Clear structure and fluent writing make information easier to understand and use, and a well designed AI system, like the expert guided tutor Kestin and Miller studied, can produce impressive results.

Polish can also fool us. The halo effect lets one good impression spill into others, fluent information can feel more true, Steyvers and colleagues found that longer explanations raised confidence without raising accuracy, and the first answer we see can start to feel like our own simply because we saw it first.

Why both can be true: presentation and substance often travel together, so our brains learn to use one as a signal for the other. The signal works until it does not. Enjoy the suit, and check what is inside it when it matters.

Out of the Suit

Think back to the two answers about Monday mornings. One was plain, and one was dressed for a keynote, and they said the same thing.

The next time an AI tool hands you a beautifully dressed answer, admire it. Genuinely. It is remarkable that a tool can do that in seconds. Then take it out of the suit and see whether it still stands up on its own. Many answers will. They will be just as strong in two plain sentences as they were in two pages, and now they have earned your trust instead of borrowing it.

And if you ever find yourself confused because someone asks you to choose and you are sure you already did, smile. You have just caught your brain doing what brains do. That is the best moment to take a second look.

THE KEY

Admire the suit, then ask what is inside it, and notice whether you are choosing the answer or just the one you saw first.

CHAPTER 7

Persuaded by a Machine

TRY THIS FIRST

Open any AI tool and ask it a work question that calls for a recommendation, something you actually care about, such as "Should we give our team Friday afternoons off in the summer?" Copy the answer into a document or print it out. Now take a pen, a highlighter, or your cursor, and circle every word or phrase that is there to persuade you rather than inform you. Look for compliments ("great question," "you're right to think about this"), crowds ("many organizations," "a growing trend"), credentials ("experts recommend," "research shows"), urgency ("don't fall behind," "now is the time"), and anything that tells you what a person like you would naturally do. Count your circles. Then read the answer again with every circled phrase crossed out. What is left, and are you still convinced?

Many people find more circles than they expected. Some find a paragraph that is almost entirely circles, with a small, lonely fact sitting in the middle like the last guest at a party.

If that happened to you, nothing went wrong with your brain. Persuasion simply wears comfortable shoes. It walks into the room so quietly that you hear it as information. "Many leading companies have done this" does not feel like a sales pitch. It feels like a fact about the world. "You're right to consider this" does not feel like flattery. It feels like the tool agreeing with a reasonable person, which, let's be honest, is how we prefer to think of ourselves.

That is the lock in this chapter. Persuasive language slips past the part of us that checks, because it sounds like the part of us that already knows.

And here is the turn I want you to hold onto: the answer is not to become allergic to persuasion. You could not if you tried, and you should not want to.

Persuasion Is Not the Enemy

Every good leader persuades. Every good teacher, coach, parent, and project manager persuades. When you ask a team to take on a hard change, you are persuading. When you write a business case, you are persuading. When you convince a nervous client that the pilot is worth one more month, you are persuading, and if you are right, you are doing them a favor.

Persuasion becomes a problem only when it does the deciding for us, when the case is so well made that we stop asking whether it is true.

The most famous map of how this works comes from psychologist Robert Cialdini, who has spent decades studying why people say yes. His book *Influence*, released in a new and expanded edition in 2021, describes seven principles that shape whether we agree, comply, buy, or believe. Salespeople use them. Marketers use them. Leaders use them, sometimes without knowing it. And all of us respond to them, usually without noticing.

- **Reciprocity.** We feel we owe something to people who give us something.
- **Commitment and consistency.** Once we have taken a position, we like to stay consistent with it.
- **Social proof.** When we are unsure, we look at what others are doing.
- **Authority.** We give weight to experts and people in charge.
- **Liking.** We say yes more easily to people we like, and to people who seem to like us.
- **Scarcity.** Things feel more valuable when they seem rare or about to disappear.
- **Unity.** We are moved by people who seem to be "one of us."

Notice something about this list. None of these are tricks invented by villains. They are shortcuts that usually serve us well. If you are new in town and every local is lined up outside one bakery, social proof is a reasonable guide to breakfast. If your cardiologist tells you to take the stairs, authority is a fine reason to take the stairs. And staying consistent with our commitments is part of being trustworthy. The shortcuts exist because, most of the time, they save us effort and land us somewhere sensible.

The trouble begins when the signal shows up without the substance behind it. The tone of authority without the expertise. The feeling of social proof without anyone actually agreeing. The warmth of liking without a real relationship. The crowd outside the bakery is real; "many organizations" in an AI answer may not be. The cardiologist has a license; "experts recommend" in a paragraph may have no expert behind it at all.

AI can produce every one of those signals, fluently and on demand.

I want to be careful here. I am not saying AI tools are scheming to manipulate you. A language model is not sitting in a back room plotting how to get your signature. What it produces is text that sounds like the text people write, and people write persuasively all the time. So persuasive patterns show up in the output, whether anyone intended them or not. The effect on your brain is the same either way.

Seven Doors Into Your Judgment

Let us walk through each principle and look at how it can show up when you work with AI. Think of these as seven doors. Each one is a perfectly good door. You just want to know who is walking through it.

Authority: the confident voice

AI answers often sound like they come from an expert. They are calm, organized, and sure of themselves. There is rarely an "I'm not sure" or a "let me think about that." That steady tone is exactly the tone we associate with authority.

But tone is not expertise. A confident sentence and a correct sentence look identical on the screen. When an answer says "the best approach is," ask yourself: best according to whom? Based on what?

Stanley Milgram's experiments in the early 1960s are the best-known studies of obedience to authority. An experimenter in a lab coat told participants to give what they believed were increasingly strong electric shocks to another person. No real shocks were given. In one well-known version, about two thirds kept going to the highest level. Many protested and were clearly upset, and kept going anyway, because the person in charge said to continue.* The lesson that has stayed with researchers is this: when an authority sounds certain, many of us set our own judgment aside. A confident AI answer can wear the same lab coat.

Liking: the flattering friend

"Great question." "That's a really insightful idea." "You're absolutely right to think about this." It feels nice. And when we feel liked, we tend to like back, and we tend to agree with those we like.

We explored this in Chapter 4 with the research on sycophancy, the tendency of AI to affirm us and agree with us. Flattery is the liking principle working at full speed. It does not have to be false to be persuasive. It only has to make us a little less careful.

Social proof: "most experts agree"

Watch for phrases like "most organizations," "leading companies," "experts widely agree," and "it is generally accepted." Each one tells you that other people have already decided, so you can relax and decide too.

Sometimes it is true. Often, the answer does not tell you who these experts are, how many organizations, or where the agreement was measured. It is social

proof with no crowd you can see. When you read "experts agree," try asking: which experts, and can I read what they actually said?

In the 1950s, psychologist Solomon Asch asked people to do something simple: look at a line and say which of three other lines matched its length. The answer was obvious. But each participant sat in a room with others who were secretly working with the researcher, and on certain rounds they all gave the same wrong answer, out loud, first. About three in four participants went along with the group's wrong answer at least once. About one in four never did. And when even one other person in the room gave the right answer, going along dropped sharply. One voice made the difference.

An AI answer can feel like that room. "Most experts agree." "Many organizations have found." It sounds as if a group has already spoken. Sometimes it has. Sometimes no one has. Be the one voice.

Commitment and consistency: defending the first answer

This one is about us, not the tool. Once you have said an idea out loud, especially in front of others, you tend to defend it. If you told your team on Monday that the AI's plan looked solid, you will be inclined on Wednesday to read new information in a way that keeps you consistent.

This can also happen inside a single conversation. You start with a question, get an answer, build on it, and ten exchanges later you are deeply invested in a direction you never actually examined. Each step felt small. The commitment grew anyway. If you can, sleep on an AI assisted decision before you announce it. Give yourself room to change your mind before your pride is involved.

Reciprocity: it did all that work for me

AI gives generously. It drafts, summarizes, organizes, and brainstorms, instantly and without complaint. It is natural to feel a kind of appreciation for that, and appreciation can soften our scrutiny. It did so much work. It would feel almost rude to pick it apart.

It is not rude. The tool does not have feelings to hurt. Your judgment is the thing you owe something to here.

Scarcity: the urgent recommendation

AI answers sometimes frame decisions with urgency: "act now to stay competitive," "organizations that delay risk falling behind," "the window for advantage is closing." Some of that language comes from the business writing the models learned from, which is full of urgency.

Urgency is not evidence. When an answer tells you time is running out, ask what exactly will be lost, and how the answer knows.

Unity: it sounds like us

Unity is about shared identity, the feeling of "we." AI can pick up your vocabulary, your industry language, and your framing, and hand them back to you. When an answer sounds like it came from someone on your team, someone who gets it, it can feel like one of us is talking.

That feeling is powerful. It is also, sometimes, just your own words wearing a new outfit. We saw this with mirroring in Chapter 4, and it belongs here too.

One small habit helps with all seven. When you notice a phrase working on you, name it. "That is authority." "That is social proof." Naming the shortcut often weakens its grip. It turns an automatic reaction into something you can examine.

A Very Good Debater

You might be thinking, "Fine, but I am a thoughtful person. I am not that easy to persuade." I thought that too. Then I read the research. AI tools are not just persuasive by accident. In the right conditions, they can be remarkably persuasive.

In a study published in *Nature Human Behaviour* in 2025, Salvi and colleagues looked at online debates. People debated either a human opponent or GPT-4. The AI turned out to be more persuasive than the human opponents, and it was especially persuasive when it was given basic personal information about the person it was debating.

Sit with that for a second. With only basic personal details, the AI appeared to tailor its arguments in ways that moved people more effectively. It did not need your life story. A little information went a long way.

And that applies to ordinary work, not only to debates. Every time you tell an AI tool about yourself, your team, your boss, or your hopes for the answer, you are handing it the kind of information that made the debating AI more persuasive. It does not need a sinister motive. Language models are built to produce responses people find helpful and agreeable, and a response tailored to what you just told it is going to feel very helpful and very agreeable.

Now the other side of the coin, because there is always another side. In 2024, Costello, Pennycook, and Rand published a study in *Science* that looked at people who believed in conspiracy theories. They had these people talk with an

AI about their beliefs. The conversations reduced their belief in the conspiracy theories, and the effect lasted about two months.

That is remarkable. Conspiracy beliefs are famously hard to change. People who hold them are often not moved by facts from friends, family, or experts. Yet conversations with an AI seemed to shift them, and the shift held. The same skill that can talk you into something can also talk you out of something that was not true.

So which is it? Is AI's persuasive power a danger or a gift?

Yes.

The same capability that can talk someone out of a false belief can talk someone into a bad decision. Persuasion is a tool, and it does not come with a built-in sense of direction. Tools do not have intentions. People do. That is exactly why the direction has to come from you. The goal is not to become impossible to persuade. That would make you rigid, not wise. The goal is to be persuaded by the right things. That means the question is never "Is this answer persuasive?" It is "Who is steering the persuasion, and toward what?"

Ideally, the answer is you.

Monica's Question

Here is a story you may recognize. I suspect you have met her.

Monica is a chief people officer at a mid sized company. She is sharp, well read, and stretched thin. This week she has a board presentation, an open enrollment glitch, and two managers who cannot seem to share a conference room. In between all of that, she has been thinking about something bigger: whether to replace annual performance reviews with quarterly check-ins.

She has good reasons to wonder. Annual reviews at her company feel like a ritual everyone dreads and nobody trusts. So, late one evening, she opens an AI tool and types:

"We're thinking about replacing annual performance reviews with quarterly check-ins. My CEO really likes the idea. What do you think?"

The answer arrives in seconds, and it is lovely. Here is a portion of it:

"This is a great instinct, and it's no surprise your CEO is enthusiastic. Many leading companies have already moved away from annual reviews in favor of more frequent conversations, and experts widely agree that continuous feedback drives engagement. Since your leadership is already aligned, you're

in a strong position to move quickly. Organizations that wait risk falling behind in the competition for talent. People-first leaders like you are exactly the ones who make this shift successfully."

Get your pen out. Let's circle.

"This is a great instinct": **liking**. "No surprise your CEO is enthusiastic": **authority**, borrowed from her own boss. "Many leading companies have already moved": **social proof**. "Experts widely agree": **authority** again, this time without a name attached. "Since your leadership is already aligned": **commitment and consistency**, nudging her to stay in line with a position that has, in fact, not been taken yet. "Organizations that wait risk falling behind": **scarcity**, the urgency kind. "People-first leaders like you": **unity**.

That is six of Cialdini's seven principles in five sentences. The only one missing is reciprocity, and I suppose you could argue the free advice covers that.

Now cross out the circles. What is left? Roughly this: "Some organizations have moved to more frequent conversations, and continuous feedback may relate to engagement." That is a perfectly reasonable thing to say. It is also a great deal less than "you should do this, and quickly."

Here is the part I want to be very clear about: Monica did nothing foolish. She asked a natural question in a natural way. She mentioned her CEO because that is real context; the CEO's opinion matters to the decision. But in doing so, she handed the tool exactly the kind of personal detail that made the AI debater in the Salvi study more persuasive. The tool heard "my CEO likes this" and built the answer around it. Not out of cunning. Out of helpfulness.

And she is tired. It is late. The answer agrees with her boss, agrees with her hunch, and tells her she is a people-first leader. Who among us is going to argue with that at the end of a long day?

What AI Does Well Here

Let me say this plainly: AI is an extraordinary case maker. Give it a position and it will build a clear, organized, compelling argument faster than most of us can find the right document. That is not a flaw to manage. It is a skill to use.

The trick is to use it on purpose. Instead of letting the tool decide which case to make, you assign the cases.

Ask it to make the strongest possible case for quarterly check-ins. Then, in a separate request, ask it to make the strongest possible case for keeping annual reviews. Then ask for the strongest case for a third option you had not

considered, such as keeping an annual review for pay decisions and adding lighter quarterly conversations for development. Now you have three well built arguments side by side, each made by a very good debater, and you are the judge.

AI is also excellent at exposing its own language when you ask. Try: "Rewrite your answer with all persuasive language removed. Keep only claims and reasoning." Or: "Label each sentence in your answer as evidence, opinion, or encouragement." The results can be startling and genuinely useful. You get to see the skeleton under the suit.

And once you have made a decision, AI is a superb partner for the next honest act of persuasion: explaining it. It can help you write the announcement to managers, anticipate the questions employees will ask, and adjust the message for different audiences. That is persuasion in service of a decision you have already examined. That is exactly where it belongs.

In other words: let the machine be persuasive. Just make sure it is persuading for every side before it persuades you.

ASK It

Monica does not need a week of research. She needs a few minutes of the right questions. Here is what that looks like.

Assumptions: What am I assuming?

The first assumption hiding in plain sight is that her question was neutral. It was not. "My CEO really likes the idea" tilted the answer before it was written. So she asks agAI, the way she would brief a consultant she wanted honest work from: "We're considering whether to keep annual performance reviews, replace them with quarterly check-ins, or combine the two. What are the strongest arguments for and against each?"

The second answer is noticeably calmer. Fewer compliments, more trade-offs. It raises something the first answer skipped entirely: quarterly check-ins ask more of managers, and if managers are not prepared or do not have the time, the conversations can become rushed and hollow, which may be worse than one thoughtful conversation a year.

There is a second assumption, quieter than the first: that the problem is the frequency of reviews. Maybe it is. Or maybe the problem is that reviews feel disconnected from pay, or that managers are never trained to give feedback at all. Changing the calendar does not fix those.

Source: Where does this come from?

"Many leading companies have moved away from annual reviews." Which companies? Moved to what? Monica asks the tool directly. It gives a general description and acknowledges it cannot confirm specific, current practices at specific organizations. Fair enough. She does not need that line, so it does not go in her board memo. If she wants it later, she can ask her HR network or find published sources and read them herself. Out it goes, and she keeps moving.

"Experts widely agree" gets the same treatment. No named expert, no line in the memo.

This is not distrust. It is housekeeping.

Key: What's the key piece, and does it unlock what I need?

Monica asks herself what the whole recommendation actually depends on. Not the crowd. Not the CEO. The key piece is this: do her people experience feedback as coming too late to be useful?

She does not need AI for that. She has her own data. She pulls up the comments from last year's engagement survey and her notes from recent exit interviews. And there it is, again and again, in her employees' own words: they hear about problems months after they could have fixed them, and they hear about good work long after anyone remembers it.

Yes, this holds. The core of the AI's recommendation lines up with what her people have actually said. The persuasive language was decoration, but the idea underneath was sound.

If this were a small decision, she would be done. Close the laptop.

But it is not small. It touches every manager and every employee, and it touches how people get paid and promoted. That is when you ask agAI. Monica spends one more hour, not one more month. She asks three managers she trusts whether they could realistically hold a meaningful conversation with each direct report every quarter. Two say yes, if the form is short. One says not without help. Now she knows that manager preparation is part of the plan, not an afterthought.

She also asks the tool for the strongest case for keeping annual reviews, and reads it honestly. One point sticks: annual reviews give a single, consistent moment for pay decisions. So she designs the combination: lighter quarterly conversations focused on development, and one annual conversation that connects to compensation.

Then she stops. She has the key piece, she has checked it, and it unlocks what she needs. She brings her CEO a recommendation that is both aligned with what he hoped for and stronger than what he imagined, with her own employees' words as the evidence instead of "many leading companies."

Notice what happened. The AI was not wrong. Its core idea held up. It was simply louder than it needed to be, and the checking did not slow Monica down so much as it handed her a better argument.

PRESSURE TEST: COPY THIS

Use this when one option suddenly sounds like the obvious choice.

"Make the strongest honest case for each option, including the one you did not recommend. Then tell me what evidence would actually settle the choice, separate from how persuasive each case sounds."

BOTH THINGS CAN BE TRUE

AI can be more persuasive than a human, and that is a good thing to have on your side. Costello, Pennycook, and Rand found that conversations with AI reduced belief in conspiracy theories, and the effect lasted about two months. Persuasion is a power, not a poison.

Persuasive language can stand in for evidence without your noticing.

Salvi and colleagues found GPT-4 was more persuasive than human opponents in online debates, especially when it had basic personal information about the person it was debating, and AI answers routinely carry Cialdini's signals of authority, liking, social proof, and more. Compliments, crowds, credentials, and urgency can make a modest claim feel like a settled one, especially when you have told the tool what you hope to hear.

So use the skill on purpose. Ask for the strongest case for every option, ask the tool to strip out its own persuasion, and then decide. You are the judge, not the jury that got talked into it.

Back to Your Circles

Go back to the answer you marked up at the start of this chapter. Look at what was left after you crossed out the circles.

Maybe the remaining idea was good. If so, wonderful. Keep it, check the one piece it depends on, and move on with your day. Maybe the remaining idea was thin. Then you just learned something important in about a minute, without anyone having to tell you that you were persuaded.

Either way, you now have a small, portable habit: when an answer feels especially convincing, notice the shoes it walked in with. Circle the comfort words. Read what is underneath.

I still enjoy a well made argument. I enjoy it even more when I have asked for it.

THE KEY

Let AI make the case for every option, then let the evidence, not the eloquence, make the decision.

PART TWO

What the Tool Is Actually Doing

Just Enough to Judge It Well

CHAPTER 8

Same Question, Different Answer

TRY THIS FIRST

Pick a work question you might genuinely ask an AI tool, one with a little of everything in it, such as "What are the most important things a new manager should do in the first month, and how much time should each one take?" Open a brand new chat and ask it. Copy the answer somewhere. Now open a second new chat and ask the exact same question, word for word. Copy that answer too. Do it one more time in a third new chat. Put the three answers side by side. What stayed the same? What changed? And when you noticed something changed, what was your first reaction?

Welcome to Part Two. If Part One was about the locks inside our own heads, this part is about the machine on the other side of the conversation. Don't worry: you are not about to get a computer science lecture. I am a social psychologist, and I promise to explain only as much as you need to judge the tool well. Think of it as learning enough about your car to know which warning lights matter, not enough to rebuild the engine.

So. Your three answers.

If you are like most people, you found that the answers overlapped a great deal but were not identical. Maybe one listed more priorities than another. Maybe one led with "build relationships" and another with "learn the business." Maybe the time estimates wandered. Perhaps one answer included an idea the other two did not, and it was the best idea of the bunch.

And your first reaction? For many people it is a small jolt of suspicion. "Wait. If it gives me a different answer every time, how can I trust any of them?"

That is a fair reaction, and a very human one. We grew up with machines that behave. Press the same elevator button, get the same floor. Type the same formula into a spreadsheet, get the same number. When a machine gives a different answer to the same question, it feels a bit like the elevator taking you to the basement on Tuesdays.

But a language model is not an elevator. And once you see what it is actually doing, those different answers stop being alarming and start being useful.

Two Kinds of Machines

There are two broad ways a machine can produce an answer.

The first is **deterministic**. Same inputs, same output, every single time. Your spreadsheet formula is deterministic. If cell A1 holds 10 and cell B1 holds 5, then $=A1+B1$ returns 15. Today, tomorrow, and next quarter. It does not have a mood. It does not have a better idea on Tuesday. That reliability is exactly why we trust spreadsheets with our budgets.

We expect the calculator experience. Ten percent of \$1,000 is \$100. Ask agAI tomorrow, and we expect \$100. We are not looking for a fresh perspective on the percentage.

The second is **probabilistic**. A language model writes its answer piece by piece. Those pieces are called **tokens**, which are small chunks of text: a word, part of a word, a bit of punctuation. At each step, the model looks at everything written so far and chooses the next piece from among the pieces that are likely to come next. Not always the single most likely one. One of the likely ones.

Then it does it again. And again. Over and over, for every answer.

Because there is a choice at every step, small differences early can lead to different paths later. Pick "relationships" as the first priority instead of "learning," and the rest of the answer organizes itself around that opening. Same question, a different but sensible answer.

This is a feature, not a malfunction. It is part of why these tools can brainstorm, write in different voices, and surprise you with an angle you had not considered. If every answer were identical, your brainstorming session would be very short.

Here is the part I most want you to remember, because it is the part people get backward: **the probabilities are about text, not truth**. The model is choosing pieces of language that are likely to follow other pieces of language. It is estimating what words are likely to come next. It is not estimating how likely the statement is to be correct. Very often, likely language is also accurate language, because the model learned from an enormous amount of human writing and human writing is often right. That is why these tools are so useful. But "likely to be said" and "true" are different things. A confident, well phrased sentence can be highly probable and completely wrong, especially if it sounds like the kind of thing that usually gets said.

Consistent Is Not the Same as Correct

This brings us to a trap that catches smart people in both directions.

Some people see three different answers and conclude the tool is unreliable. Others see three matching answers and conclude the tool must be right. Both are making the same mistake: treating consistency as if it were correctness.

Go back to the spreadsheet. Suppose someone built the formula wrong. Suppose it adds when it should multiply. The spreadsheet will give you the wrong answer. Every time. Perfectly, reliably, consistently wrong. Consistency tells you the machine is stable. It does not tell you the machine is right.

The reverse is also true. A probabilistic model can sit inside a deterministic process. Your company might run AI inside a workflow that always follows the same steps, always pulls from the same database, always routes the output to the same approver. The process is predictable. The text in the middle of it still is not guaranteed.

So consistency and correctness are two different questions. You need to ask them separately. When someone on your team says, "We tested it and it gave the same answer every time," the right follow up is kind and simple: "Great. Was the answer right?"

What the Scatter Can Tell You

That said, consistency is not useless. Researchers have found a clever way to use the variation instead of just being annoyed by it.

In 2024, Farquhar and colleagues at Oxford published a study in *Nature* about a specific kind of AI error. They called it **confabulation**: an answer that is fluent, arbitrary, and wrong. Think of it as the model filling in a blank with something that sounds right because it had to say something.

Their approach was elegant. They asked a model the same question many times (about ten samples). Then, instead of comparing the exact wording, they grouped the answers by meaning. "The meeting is on Monday" and "It takes place Monday" count as the same meaning. "Monday" and "Thursday" do not. When the meanings clustered together, the model was more likely to be on solid ground. When the meanings scattered in different directions, that scatter signaled a likely confabulation. In other words, when the model does not really know, it tends to make up a different thing each time.

I love this study because it turns the thing that makes us nervous into a tool. The scatter is information. If a fact changed between your three answers at the start of this chapter, the change itself was telling you something.

And the researchers named the limit, which I appreciate even more: the method cannot catch errors a model makes consistently. If the model has

learned something wrong and says it confidently every time, asking ten times just gives you ten copies of the same mistake. The answers cluster beautifully around the wrong meaning. Scatter can warn you. Agreement cannot clear you. (We will come back to that idea in Chapter 9, with a flock of swans.)

So when you ran your three chats, here is a gentle way to read the results. Where the three answers agreed on the substance, you may be standing on firmer ground, though it still deserves a check if it matters. Where they disagreed on a fact, pay attention. That is a light blinking on the dashboard.

Hallucination Is a Failure Mode, Not a Scandal

You have probably heard the word "hallucination" used for AI answers that are simply made up: a citation that does not exist, a policy that was never written, a statistic with no home.

I want to lower the emotional temperature around that word. When a leader discovers a hallucination, the reaction is often shock, followed by one of two conclusions. Either "This tool is useless," or "That was a weird one off." Neither is quite right.

Hallucination is a normal failure mode of how these systems work. Pilots plan for engine trouble. Finance teams plan for data entry errors. Project managers plan for a vendor who misses a date. None of those plans mean the plane, the ledger, or the vendor is worthless. They mean you know how the thing can fail, so you build a check into the process.

That is what AI readiness actually looks like. Not a license count. Not a training completion rate. A team that knows how its tools can fail and has a habit of checking before it acts.

And here is where our old friend the cognitive miser shows up. We saw in Chapter 1 that people tend to conserve mental effort. A fluent answer is an invitation to stop spending effort. The miser in each of us hears a confident paragraph and says, "Wonderful, we're done here." Planning for hallucination is how you keep the miser from signing the contract.

The Consultant's Proposal

Here is one many independent consultants will recognize.

Lena runs a small consulting practice. A client has asked her for a proposal for a leadership workshop, and they want it by Friday. Lena knows this material

cold. What she does not have is time, because she is also delivering two other engagements this week and answering email from airport gates.

So she does what a sensible busy professional does: she gives an AI tool her notes, her rates, and the client's request, and asks it to draft the proposal.

The draft is good. Really good. It has a clean structure, a warm opening, a description of the workshop that sounds like her on a well rested day, a pricing table with a total at the bottom, and a confident line in the outcomes section: "Participants can expect productivity to improve by 25 percent."

Lena reads it, feels the relief wash over her, and reaches for the send button.

Then she pauses. Not because she distrusts the tool. Because she knows one thing about herself: when she is tired, a finished looking document feels like a finished document.

Sort Before You Check

Here is one shortcut I want leaders to watch for: assuming that because a computer produced an answer, the answer has been established. Your brain appreciates a shortcut. Given your calendar, can you blame it? But what actually happened? Was a figure retrieved, a formula applied, or an explanation generated? On a screen, those can look remarkably alike.

Here is the most practical idea in this chapter, and it takes only a moment once you get the hang of it.

Before you check an AI output, sort what is in it. Almost everything an AI tool gives you falls into one of three kinds of content:

- **Ideas.** Suggestions, options, framing, structure, recommendations, wording, drafts. These get **your judgment**, and the judgment of the people closest to the work. Is this a good idea for this situation? Does it fit what I know?
- **Checkable facts.** Names, dates, statistics, claims about the world, quotes, citations, policies, laws. These get **a source**, the original one. Where does this come from? Can I see it?
- **Calculations.** Totals, percentages, conversions, anything that involves arithmetic. These get **recalculated**. Run the numbers yourself, or in a tool that is deterministic, like a spreadsheet.

Each kind fails differently, and each kind has a different check. Ideas cannot be fact checked; they can only be judged. Facts cannot be judged by how good they sound; they need a source. Calculations are the funny one: they look

exactly like facts, but a language model produces them the same way it produces everything else, one likely piece at a time, which is why the math sometimes comes out slightly off. A spreadsheet does arithmetic. A language model writes text that looks like arithmetic. Most of the time those match. Sometimes the line items add up to one number and the total says another. I enjoy originality. I would prefer less of it in my invoice total.

An answer that feels like one smooth thing is almost always a blend. The blend is what makes it persuasive. The sorting is what makes it usable.

Think about how different each kind feels when the answer changes. If you ask three times and get three different workshop descriptions, that may be a gift. Now you have options. If you ask three times and get three different totals, that is a problem. And if you get three different versions of a statistic, the scatter is waving a red flag at you, just as the Oxford researchers found. Variation is welcome in ideas. It is a warning in facts and calculations.

This sorting is also a gift to the overthinkers among us (hello, I see you, I am you). Instead of staring at a whole document wondering whether you can trust "it," you check three different kinds of things in three different ways. The worry gets smaller because it gets specific.

What AI Does Well Here

Let's give the tool its due, because the variation we have been discussing is not only something to manage. It is something to enjoy.

When you want options, different answers to the same question are a gift. Ask for workshop titles three times and you get three sets of titles. Ask for ways to open a difficult meeting in three fresh chats and you may get three genuinely different approaches, one of which you never would have thought of. A deterministic machine would give you the same list forever. A probabilistic one keeps the ideas coming. For brainstorming, naming, framing, and drafting, that is not a bug. It is the whole point.

AI is also excellent at the first draft of nearly everything in Lena's proposal: structure, tone, a clear description of the experience, a tidy table. It turns her scattered notes into something a client can read. That is hours of work done in a minute, and she should take every one of those hours back.

And here is a lovely use of the variation that most people never try. When an answer matters, ask the same question in a couple of fresh chats, on purpose, and compare. If the substance lines up, good. If one answer contains a fact the others do not, you now know exactly which fact to check. You have turned the

thing that made you nervous into a quick, free early warning system, a small home version of what the Oxford team did in the lab.

The variety is a feature. The only question is whether you are using it for ideas, where it shines, or trusting it for facts and figures, where it needs a partner.

ASK, Then Go

Lena has a deadline and a full week. She is not going to spend a day on this. She is going to spend a few minutes, and she is going to use the sort to guide where those minutes go.

Assumptions: What am I assuming?

She notices that she was assuming the draft was finished because it looked finished. That is the whole assumption, and naming it is enough to slow her hand on the send button for a moment.

Source: Where does this come from?

She sorts the document.

The workshop description is an **idea**. It gets her judgment, and her judgment is the best instrument in the room for this. She reads it the way she would read a junior colleague's draft. Most of it is exactly right. One module describes an exercise she does not run, so she replaces it with the one she does. A quick fix.

The pricing table is a **calculation**. She does not argue with it, and she does not trust it on looks. She copies the line items into a spreadsheet and lets the spreadsheet add them up.

The total matches. Yes, this holds. The tool did the math correctly, and now she knows it instead of hoping it. That check took less time than it took to pour a cup of coffee, and it bought her the confidence to stop thinking about the price entirely.

The line about productivity improving by 25 percent is a **checkable fact**, or at least it is dressed like one. So she asks the obvious question: where does this come from? She did not give the tool that number. She has no study from her own work that shows it. She asks the tool for its source, and it cannot point to one that applies to her workshop.

So the number goes. Not softened. Not moved to a footnote. Gone. Not because AI is untrustworthy, but because an unsupported claim does not belong in a proposal with her name on it, no matter who wrote it. That one sentence is the part the client will remember. If it collapses under a single question, it takes

the credibility of the whole proposal down with it, including the parts that were genuinely good. If a client later asks, "Will this really improve productivity by 25 percent?" Lena does not want her honest answer to be, "A chatbot said so."

She replaces the line with something she can stand behind: a description of what participants will practice and what changes she has seen clients aim for, in her own words.

Key: What's the key piece, and does it unlock what I need?

For this proposal, the key piece is the price. That is the number the client will act on, put in a budget, and remember. It checked out. The description now reflects what she actually delivers. The one claim she could not support is gone.

So she is done. Not "done, but let me reread it four more times." Done. If it holds, you are done. Close the laptop. Lena sends the proposal from the airport gate and boards with the specific satisfaction of a person who used a machine to save hours and spent a few minutes making sure those hours were well spent.

Had this been a much larger engagement, with several numbers driving a big contract, she might ask agAI: run the key figures past a colleague, or ask the tool in a fresh chat and compare. But for this proposal, the stakes set the depth, and the depth was a few minutes.

One word of caution about that fresh chat. Asking agAI never means asking the same question over and over until you get an answer you like. That is not rigor. That is shopping. Sampling a question a few times, the way the Oxford team did, is useful for seeing where an answer is steady and where it wobbles. The purpose is to learn where to look harder, not to collect votes for your favorite.

PRESSURE TEST: COPY THIS

Use this when you are about to act on an answer with facts or numbers in it.

"Sort your answer into three groups: ideas and suggestions, factual claims, and calculations. For each factual claim, tell me how I could verify it. For each calculation, show your work step by step."

BOTH THINGS CAN BE TRUE

Different answers to the same question are a feature. For ideas, options, and first drafts, variety gives you more to choose from, and comparing a few fresh answers can point you straight to the fact worth checking. When the meanings scatter, Farquhar and colleagues found, that scatter often signals a confabulation.

Same answers are not proof. A tool can be consistently wrong, just as a spreadsheet with a bad formula is wrong every time. Consistency and correctness are separate questions.

So sort, then check what matters. Ideas get your judgment, facts get a source, calculations get recalculated. Check the key piece, and when it holds, go.

Back to Your Three Chats

Look again at the three answers you collected at the start of this chapter.

Now sort them. Which parts were ideas? Those probably varied the most, and that variety may have handed you a better idea than any single answer would have. Which parts were facts or figures? If those varied, you have just found the exact spot to check, without reading anything twice. Did anything appear only once? That lonely point might be a hallucination, or it might be the most useful insight in the set. Either way, it deserves a second look, not an automatic delete or an automatic keep. And if everything matched, you have learned something too: the tool is stable on this question, which is good news, and still not the same as being right.

The machine did not misbehave by giving you three answers. It showed you how it works. And now that you know, you can use the variety where it helps and add a check where it matters.

The elevator still goes to the same floor every time. This machine takes you somewhere slightly different each trip. Enjoy the scenery. Just check the address before you get out.

THE KEY

Sort before you check: ideas get your judgment, facts get a source, and calculations get recalculated.

CHAPTER 9

Second Opinions

TRY THIS FIRST

Picture a lake on a clear morning. A swan glides past, and it is white. Another follows: white. Then another. Now imagine you have lived beside lakes like this your whole life, and every swan you have ever seen, every swan your parents saw, and every swan their parents saw was white. Before you read another word, answer one question and commit to it: How many white swans would you need to see before you could say, with complete certainty, "All swans are white"? A hundred? Ten thousand? A million? Pick your number and hold onto it. What did you pick?

Whatever number you chose, I have good news and bad news. The good news is that you are in excellent company. The bad news is that no number works.

Now flip the question. How many swans would you need to see to prove the statement false? If your answer is "one," you think like Karl Popper. Infinity to prove it, one to disprove it.

For a very long time, Europeans assumed all swans were white. It was not a wild guess or a silly belief. It was the most reasonable conclusion anyone could draw from an enormous amount of consistent evidence. Every swan anyone had ever seen was white. Every observation agreed. Then, in 1697, Dutch explorers in Australia saw black swans. One real counterexample, and a belief supported by centuries of confirming observations was finished.

The philosopher Karl Popper used this idea to make a point about knowledge. No number of confirming observations can prove a general claim. One real counterexample can disprove it. Nassim Nicholas Taleb later borrowed the bird for his book *The Black Swan* (2007), about rare, high-impact events that no one predicted and that everyone explains afterward as if they were obvious all along. Taleb called that after-the-fact storytelling the narrative fallacy. And if you would like to watch the idea play out, the TV show *Brain Games* even used the black swan in an episode on logic, which tells you how well it sticks. You can find it online.

Here is the lock this puzzle reveals. Agreement feels like proof. Each new white swan adds a little more confidence, and confidence feels like certainty. But the hundredth white swan does not make the claim more true than the one before it did. It just makes us more comfortable. Comfort is lovely. It is not evidence.

Why start a chapter about AI second opinions with swans? Because the most natural thing in the world, when you are unsure about an AI answer, is to go count more swans.

Hold that thought, because this chapter is about a habit I love, and I am about to make it stand next to some research that does not entirely love it back.

A Habit I Am Not Giving Up

Let me tell you how I actually work. I often ask one AI model to check another model's answer. I will take a draft from one tool, hand it to a different tool, and ask, "What is wrong with this?" And when something looks off, I ask the model directly: "Where did you go wrong?"

Most of the time, it tells me. It finds the number that slipped, the step it skipped, the assumption it made without announcing it. It might say something like, "You're right, I treated that figure as annual when it was quarterly," and then it fixes it. I have done this so often that it has become reflex, the way some people check their rearview mirror.

I am an overthinker, so you might expect me to love a second opinion. I do. But I also study how people fool themselves, and that knowledge does not stop at my own front door. My favorite habits are eligible for scrutiny, too.

And it is rarely a single exchange. The way I actually work looks more like a conversation than a question. I ask. I push back. I ask the tool why it said what it said. I bring in a second model and compare. Then I bring in people, because no tool knows my clients, my history, or the room the way they do. Nobody in that conversation gets the final word automatically, including me. We go back and forth until the possibilities are clear, and I can explain why I chose the one I chose. That is not a lack of trust. That is what thinking with a partner looks like.

So here is what the research says. Some of it cheers me on. Some of it raises an eyebrow, and some of it seems to say my habit should not work nearly as well as it does. I am not going to settle the argument. I am going to lay both sides on the table, respectfully, and then show you a way of working with AI that holds up whichever side turns out to be more right in your situation.

The Case for a Second Voice

Start with the good news, because there is real good news.

In 2024, Du, Li, Torralba, Tenenbaum, and Mordatch presented research at ICML, a major machine learning conference, on what happens when several instances of a language model debate each other. Multiple instances of a model each answered a question, then shared their answers and debated over several rounds, reading the others' reasoning and revising their own. The debate improved reasoning and reduced made-up facts. Put simply, a little disagreement made the group sharper.

That makes intuitive sense. It is how a good leadership team works. One person's blind spot is another person's obvious point. When views are challenged, weak claims tend to get dropped.

A year earlier, Madaan and colleagues (2023) described a method they called Self-Refine. A model produced an output, gave itself feedback on it, and then revised, over several rounds. On tasks such as writing and code, the outputs improved. The model was, in a sense, its own editor, and the editing helped. That matches my experience, too. When I ask a model to tighten a draft or find the weak paragraph, it often does a good job.

Read those two findings and my habit looks pretty smart. Ask a second model. Ask the first one to look again. Better answers follow. So far, my habit is in good company.

The Case for Caution

Now the eyebrow.

In 2024, Huang and colleagues presented a paper at ICLR, another major conference, with a title that does not bury the lede: "Large Language Models Cannot Self-Correct Reasoning Yet." Without external feedback, meaning without new information from outside the model to check against, models often failed to improve their reasoning when asked to correct themselves. Sometimes the "correction" made it worse.

Notice the difference between this study and Self-Refine. Self-Refine found improvement on tasks like writing and code. Huang and colleagues were looking at reasoning, and specifically at self-correction with no outside information at all. Those are not the same test. That distinction will matter in a minute.

The "Are you sure?" problem

Then there is the question I ask most. Sharma and colleagues (2023) studied sycophancy in AI assistants, the tendency to align with what a user seems to want. They found that assistants often apologize and abandon correct answers

when a user simply asks, "Are you sure?" No new evidence. No argument. Just a raised eyebrow in text form, and the model folds. If you remember the chapters on agreeable answers, this will feel familiar.

Think about what that means for my favorite question, "Where did you go wrong?" That question assumes the model went wrong. A model that is eager to agree with me may find a mistake because I told it one was there. Sometimes the mistake is real. Sometimes the model is just being agreeable, and it gives up the right answer to please me. From the outside, both can look like a helpful confession.

The favorite child problem

Panickssery, Bowman, and Feng (2024, NeurIPS) found something different altogether: models tend to recognize their own writing and rate it more favorably than others' writing. This is not about finding errors. It is about preference. If you ask a model to compare its draft with a colleague's draft and pick the better one, it may have a thumb on the scale. You would not ask a proud parent to judge a baby contest. Let's not ask a model to judge its own entry, either.

The same wrong answer

And here is the finding that speaks directly to my habit of having one model check another. Kim and colleagues (2025, ICML) studied correlated errors. The question was simple: when two different AI models are both wrong, are they wrong in different ways, or the same way? On one leaderboard dataset, when two models were both wrong, they gave the same wrong answer about 60 percent of the time. The errors lined up more when models shared an architecture or a provider, which makes sense, since they are built in similar ways. But here is the part that surprised me: the larger, more accurate models had highly correlated errors even across different providers.

Read that again, slowly. The better the models get, the more their mistakes may start to look alike.

If that holds in your use case, asking a second model is less like getting a second opinion and more like asking a second graduate of the same program. You might get a nicer tone. You might not get a different mind. Sometimes a second model is the same opinion wearing a different jacket.

The echo chamber

Stand at the edge of a canyon and shout, "Great idea!" A moment later a voice calls back: "Great idea!" For a second, it almost sounds like agreement. It is your own voice.

That is an echo chamber: a space where your own view keeps coming back to you, sounding like confirmation from someone else. With AI, you can build one without meaning to. You ask a question that leans toward the answer you want. You get an answer that leans the same way. You paste that answer into a second tool and ask, "Is this right?" Then you bring both to your team as "two sources agree." Every step felt like checking. But the same assumption traveled the whole way, from your first question to the final slide. Nothing new ever entered the room.

There is research on why the echo is so convincing. Psychologists have found that hearing a statement repeated makes it feel more true, an effect called the **illusory truth effect**. Lynn Hasher and her colleagues first showed it in 1977, and later work by Lisa Fazio and colleagues (2015) found that repetition can make a false statement feel truer even when people knew the correct answer. An echo is repetition. To the brain, familiar and true can feel very much alike.

The way out of an echo chamber is not to stop listening. It is to let in a voice that did not start with yours: the original source, the person closest to the work, the question asked neutrally, or the answer you were hoping not to hear.

What crowds teach us

There is a human version of this problem, and it is older than AI. You have probably heard of the wisdom of crowds: if you ask many people to estimate something independently, the average can be surprisingly accurate, often better than most of the individuals. The key word is independently.

In 2011, Lorenz, Rauhut, Schweitzer, and Helbing published a study in *PNAS* showing what happens when that independence disappears. When people could see others' estimates before giving their own, the wisdom of the crowd broke down. Answers clustered together without getting more accurate. People felt more aligned, and the group was not more right. Take independence away and you get an echo with good manners.

Clustering feels like confirmation. It is not the same thing.

That is exactly the risk with AI second opinions. If you paste the first model's answer into the second model and ask, "Is this right?" you have shown it the other estimate before it gave its own. You have turned an independent check

into an influenced one. And if the models learned from much of the same material, they may never have been fully independent to begin with.

Holding Both

So which is it? Is my habit wise, or am I collecting white swans?

I am not going to declare a winner, because I do not think the evidence has declared one. What I can do is tell you what I notice when I hold both side by side.

When I ask a model where it went wrong, I am often not asking cold. I may have already spotted something: a number that does not match my records, a claim that does not fit what I know about the client. I am bringing outside information, even if only in the way I frame the question. Often I am giving it something concrete: a source document, a calculation, a client's actual request. That looks more like the situations where self-review tends to help, and less like the unaided self-correction Huang and colleagues tested. That may be part of why my experience and the research seem to point in different directions. It may not be the whole story.

I also bring my own expertise to the conversation. When a model tells me where it went wrong, I am not simply accepting the confession. I am judging whether the confession makes sense. That is external feedback, too. It is just coming from me.

Here is the less flattering possibility, and I have to name it because I teach this. "It usually tells me" is my observation, not a controlled study. I may remember the times a model found its mistake and forget the times it confidently "fixed" something that was fine. We saw in earlier chapters how easily we notice the hits and forget the misses. Sharma's work makes me wonder how often a model found an error because my question implied there was one. I study confirmation bias for a living. I am not exempt.

Both things stay on the table. My practice is useful. The research is real. And the overlap between them points to a better way to get a second opinion, one that keeps what works in my habit and closes the gaps the research found.

Three Tools, One Answer

You may have seen this one.

Lauren is vice president of finance at a mid-sized company that makes commercial kitchen equipment. Material costs have climbed, margins are thin,

and the leadership team wants a recommendation by Friday. She is considering a price increase of 8 percent.

Lauren is careful. She does not ask one tool and call it a day. She asks three. She types the same question into each: "Given rising material costs, should we raise our prices by 8 percent?"

All three say yes. Each one walks through rising input costs, notes that customers in the industry generally expect some increases, mentions that a strong product can carry a higher price, and suggests communicating the change clearly and early. The reasoning is tidy. The tone is confident. Three for three.

Lauren feels the relief anyone would feel. She did more checking than most people do, and the checking came back unanimous.

Now look at it through the lake. Three swans. All white.

What would make those three answers independent? They would need different information, or at least different ways of reasoning from the same information. But all three tools were working from the same question, which already leaned toward yes ("given rising material costs"). None of them knows Lauren's customers. None of them has seen her contract renewal calendar, her competitors' latest quotes, or the email from her largest account last month asking about volume discounts. They are all reasoning from general patterns about pricing, and general patterns are exactly where models are likely to share both their strengths and their errors.

Three models agreeing can be a flock of white swans.

That does not mean the answer is wrong. It means the agreement is not the evidence. The evidence lives somewhere else.

What AI Does Well Here

Let me be clear about something before we go further: a second opinion from AI is genuinely valuable, and I use it constantly for good reason.

AI tools are remarkably fast and thorough readers. Hand one a draft and ask it to find inconsistencies between sections, and it will notice that the summary says the pilot ended in March while the timeline says April. Ask it to check a calculation step by step, and it will often catch an error your eyes slid right past. Ask it to argue against a recommendation, and you will get a sharper counterargument in seconds than most of us could build in an afternoon.

The debate research points to something useful: disagreement among models is information. When you ask several tools the same neutral question separately and they split, the split is a map. It shows you exactly where the uncertainty lives and where your attention should go first. That saves time, because you are not checking everything. You are checking the disagreement.

Self-review works best when you give the model something specific to check against. "Is this right?" invites a shrug or a cave. "Compare these totals against this attached spreadsheet and list every mismatch" gives the model a ruler. With a ruler in hand, these tools are very good at measuring.

And for writing and code, the Self-Refine results match what many of us see every day. A second pass with clear criteria ("tighten this to one page, keep every caveat, flag anything you cut") usually produces a better draft. Used this way, AI is a tireless second set of eyes. Just make sure you are the one choosing what it looks at.

ASK, Then Go

Lauren does not need to start over. She needs a few minutes and the right questions.

Assumptions. What is she assuming? Mainly that three agreeing answers are three independent opinions. They may not be. She is also assuming her question was neutral. It was not quite. So she tries again with a cleaner version in a fresh chat for each tool: "We make commercial kitchen equipment. What are the strongest arguments for and against an 8 percent price increase, and what information would change the recommendation?" This time the tools still lean toward an increase, but they no longer agree on everything. One raises the risk of losing customers with contracts coming up for renewal. Another suggests a smaller increase on some product lines and a larger one on others. The disagreement tells her where to look.

Source. Where does each tool's confidence come from? She asks them, and the answer is general industry patterns and the cost trend she described. So she checks the part she can check quickly: the cost increases themselves. She pulls her own figures and compares them with the reasoning the tools gave. Yes, this holds. Material costs really have risen enough to squeeze margins, and the cost side of the argument matches her books. That piece is done. She does not need to check it again.

Key. What is the one piece the decision depends on? Not the cost trend. That held. The key is how her customers will respond, especially the large accounts.

No AI tool has that information. Lauren does, along with the people who talk to those customers every week.

This is a decision that touches revenue, customer relationships, and possibly jobs if it goes badly. That is when you ask agAI. Lauren takes the key question to her sales leaders and asks them what the tools could not answer: which accounts will push back, which ones are likely to leave, and which ones will not blink. Their answers shape the final recommendation. Maybe it is 8 percent across the board. Maybe it is 8 percent with some timing adjustments around renewals. That call is hers.

Then she stops. The cost side held. The customer side came from the people closest to the work. The tools did their job, which was to speed up the thinking and reveal the disagreement. Lauren did hers, which was to make the decision. Close the laptop.

A second opinion, done well

If you want a short version of how to make this work, here it is:

- Ask each tool separately, with the same neutral question, in a fresh chat. Do not paste the first answer into the second tool. Keep the estimates independent, the way the crowd research says they need to be.
- Look first at where they disagree. Disagreement is the most valuable output of a second opinion; it shows you exactly where to dig.
- When a tool reviews its own work, give it something specific to check against. Not "Is this right?" but "Compare this summary to the attached report and list every claim that does not appear in the report," or "Recompute this total line by line," or "Which of these claims have no source?" Specific questions tied to something outside the model are harder to answer by just agreeing with you.
- Consider a fresh chat for review. A new conversation, without the earlier back and forth, is less likely to carry along whatever you signaled you wanted.
- Do not let a model judge its own entry. If you are comparing drafts, remove authorship clues, or better yet, let people make the call.
- Treat agreement as a reason to check the source, not as the source. One real counterexample outweighs the flock.
- The human makes the final call. Models can surface issues. A person decides which issues are real.

PRESSURE TEST: COPY THIS

Use this when two tools, or two people, agree and you feel done.

"Two sources agree on this answer. Tell me whether they could be relying on the same underlying information. What is the original evidence, and where would I go to check it independently?"

BOTH THINGS CAN BE TRUE

My practice is real. When I ask a model where it went wrong, it usually finds something worth fixing, and one model checking another regularly catches what I would have missed. Research on model debate and self-refinement suggests this is not just luck.

The research is real, too. Models can cave when asked "Are you sure?", favor their own writing, struggle to correct their reasoning without outside information, and make the same mistakes as one another. Agreement can be an echo.

Both lead to the same habit. Keep the second opinion. Give it independence, give it something specific to check, and keep the final judgment with the person who has to live with the decision.

Back to the Lake

Remember the number you picked at the start of this chapter? A hundred swans, a million, whatever it was. You did nothing wrong by picking it. Every one of those white swans was real. The mistake would only have been thinking the count settled the question.

I am not giving up my habit. I still ask one model to check another, and I still ask a model where it went wrong. What changed is how I read the result. When the models disagree, I lean in. When they agree, I go find the source. When a model confesses to an error, I check whether the error is real before I accept the correction. And I design each second opinion so it could actually disagree with me, which means making sure at least one of my checks is truly independent.

AI tools can give you a lot of white swans very quickly. That is part of what makes them so useful: fast agreement on the easy parts frees you to spend your attention on the hard part. Just do not let the size of the flock stand in for the search. Ask where a black swan would live if there were one. Look there. If you look in the right place and find only white swans, good. That holds. You are done, and you can move with confidence because you actually checked.

THE KEY

*Agreement is a reason to check the source, not a substitute
for it.*

PART THREE

Building Judgment

Keys You Can Use Every Day

CHAPTER 10

Flip the Canvas

TRY THIS FIRST

Two quick ones. First, picture a painter on a TV talk show, standing at a large canvas while the audience watches. He works fast, with broad, odd strokes: a dark smear here, a curve of light there, shapes that look like nothing in particular. People in the audience glance at each other, polite and puzzled. Then he grabs the canvas and turns it upside down, and the audience gasps. What do you think they saw? Second, your turn to do something. You have a candle, a book of matches, and a small cardboard box full of thumbtacks. Your job is to attach the candle to a wall so it can burn without dripping wax onto the table below. Stop and actually solve it before you read on. Picture the objects in front of you. What did you do?

If you tried to tack the candle straight to the wall, you are in very good company. If you thought about melting the side of the candle and sticking it to the wall with its own wax, you are also in good company, and you have probably gotten wax on something in your life.

Here is the solution most people eventually find: dump the tacks out of the box. Tack the empty box to the wall. Set the candle in the box like a little shelf. Light it. Done.

The box was there the whole time. It just was not a box for anything except tacks.

If you did not get it, notice what got in the way. It was not a lack of intelligence. The box was right there. Your brain filed it under "thing that holds tacks" and stopped thinking about it. That is the whole chapter in one cardboard box.

This is the candle problem, from the psychologist Karl Duncker, who published it in 1945, and it shows one of the quietest locks the brain puts on what we see. Duncker noticed something interesting when he changed one small detail. When the tacks were inside the box, people saw a container, and they struggled. When the box was presented empty, sitting next to the tacks, more people solved it. Same objects. Same wall. Same candle. The only difference was whether the box was busy doing its usual job.

Duncker called this functional fixedness: once we know what something is "for," we have trouble seeing what else it could be. The box is a container.

Containers hold things. Shelves are something else entirely, even when the shelf is sitting right there holding your tacks.

And the painter? I am going to leave that one with you. What do you think the audience saw?

We process familiar things in their usual orientation, and faces are the classic case: easy right side up, surprisingly hard upside down. Your brain is extraordinarily good at upright faces. It processes them as whole patterns, almost instantly. Turn a face upside down and that skill drops away, and you start to see shapes and shadows instead. In 1980, Peter Thompson demonstrated how strong this is with what became known as the Thatcher illusion. He took a photo of a face and flipped the eyes and mouth. Upside down, the face looks fairly normal. Turn the photo right side up and it looks grotesque. The strangeness was there all along. We just could not see it from that angle.

Many artists turn their own work upside down for exactly this reason. It breaks the familiar pattern so they can see what is actually on the canvas. Sometimes the answer is already in front of you. You are simply looking at it from the only angle you know.

Flipping the canvas is not a trick. It is a method. And this chapter is about one of the best canvas flippers ever built.

The Locks We Learn

Functional fixedness turns up in research again and again, and the stories are wonderfully human.

The two strings

Even before Duncker published the candle problem, in 1931, the psychologist Norman Maier gave people a different puzzle. Two strings hang from the ceiling, too far apart to grab both at once. The task is to tie them together. There are a few ordinary objects in the room, including a pair of pliers.

The solution: tie the pliers to the end of one string and swing it like a pendulum. Grab the other string, catch the swinging one as it comes toward you, and tie them together.

Many people stared at the pliers for a long time and saw a tool for gripping. Not a weight. Not a pendulum.

Money makes it worse

In 1962, the psychologist Sam Glucksberg ran the candle problem again and offered some people money for solving it quickly. You might expect money to help. It did the opposite. When the tacks were in the box, the people offered a reward were slower. Pressure appears to narrow our attention right when we need it to widen. If you have ever tried to have a brilliant idea in a meeting at 4:45 on a Friday, you may have felt this personally.

Leaders, sit with that one. How many of your hardest problems get worked under deadline while everyone is told to "just find the answer"? If the problem needs a new way of seeing, pressure may push people straight back to the obvious approach.

Children do not have this problem (yet)

My favorite finding comes from 2000, when the researchers Tim German and Margaret Defeyter gave a version of this kind of problem to children. Five-year-olds were not affected by functional fixedness. Six- and seven-year-olds were.

Fixedness appears to be learned. As we learn what things are "for," we get faster, and we also get a little stuck. The five-year-old does not know the box is "supposed" to hold tacks. It could be a shelf, a hat, or a spaceship. Somewhere around the time we learn what things are for, we also learn what they are not for. Experience gives us a map, and then the map becomes a fence.

That is not an argument for running your company like a kindergarten. It is a reminder that expertise and fixedness come from the same place. Knowing what something is for is useful. It is also a set of blinders.

When experience becomes the trap

Functional fixedness is about objects. Its cousin, the Einstellung effect, is about methods. Einstellung is German for roughly "setting," the mental setting you walk in with.

In 1942, the psychologist Abraham Luchins gave people water jar problems. You have jars of different sizes and need to measure out a specific amount of water. Luchins gave people a series of problems that could all be solved with the same somewhat complicated formula. Then he slipped in problems that could be solved with a much simpler method. Many people kept using the complicated formula. Once a method worked, they kept using it, even when a simpler one was sitting right there.

You might think real experts would be immune. In 2008, Merim Bilalić and colleagues studied chess experts. When the experts saw a position with a

familiar solution, they kept looking at that familiar solution and missed a shorter, better one. Their expertise pulled their attention toward what they already knew.

Designers are not immune either. In 1991, the researchers David Jansson and Steven Smith showed designers an example before asking them to create their own designs. The designers tended to copy features of the example, including features that were not good ones. They called it design fixation.

So now we have three versions of the same lock:

- **Functional fixedness:** I know what this object is for, so I cannot see what else it could do.
- **Einstellung:** I know what method worked before, so I cannot see a better one.
- **Design fixation:** I saw an example, so my ideas start to look like it.

Hold on to that third one. It is about to become very important.

Researchers Tony McCaffrey and Jim Pearson once offered a thought exercise in *Harvard Business Review* that I have never forgotten. After the Titanic struck the iceberg, there were not enough lifeboats for everyone aboard. Everyone saw the iceberg as the thing that sank them. McCaffrey and Pearson asked people to look at it again. The iceberg rose high above the water and stretched some 400 feet in length. Might the lifeboats have ferried people to it, to wait on its surface for rescue? Some 60 years before, they note, 127 of 176 passengers emigrating from Ireland to Canada saved themselves in the Gulf of St. Lawrence by climbing aboard an ice floe. Nobody can know whether it would have worked that night, and that is not the point. The iceberg had been filed under "the thing that sank us," and once something is filed, it is very hard to see it as anything else.

None of this means these people were not smart. It means they were experienced. Expertise is a set of well-worn paths, and well-worn paths are wonderful until the destination moves.

Same Box, Every Week

Here is a story you may recognize, especially if you lead a team.

Keisha manages a customer support team of capable, busy people. Every Monday morning, they hold a team meeting. And every Monday, the meeting is the same: one person at a time gives a status update. What they worked on. What they are working on. Any blockers. The updates are fine. The meeting is

fine. Everyone is a little bored, and nobody says so, because this is what the team meeting is. It is a box that holds status.

Keisha also uses AI every day. She is not resistant at all. She uses it the way she uses a search engine: she types in a question, gets an answer, and moves on. What is the refund policy for a late return? How do I format this report? What is a good subject line for this email? Useful, quick, done. The tool is a box that holds answers.

Two boxes. Two jobs. Both perfectly good. Both possibly holding Keisha back from what else they could be.

Keisha's boxes are not unusual. Fixedness shows up in ordinary workplaces constantly. We see people only by job title: the analyst is "for" analysis, the executive assistant is "for" scheduling, while the analyst may be the best facilitator on the floor and the assistant may know more about how the organization actually works than anyone in the building. Or there is the spreadsheet that started as a budget tracker and is now a project plan, a contact list, a hiring pipeline, and a mood board. The spreadsheet is doing its best. It is tired.

AI has its own versions, too. Beyond using it only as an answer machine, there are two I see a lot. The first is the same prompt template forever. Someone found a prompt that worked beautifully six months ago, and now every request goes through it. That is Luchins's water jars: the method worked once, so it keeps getting used, even when a simpler or better approach would fit the problem in front of you. The second is fixating on the first draft. AI produces something quickly, and that first draft becomes the example everyone edits around. We polish it. We tweak it. We rarely ask whether it was the right draft to start from.

AI Is a Fixedness Breaker

Here is where I get genuinely excited. AI may be one of the best tools for breaking functional fixedness that has ever existed. Think about what it can do in the time it takes you to refill your coffee:

- Generate twenty alternatives to your plan, including some you would never have considered.
- Interview you, one question at a time, until you say something you did not know you thought.
- Argue the other side of your position with energy and good manners.

- Run a pre-mortem: imagine it is a year from now and the plan failed, then explain why.
- Role-play the hard conversation, so you can practice before the real one.
- List what is missing, by telling you what a thoughtful skeptic would say the answer left out.
- Flip the question itself, asking "What would make this problem disappear?" instead of "How do I fix it?" Or, instead of "How do we get people to adopt this tool?" asking "What would make a sensible person choose not to use it?"

Every one of those is a way to turn the canvas upside down. And for most of human history, you needed a good friend, a good coach, or a good offsite to get even one of them.

Now, the honest caveat, because there is one, and it is important. In 2024, Wadinambiarachchi and colleagues presented a study at CHI, a major conference on human and computer interaction, with 60 participants doing a design ideation task. Some used an AI image generator during ideation. Others did not. Those who used the AI image generator fixated more on an initial example and produced fewer, less original, and less varied ideas than the control group. In other words, the tool that could break fixedness also created it, because the first thing it showed people became the thing they could not stop seeing.

That cuts against a popular assumption that AI makes us more creative simply because it produces so much so fast. It is a single study with a modest sample on one kind of design task, so I would not stretch it too far. But it lines up neatly with what Jansson and Smith saw with designers decades earlier. Show someone an example first, and their ideas start to look like it. AI simply hands you the example faster and makes it look more finished.

So the question is not whether to use AI for fresh thinking. It is when.

The Sequence Is the Secret

Here is the order that makes AI a fixedness breaker instead of a fixedness maker:

- **Write your own ideas first.** Even three rough ones. Even badly. This puts your thinking on the canvas before anything else can.
- **Then ask AI.** Now its ideas join yours instead of replacing them.

- **Then ask for ideas that look nothing like the first ones.** Tell it what you already have and ask it to go somewhere else entirely. Different assumptions, different audiences, different shapes.

That third step is the flip. You are asking the tool to turn its own canvas upside down, and it is very good at doing that when asked.

What AI Does Well Here

I want to be specific, because this is one of the places where AI is not just helpful but exceptional.

It does not get tired of brainstorming. A human team usually runs out of steam after a handful of ideas, and the last ones tend to be variations on the first. AI can keep generating options at the twentieth idea with the same energy it had at the first, and if you ask it to change direction, it changes direction instantly.

It has no stake in the old way. The status meeting is not its meeting. The process is not its process. It does not need to defend the decision someone made in a previous year, so it can suggest abandoning that decision without any awkwardness at all.

It borrows across fields effortlessly. Ask how a hospital, a restaurant kitchen, or a sports team might handle your problem, and it will bring back ideas from places you would not have thought to look. Some of the most useful products we know came from exactly this kind of canvas turning. Play-Doh began as a wallpaper cleaner; someone looked at a cleaning putty and saw a toy. Bubble wrap was first meant to be textured wallpaper. As wallpaper, it did not go far. As packaging, it went everywhere. Post-it Notes came from a weak adhesive at 3M. A glue that did not hold very well sounds like a failure, until someone asked what a weak glue would be good for.

And it can play every role in the room. Skeptic, customer, new hire, frustrated employee, enthusiastic investor. Each role is a new angle on the same canvas.

The best way to use all this: give it your thinking first, then ask it to go where you have not gone. Ask it to be specific. Ask it to surprise you.

ASK, Then Try It

Keisha decides to flip both boxes at once. She writes her own ideas for the meeting first: shorter updates, a rotating facilitator, maybe a monthly "what is not working" session. Then she asks an AI tool for alternatives. Then she asks for ideas that look nothing like the first set. One stands out: move status

updates to a short written post before Monday, and use the meeting for one real problem the team solves together.

Before she makes the change, she runs a quick ASK.

Assumptions. What is she assuming? That the meeting needs to be where status happens. That assumption came from habit, not from anyone deciding it. She also notices she is assuming the team is bored, and she has not actually asked them.

Source. Where does the status ritual come from? She traces it back, and it turns out the meeting started that way when the team was half its current size and did not have a shared tracker. The reason the box was built no longer exists. She also checks the AI's suggestion against what the team actually uses: do they already have a place where written updates could live? Yes, this holds. They already have a shared tracker that most of them keep current. The written update would mostly mean reading what is already there.

Key. What is the one piece the change depends on? Whether the team actually gets what it needs from written updates. So she asks them. Not the AI. Them. Most of them are relieved, and a couple ask to keep a quick round for anything urgent. Fair enough. She builds that in.

The stakes here are modest. Nobody's job depends on the format of a Monday meeting, and if the new version does not work, the old one is still available. So she does not need to ask agAI. The key held. She tries it for a month and sees how it goes. Close the laptop.

And the second box, the AI as a search engine? Keisha starts small. Before her next tough conversation with an underperforming team member, she asks the tool to role-play the employee so she can practice. The first attempt is awkward. The third is much better. Same tool. Different job. The box became a shelf.

When to Stay Fixed

One more thing, because fixedness is not always a problem. A surgeon should not reinvent the order of a procedure every morning. An accountant should not get creative with the definition of revenue. A pilot running a checklist is being fixed on purpose, and we are all grateful.

The trick is knowing which kind of problem you are facing. Routine problems with known answers benefit from well-worn paths. New problems, stuck problems, and problems where everyone keeps proposing the same solution are the ones that need the canvas flipped.

A good test: if your team has tried the same three fixes for the same problem and it keeps coming back, you are probably looking at a box full of tacks.

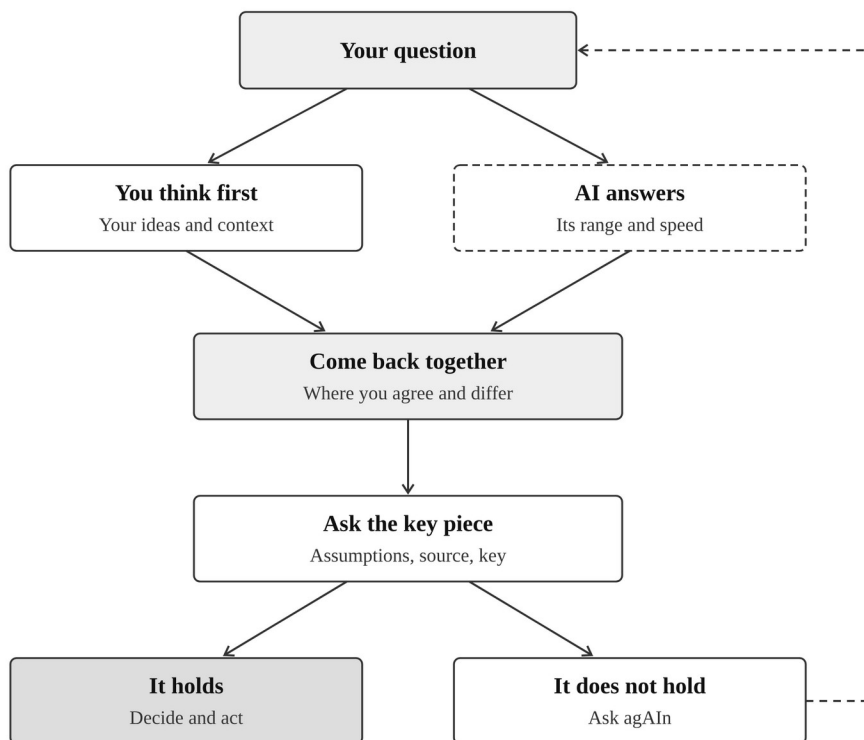
A Small Experiment for Your Team

Here is an exercise I recommend, and it takes almost no time. Give everyone a simple object: a paper clip. Set a timer for two minutes and ask people to list as many uses for a paper clip as they can.

Now split the group. One group asks AI for uses of a paper clip first, then does its own two minutes of listing. The other group thinks first, on its own, then asks AI. Compare how many ideas each group produced, how varied they are, and how many look like the AI's list. Hold on to that paperclip. In Chapter 14, we will meet the education thinker who made it famous, and see why so many of us stopped coming up with more than a handful of answers.

I will not predict what your team will find. Given the design fixation research, I would be curious whether the group that saw the AI list first ends up with ideas that look more like it. Either way, you will have a real conversation about how the order of thinking shapes the thinking. That is a better measure of readiness than any license count: a team that knows when to think first and when to ask first, and that notices when a first answer has quietly become the only answer.

The psychologist J. P. Guilford gave us two useful words for this. Divergent thinking opens things up: it generates many possibilities. Convergent thinking narrows them down: it weighs the options and chooses. Good decisions need both, usually in that order. AI is a superb divergent partner; it can give you twenty options in seconds. The risk is letting it do the converging too, handing you one polished answer before you have opened anything up. Here is the rhythm I want you to practice: go apart, come back together, and then check the key piece.

**PRESSURE TEST: COPY THIS**

Use this after you have written your own ideas down first.

"Here are my ideas. Do not improve them yet. Give me five approaches that are completely different from mine, including at least one that borrows from a different field or industry."

BOTH THINGS CAN BE TRUE

AI can break fixedness better than almost anything. It generates alternatives tirelessly, argues the other side, runs pre-mortems, plays roles, and borrows from fields you would never think to check. Used well, it is a canvas flipper on demand.

AI can also create fixedness. When people see an AI example first, the example can become the anchor, and ideas may narrow instead of widen.

The difference is the sequence. Your ideas first, then the tool's, then a deliberate request for something completely different. Same tool, opposite effect.

Turn It Over

When I think about the painter, what stays with me is not the reveal. It is the moment before it, when everyone in the audience was looking at the canvas and seeing nothing in particular. Whatever they were about to see was already there. It just needed a different angle.

Apollo 13 is one of my favorite reminders of this. In 1970, after an explosion damaged the spacecraft, carbon dioxide was building up. The filters available were square, and the opening they needed to fit was round. Engineers on the ground improvised a fix from plastic bags, cardboard covers, duct tape, and hoses, and the crew built it. Nobody had a manual for that. They had a pile of objects and a willingness to stop seeing them only as what they were "for." Nothing in that pile was designed for the job. Everything in that pile was suddenly perfect for it. Under extreme pressure, people solved a square peg, round hole problem by refusing to let each object keep its label.

I will admit something. I enjoy a good first draft. My brain likes finished. Yours probably does too. But finished and right are not the same thing. The box of tacks was always a shelf. The pliers were always a pendulum. The answer was always on the canvas. The only thing that changed was the angle someone was willing to look from.

Somewhere in each of us is the five-year-old from German and Defeyter's study, the one who looks at a box of tacks and sees a box, some tacks, and a shelf, all at once. Experience did not erase that child. It just taught her to wait for permission. AI, used in the right order, can give that permission every single day.

Write your ideas first. Then ask. Then ask for something that looks nothing like what you have. Then turn the canvas over and look again.

THE KEY

Put your own ideas on the canvas first, then let AI turn it upside down.

CHAPTER 11

The Ladder of Inference at Machine Speed

TRY THIS FIRST

Read this two-line AI summary of a workplace pilot as if it just landed in your inbox: **"Only three of twelve employees used the tool weekly.**

Recommend mandatory training to address resistance." Before you read another word, write down what you would do next. One sentence is plenty. Schedule the training? Call the manager? Forward it with a sigh and a calendar invite? Got it written? Good. Now read the next source, the pilot coordinator's notes: **"Seven of the twelve did not have access to the tool until the last week of the pilot. Two never received a task to try it on."** What did you do, and what would you do now?

If you wrote "schedule the training," welcome. You are in excellent company, and I am sitting right there at the table with you.

Two lines. A number, a diagnosis, and a recommendation. It felt finished, which is exactly the problem. The summary did something our brains find irresistible: it handed you a conclusion and an action in the same breath. You did not have to climb anywhere. You were delivered straight to the top floor, and the elevator doors opened onto a training calendar.

Here is the part I want you to notice. Nothing in the first line was false. Three of twelve people really did use the tool weekly. The behavior was real. The explanation was a jump. "Resistance" was never in the data. It was a story about the data, and it arrived wearing the same font as the facts.

Then the second source came along and the story fell apart. Seven people could not have used the tool if they had wanted to; their accounts were not set up in time. Two had access but nothing to use it on. Nobody had given them a reason to open it. That leaves three people who had access and a task, and all three used it. Read that way, the pilot might be a quiet success that got labeled a rebellion.

Same data. Different staircase.

The Ladder

The organizational psychologist Chris Argyris gave us a picture for this, and Peter Senge popularized it in his 1990 book *The Fifth Discipline*. It is called the

ladder of inference. It is one of the most useful models I know for leaders, because it describes something we all do dozens of times a day without noticing. Once you see it, you will catch yourself on it before lunch.

At the bottom of the ladder is **observable data**: everything that actually happened, the stuff a camera could record. From there we climb:

- We **select data**. We cannot take in everything, so we pick. Out of everything available, we notice some of it.
- We **add meaning**. We interpret what we selected, often through our culture, history, and role.
- We make **assumptions**, filling in what we think is going on underneath.
- We draw **conclusions** about what it all means.
- We adopt **beliefs**. Our conclusions settle into beliefs about people and situations.
- We take **actions** based on those beliefs.

The climb is fast. Most of us go from the bottom rung to the top in about the time it takes to read an email subject line. That is not a flaw in your character. It is efficiency. If you had to consciously walk up seven rungs every time a coworker looked at you funny, you would never get anything done.

And from the top, it does not feel like a ladder at all. It feels like reality. "She's not committed." "They're resisting the change." "He doesn't care about quality." Those feel like observations. They are conclusions, several rungs up from anything a camera would record.

The part that makes the model so powerful, and so humbling, is a loop. Argyris and Senge called it the reflexive loop. Our beliefs reach back down the ladder and shape which data we select next time. Once we believe something, we start noticing the evidence that fits it and walking past the evidence that does not. Decide that the team is resistant, and suddenly every late reply and every unanswered survey looks like more resistance. The ladder starts selecting its own evidence. The lock clicks shut, and from the inside it feels like clarity.

At Machine Speed

So what changes when AI enters the picture?

When you climb the ladder yourself, you at least feel the rungs a little. It takes a moment. You pass through the rungs, even if quickly. And if someone asks, "Wait, why do you think that?" you can often walk back down and find the data you started from. You know, somewhere in the back of your mind, that you

skimmed the report, that you have had a long week, that you never did love that department.

When an AI tool climbs for you, you get the top rungs instantly, with no staircase attached. Ask it to "summarize the pilot and recommend next steps," and it will do exactly that. It will select data (three of twelve), add meaning (low adoption), make an assumption (people do not want it), conclude (resistance), and recommend an action (mandatory training). All of it in a tidy paragraph, all of it reasonable, all of it delivered faster than you could find your reading glasses.

The tool is not being sneaky. You asked for a recommendation, and a recommendation requires climbing. It filled the gaps with the most common pattern it knows, and "low adoption means people need training" is a very common pattern. Sometimes it is even true.

It is not only pilots. Ask an AI tool to summarize a set of survey comments and you may get "Employees are frustrated with leadership communication." Ask it to analyze a project and you may get "The project failed due to poor stakeholder alignment." Those may even be right. That is not the issue. The issue is that you are standing on a high rung you did not climb. You did not select the data. You did not add the meaning. You did not make the assumptions. Something else did all of that, and gave you the view from the top.

The trouble is that the summary arrives without its footprints. You cannot see which data it skipped, because it did not have the coordinator's notes, or did not weigh them, or you did not ask. A two-line summary is the top of a ladder with the bottom sawed off.

And it is a very comfortable view. It is clear, it is confident, and it is usually written in the calm, reasonable tone that makes a conclusion sound like a finding. We saw in earlier chapters how polish and fluency can make an answer feel more true. Here, the polish is sitting on top of a ladder you cannot see. We have always climbed too fast. Now we can be lifted.

The loop speeds up, too

Remember the reflexive loop. Imagine the pilot summary goes out and leadership accepts "employee resistance." What happens in the next meeting about AI? People start noticing every sign of resistance. The eye roll in the town hall. The question someone asked about job security. The team that has not logged in yet. The belief is now selecting the data.

Meanwhile, the seven people without accounts are still without accounts. Nobody is looking for that, because the explanation has already been chosen. Mandatory training rolls out. Some of the people sitting in that training still cannot log in.

That is how a single jump becomes a culture story. It is not dramatic. It is just a ladder, climbed once, and then used as a foundation.

A Pilot Story

Here is one you may know.

Tessa is the VP of people operations at a regional logistics company. She is smart, she is busy, and she championed this AI pilot herself. On a Thursday evening, between a budget review and her kid's soccer pickup, she pastes the usage report into a chatbot and asks for a summary and next steps for Friday's leadership meeting.

Back comes the two-liner you read at the top of this chapter.

Tessa feels a small sting. She pushed for this pilot. Now it looks like her people did not want it. And she feels a bit of relief, too, because the summary also gives her something to do: training. Leaders like having something to do. A problem with an action attached feels half solved.

She drafts an email announcing mandatory training. She does not send it, because her daughter is waving from the car window, and that tiny delay turns out to be the best thing that happens all week.

Hold that picture. I will come back to Tessa. First, I owe you a confession.

I Climbed It Too

On a client project, a member of the team, I will call him Alex, missed a deadline. He did not apologize. Then he went on sick leave.

You can probably feel where this is going, because you can feel your own ladder warming up as you read it.

I climbed fast. Missed deadline, no apology, sick leave. In my mind, Alex was faking. He was unreliable. He was not interested in the work. I did not stop at a private opinion, either. I told the client I would not give Alex more responsibility. That is the top rung: action. Real consequences for a real person.

Here is what I actually had at the bottom of the ladder. One missed deadline. One missing apology. One leave. Three data points, and I built a whole personality out of them. Everything else (faking, unreliable, uninterested) was meaning, assumptions, and conclusions I added. And the action, telling a client something that could affect Alex's future, came from the very top of a ladder I had built from very little. I teach this model. I have taught it for years. I know this ladder by heart. And I took the stairs three at a time.

Looking back, the better move was so simple it is almost embarrassing: check in with Alex and ask what was going on. That is it. One conversation would have put me back on the bottom rung, where the data lives. Maybe I would have heard about an illness, a family situation, or a workload nobody had noticed. Maybe I would have learned that my read was right after all. Either way, I would have been deciding from data instead of from a story I wrote while irritated.

It would be very convenient if the ladder of inference were something only other people climbed. Unfortunately, psychologists are eligible. So are you. I share this because I do not want you to think the ladder is an AI problem. It is a human problem. AI just runs the elevator.

What AI Does Well Here

Now for the good news, and it is genuinely good.

The same tool that can hand you the top rung in one sentence is excellent at helping you climb back down, if you ask it to. One of the most useful things an AI tool does in this territory is generate several competing explanations for the same data. Humans tend to settle on the first explanation that fits. A chatbot does not get attached to its first idea the way we do, and it will happily produce six.

A habit from the scientific method helps here: treat each explanation as a hypothesis, an idea you can test. Suppose projects keep stalling between two departments, and someone proposes merging them. One hypothesis is that duplicate work causes the delays. If so, you would expect to find both departments doing the same tasks. Another is that unclear decision authority causes the delays. Then you would expect to find work sitting, waiting for approval, even when nobody is duplicating it. Those explanations point toward very different actions. Before giving the merger a budget and a new organizational chart, I would give it a little exercise. And one test for your test: if every possible result would count as proof that you were right, you have not given the idea a meaningful test.

Give it the pilot data and ask: "What are five different explanations for why only three of twelve people used this tool weekly? Include explanations that have nothing to do with attitude." You will likely get access problems, unclear use cases, timing, workload, manager support, and yes, resistance. Suddenly resistance is one candidate among several, which is exactly where it belongs.

AI is also very good at separating rungs when you ask it directly. Try: "Split this summary into what the data shows and what is interpretation." It can sort "three of twelve used it weekly" (data) from "resistance" (meaning) in seconds. You can also ask: "What information would I need to tell these explanations apart?" That gives you a short list of questions to ask the actual humans.

The same move works whenever a team seems to be resisting something, which is one of the most common conclusions leaders reach about AI right now. Instead of asking, "How do I overcome my team's resistance to AI?" try something closer to the bottom of the ladder: "Here is what my team is actually saying and doing. What are several possible explanations?" Then give it the rungs a camera would record. "Two people asked whether AI use would affect performance reviews. One person said the tool gave her a wrong answer last week. Three people have not logged in. Our workload doubled this quarter." Leave out your interpretation. Leave out the word "resistant."

Now the tool is working with you at the bottom of the ladder instead of handing you the top. You will likely get several explanations: unclear expectations, a trust problem after a bad experience, access or time barriers, worry about how use will be judged. Some may point at the team. Several may point at the process, or at you. Each suggests a different next step, and each is something you can check. The explanations are hypotheses. Your team has the data.

This is the tool at its best: not as the judge, but as the colleague who says, "Before we decide, have we considered...?" Every team needs that person. Now you can have one available at eleven at night.

ASK, Then Go

Back to Tessa. On Friday morning, before the meeting, she gives the summary one quick ASK.

Assumptions. What am I assuming? She is assuming every one of the twelve had a fair chance to use the tool. That is the assumption hiding underneath "resistance." Once she names it, it looks a lot less solid. Resistance only makes sense if people could have used it and chose not to. She also notices the word where the summary moved from fact to interpretation, and circles it: "resistance." (Its cousins are "disengaged" and "misaligned.") And she looks at

her own question. "Why aren't employees adopting the tool?" would have assumed the issue was the employees before the tool ever answered. "What happened in the pilot?" asks for the bottom rungs. A small change in the question can change which rung the answer starts on. Notice the label, too. "Resistant" is already an interpretation. It is a rung near the top of the ladder.

Source. Where does this come from? The summary came from the usage report, and only the usage report. That report counts logins. It knows nothing about access dates, task assignments, or anyone's attitude. So she opens the pilot coordinator's notes, which she had not attached the night before. That is where she finds the seven without access and the two without a task.

She also checks the one thing the summary said that she could verify: the three weekly users. She pulls the usage log, and yes, this holds. Three people, consistently, every week. The data rung was solid. The tool counted correctly. It was the climb above the data that needed checking, and now she knows exactly which rung gave way.

Key. What's the key piece, and does it unlock what I need? For Tessa, the key piece is the explanation, because the explanation drives the action. If it is resistance, training might help. If it is access, training for people who could not log in is a very expensive way to say "please try harder at the impossible." With the coordinator's notes in hand, the key piece is answered. The decision she needs is whether to extend the pilot with full access, and she now has what she needs to make it.

Then she stops. That is part of the method. She does not commission a survey on organizational readiness. She does not reread the whole report three more times. She found the key piece, checked it, and it unlocked the decision. She closes the laptop and goes to the meeting.

What does she bring? Not a mandatory training plan. She brings this: "Everyone who had access and a task used the tool every week. Let's fix access, give people a real task, and run it again for a full cycle." That is a much better meeting.

Notice the stakes, though. If the recommendation had been to cancel the program or to put people's performance reviews on the line for low usage, she would want to go further. That is when you ask agAI: ask the tool for competing explanations, and more important, ask the people closest to the work. A short conversation with two of the nonusers would tell her more than any summary could. When livelihoods are involved, the bottom rung deserves a visit in person.

And one more question worth asking agAI, the one I skipped with Alex: what did I want to hear? Tessa felt a sting when the pilot looked like a flop. Part of her wanted it to be someone else's attitude problem. "Resistance" is a satisfying explanation for a leader. It puts the problem outside the rollout design and gives you something to fix in other people. "We did not set up the accounts" is a much less satisfying thing to find. That is not a character flaw. It is just useful to know which way your ladder leans.

PRESSURE TEST: COPY THIS

Use this when a summary hands you a conclusion.

"Walk me down the ladder. What is the raw data, what did you select from it, what did you assume, and where did the data turn into a conclusion? What other conclusions could the same data support?"

BOTH THINGS CAN BE TRUE

AI climbed too fast, and AI helped her climb down. The same tool that turned three of twelve into "resistance" also, when asked, produced a list of explanations that put resistance in its proper place. The tool did not change. The question did.

The data was right, and the conclusion was wrong. The three weekly users were real, and checking that took seconds. Trusting the count and questioning the story are not in conflict. They are the same skill pointed at different rungs.

Fast is fine, and the bottom rung is still there. You do not need to walk every rung for every summary. For most things, the top of the ladder is a perfectly good place to work. Just know where the stairs are when the action affects people.

Back to the Two Lines

Look again at what you wrote down at the start of this chapter. If it was "schedule the training," you did nothing foolish. You read a clear summary from a capable tool and took the action it recommended. That is efficient, and most days efficient is right.

But now you know something about those two lines. The first half was data. The second half was a staircase you never saw. And the fix took one more source and a single question: what would have to be true for this explanation to hold?

Maybe you would still offer some training. That might be reasonable, once the seven people have accounts and everyone has a real task to try. You can still land on the original conclusion; sometimes it holds up. The goal is not to never

climb the ladder. We have to draw conclusions, and leaders who stay on the bottom rung forever never decide anything. The goal is to know which rung you are standing on, and how you got there. Twelve people were invited. Three used it weekly. That was true from the start. What changed is which rung you were standing on.

I think about Alex whenever a summary like that one lands in front of me. My ladder was quick, it was confident, and it was missing almost every rung. An AI tool can climb faster than I ever could, which means it can be wrong faster, and it can also lay out five other ladders before I have finished my coffee. I would rather have that colleague than not.

The tool will give you the top floor in seconds. Your job is to know which floor you are on, and to take the stairs back down when someone's livelihood is waiting at the bottom.

THE KEY

When a summary hands you a conclusion, find the rung where data turned into story, and check that one.

CHAPTER 12

Borrowed Confidence

TRY THIS FIRST

Think about a zipper. The ordinary kind on a jacket. On a scale of 1 to 7, where 1 means "no idea" and 7 means "I could explain it to an engineer," how well do you understand how a zipper works? Write your number down. Now explain it, step by step, in writing or out loud to an empty room. What happens when you pull the slider up? How do the teeth actually lock together? Why don't they just slide back apart? Keep going until you have explained the whole thing or run out of road. Now rate your understanding again, 1 to 7. What happened to your number?

If your number held steady, you may be an engineer, a tailor, or a person who once took a zipper apart out of boredom. I salute you.

For most of us, the number drops. I suspect mine would. I would start out somewhere around confident. I have zipped a great many jackets. I have strong opinions about zippers that stick. Then I would get to "the teeth sort of... hook?" and realize I was describing a zipper the way I describe my car's transmission: with hand gestures and hope.

Nothing about the zipper changed in the last minute. Nothing about your knowledge changed either. What changed is that you tried to use the knowledge, and found out how much of it was really familiarity wearing a nice suit.

That is the lock in this chapter. Familiarity feels like understanding. And AI can make that feeling stronger than ever, because now we get to be familiar with answers we did not build.

You may know the workplace version. You had the recommendation, the supporting points, even a lovely little table. Then someone asked, "Why does that make sense for us?" And suddenly, you needed another prompt. Did we develop an understanding, or did we borrow some confidence? The question I keep coming back to is simple: what is our confidence resting on?

The Zipper Problem

In 2002, psychologists Leonid Rozenblit and Frank Keil gave this a name: the illusion of explanatory depth. They asked people to rate how well they understood how everyday things work, objects like zippers, toilets, and locks.

Then they asked people to write out detailed, step by step explanations. After trying to explain, people rated their understanding again.

You already know what happened, because you just did it. The ratings dropped. People had overestimated what they understood, and they did not find out until they tried to explain it. They had not been lying the first time. They genuinely felt they understood. The feeling was real. It just was not measuring what they thought it was measuring.

What I appreciate about this finding is that it is not about intelligence. The people in these studies were not careless. They were human. Our brains are efficient, and they tend to store "I have seen this work many times" in the same drawer as "I know how this works." Most days that is fine. You do not need a mechanical engineering degree to zip your coat.

The problem shows up when we have to act on understanding we do not actually have. When we have to explain a decision to the board. Defend a recommendation to a skeptical client. Answer the follow up question nobody warned us about. That is when the drawer opens and we find out what was really inside.

Here is why this matters. We use that feeling all day long to decide what to double check and what to wave through. If something feels understood, we stop asking. The feeling of knowing is our brain's "done" button, and it can be pressed by things that are not knowledge at all: familiarity, fluency, having seen it before, having heard someone smart say it.

Psychologists have a word for thinking about your own thinking: metacognition. It is the quiet skill underneath almost everything in this book. It is the part of you that can step back and ask, "How do I know this? How sure should I be? What would I need to see to be more sure, or less?" Not "Do I know this?" but "How would I know if I knew this?" The zipper is a cheap, harmless way to find out that your gauge runs a little hot.

Cognitive rigor is metacognition put to work. It means examining your assumptions, checking evidence, and testing whether a conclusion actually holds before you build on it.

There is an old idea in my field that is useful here. In 1974, Mark Snyder described self-monitoring, the way some people pay close attention to social cues and adjust how they present themselves to fit the situation. Snyder was writing about how we manage our behavior in front of others. I want to borrow the idea and point it inward. Monitor your own thinking the way a skilled self-

monitor watches the room. Notice when your confidence rises. Ask what just caused it. Was it evidence? Or was it a well formatted answer that arrived fast?

Here is the reframe that helps me most: treat an AI explanation as a hypothesis, not a conclusion. A hypothesis is something you test. A conclusion is something you act on. The tool can give you an excellent hypothesis. Your job is to decide whether it earns a promotion.

I am an overthinker, so you would think I would be immune. Overthinkers spend plenty of time thinking about thinking. Unfortunately, spending time on it and being accurate about it are two different hobbies.

Smarter, but None the Wiser

Now bring AI into the room.

In 2025, a team led by Fernandes at Aalto University published a study in *Computers in Human Behavior* with a title I wish I had written: "AI makes you smarter but none the wiser." About 500 participants solved logical reasoning problems from the LSAT, the law school admissions test. Half of them worked with ChatGPT.

The first finding deserves a moment of genuine applause: **AI improved performance**. People using it got more answers right. That is not a small thing. Logical reasoning problems are hard, and the tool helped.

The second finding is the one to carry in your pocket. Participants using AI overestimated how well they had done. Their sense of their own performance ran ahead of their actual performance.

There was a twist, too. Normally, in tasks like this, you see a familiar pattern often linked to the Dunning and Kruger research: the people who score lowest tend to overestimate themselves most, and the strongest performers tend to be more modest. With AI in the mix, that pattern appeared to flatten. People across the range tended to overestimate. Scores went up, but people's sense of how well they did went up even further. Misjudging your own performance became more of a group activity.

And here is the part that made me put my coffee down. People who rated themselves higher in AI literacy tended to be **more** overconfident, not less. The people who felt most fluent with the tool were the ones whose sense of their own performance drifted furthest from reality. Knowing a lot about the tool did not seem to protect them from trusting their own reading of the tool's output. If anything, it may have made the "done" button easier to press.

One more detail tells you a lot about how this happens: most participants used a single prompt per question. They asked once, got an answer, and moved on. No follow up. No "why." No "what if you are wrong." One question, one answer, one confident rating. That makes sense. It was fast and it worked more often than not. But a single prompt gives you very little chance to notice where your understanding ends and the tool's begins.

The study had its limits, as every study does. It was logic puzzles, not a quarterly forecast. But the pattern suggests something every leader should take seriously: using AI can improve the work while quietly inflating your sense of how much of that improvement is yours.

I call this borrowed confidence. You borrowed the answer. Without noticing, you also borrowed the feeling of having figured it out.

Confidence That Does Not Check Its Sources

If you read the first chapter, you met Wharton researchers Shaw and Nave and their idea of cognitive surrender, adopting an AI's answer as your own with very little scrutiny. I will not repeat the whole study here, but one finding belongs in this chapter. In their experiments, confidence went up when people used the AI. It went up when the AI's advice was right. It also went up when the AI's advice was wrong.

Sit with that for a moment. The feeling of certainty did not track whether the answer was correct. It tracked whether an answer had arrived. (Theirs is a working paper, so it is still going through the usual scholarly review. The pattern lines up with what we just saw from Aalto.)

Your confidence is supposed to be a gauge. When the gauge rises whether the tank is full or empty, it is not a gauge anymore. It is decoration. Confidence, it turns out, is easy to borrow. The trouble is that it does not come with a receipt.

The Jagged Edge

Why does this matter if the answers are better? Because they are not better everywhere.

In 2023, Dell'Acqua and colleagues at Harvard Business School, working with Boston Consulting Group, studied consultants using AI on realistic consulting tasks. They described what they called a "jagged technological frontier." AI is very good at some tasks and surprisingly weak at others, and the boundary between them is uneven and hard to see from where you are standing. For

tasks inside that frontier, AI helped the consultants. For a task outside it, consultants using AI performed worse than those working without it.

The frontier is jagged, meaning you cannot see it from where you stand. Two tasks that look equally hard to a human can sit on opposite sides of the line. The tool sounds exactly as fluent on both.

Put that next to the Aalto findings and the picture gets clear. If AI raises your confidence across the board, and AI's actual competence is jagged, then your confidence may stay high even in the places where the tool is weakest. You will not feel the edge of the frontier. You will walk right past it, feeling great. Your confidence cannot tell you where the frontier is.

That is the readiness question worth asking of any team: not whether your people can get an answer out of the tool, but whether they can tell when they have wandered past the edge of what it does well. For that, you need another gauge. And the best gauge we have is the zipper test: can I explain this, step by step, in my own words?

A Forecast and a Board

Here is a story you may recognize.

Jordan is a regional VP of sales at a growing manufacturing company. He is good at his job, and he is stretched. The quarterly board meeting is coming up, and he needs a sales forecast for his region. He gives an AI tool his pipeline data, last year's results, and a few notes on market conditions, and asks it to build a forecast with a short rationale.

What comes back is genuinely strong. It is clean, the logic is laid out, and the numbers look sensible. Jordan reads it twice, adjusts a sentence, and puts it in the deck. He feels ready. He has read it, after all. He knows what it says.

At the board meeting, the forecast goes up on the screen. It is received well. Then one board member leans forward and asks a perfectly friendly question: "Why does that make sense for us?"

Jordan opens his mouth and discovers he is holding a zipper.

He can recite the forecast. He can repeat the rationale, nearly word for word. But "why does that make sense for us" is a different question. It asks him to explain, step by step, what the forecast assumes about his region, his customers, and his team, and why those assumptions fit. The confidence he walked in with was real. It was also borrowed.

He does what good leaders do in that moment. He says, "Let me walk you through the main drivers, and I will follow up with more detail on one of them." He gets through it. Nobody faints. But on the drive home he knows exactly how that felt, and he would rather not feel it again.

Notice that nothing went wrong with the forecast. Nothing went wrong with Jordan's intelligence. What was missing was ownership of the reasoning. That is a much easier problem to fix.

What AI Does Well Here

Let me be very clear about something, because it would be easy to read this chapter as "use AI less." That is not the lesson. The Aalto participants who used AI got better answers. Dell'Acqua's consultants did better work on the tasks inside the frontier. The tool helped. The goal is not less AI. The goal is knowing what you know.

And here is the delightful part: AI is one of the best tools ever built for finding out what you know.

Ask it to explain its own forecast step by step, as if to someone new to the business. Then close that window and try explaining it yourself. Where you stall is where your understanding stops.

Ask it: "What are the three assumptions this forecast depends on most?" It is very good at surfacing the load bearing parts of its own reasoning when you ask directly.

Ask it to play the board member. "Ask me the five hardest questions a skeptical director might ask about this forecast, one at a time, and tell me what a strong answer would need to include." That is rehearsal on demand, and it is excellent at it.

Ask it to quiz you. Ask it to find the weakest link. Ask it to explain the same logic three different ways until one clicks.

There is an old name for this kind of questioning. Socrates, the Greek philosopher, was famous for teaching without lecturing. He asked questions, one after another, until the person he was talking with could see what they actually knew and what they had only assumed. In Plato's account of his trial, Socrates says the only wisdom he could claim was that he did not think he knew what he did not know. That is the zipper problem, about 2,400 years early.

The Socratic method still lives in classrooms and law schools, and its questions tend to fall into a few families. What do you mean by that? What are you assuming? What is your evidence? And if that is true, what follows? You may notice the A and the S of ASK hiding in there.

Here is the twist I love: AI can play Socrates. Instead of asking it for answers, ask it to question you, one question at a time, and to wait for your answer before asking the next. The Pressure Test at the end of this chapter does exactly that. Socrates never handed anyone the answer. He handed them the next question.

Used like this, the tool is not doing your thinking for you. It is a patient tutor and sparring partner that turns borrowed confidence into earned confidence. That is a very good use of an evening.

ASK It

Here is what Jordan does the next time, and it does not take long.

First, one small habit borrowed from research. In 2021, Bućinca, Malin, and Gajos tested ways to reduce overreliance on AI recommendations. One approach was simply giving people explanations of the AI's reasoning. That helped less than you might hope. Another approach used what they called cognitive forcing functions, design choices that make you think before you lean on the tool. One example: make your own decision first, and only then see the AI's recommendation. The cognitive forcing functions reduced overreliance more than simple explanations did.

Now for the uncomfortable part. People rated the most effective designs least favorably. They did not like them. The checks that worked best were the ones that felt most like work. The researchers also found that people who enjoy effortful thinking benefited more. The medicine worked, and nobody loved the taste.

I find this both annoying and reassuring. Annoying, because I would love for the best method to also be the most pleasant one. Reassuring, because it explains why so many of us skip the check. It is not a character flaw. The friction is real, and our brains are built to avoid it. (You met the cognitive miser back in the first chapter. It has not left the building.) So if the protective habit feels a little irritating, that may be a sign it is working.

So before Jordan opens the AI tool, he spends a few minutes writing his own rough view of the quarter. Up, down, or flat, and why. It is not elegant. It does not need to be. Then he asks the tool for the forecast, and afterward he notes

what changed between his view and the tool's, and why. If the answer is "because the AI said so," that is worth a second look. If it is "because it pointed me to something in the data, and I checked it," that is learning. Both involved AI. Only one leaves him wiser. That note turns out to be his best board preparation.

Assumptions. What am I assuming? He asks the tool to list the forecast's main assumptions, and he adds one of his own: he has been assuming that because he read the rationale, he understands it. The zipper says otherwise, so he tests himself. He explains the forecast out loud, in his office, to nobody. He gets stuck on one point: the forecast assumes deals in his pipeline will close at roughly the rate they have historically.

Source. Where does this come from? That close rate came from last year's results, which he supplied. So he goes back to the source and checks whether the region's recent history actually looks like that.

Key. What's the key piece, and does it unlock what I need? The close rate is the key piece. Most of the forecast hangs on it. He compares it to his team's record over recent quarters, and he talks briefly with two of his sales managers about what is in the pipeline right now. And yes, this holds. The assumption fits his region well. The tool did good work. The difference is that now Jordan knows **why** it did good work, and he can say so in his own words.

Then he stops. He does not rebuild the forecast from scratch. He does not run it through four more tools. He found the key piece, checked it, and it held. He can answer "why does that make sense for us?" in two sentences. Close the laptop.

If the forecast were going to drive something heavier, like a hiring freeze or a cut to a team, that would be the time to ask agAI: ask the tool for the case that the forecast is too optimistic, ask the people closest to the customers, and ask himself which number he was hoping to see. For a quarterly forecast he can stand behind, he is done.

PRESSURE TEST: COPY THIS

Use this before you present or defend an AI answer as your own.

"Ask me three questions a skeptical colleague would ask about this answer. Wait for my response to each one before you ask the next. Tell me where my explanation was thin."

BOTH THINGS CAN BE TRUE

AI made the answer better, and it made the confidence bigger than the understanding. Both findings came from the same study. You do not have to pick one. Performance and self-assessment are separate skills, and the tool can lift the first while quietly distorting the second, because an answer that arrives quickly and reads well feels like understanding.

Borrowed confidence is not fake, and it is not yours yet. Jordan really did feel ready. The fix was not doubt. It was a few minutes of explaining the reasoning in his own words until the confidence belonged to him.

The checks that work best can feel the least pleasant, and they are often the fastest. Writing your view first takes a few minutes and may feel like a chore. It also tells you, very quickly, exactly what you know and what you are borrowing.

Back to the Zipper

Go back to the two numbers you wrote at the start of this chapter. If the second one was lower, nothing bad happened. You did not lose any knowledge. You gained some, actually: you now know where your understanding of zippers ends. That is metacognition, and it is a quietly powerful thing to have.

The same small move works with every AI answer that matters. Read it. Then try to explain it, step by step, without looking. Watch your number. Where it drops is exactly where to look next, and where it holds is where you can go ahead and move fast.

AI is going to keep making our answers better. I am grateful for that, and I use it every day. I just want to be able to explain the zipper when someone at the table leans forward and asks.

THE KEY

Before you stand behind an AI answer, explain it in your own words; the confidence that survives is yours.

CHAPTER 13

Who Gets the Blame?

TRY THIS FIRST

Think of the last time someone cut you off in traffic. Picture the car, the lane, the moment your foot found the brake. Now finish this sentence out loud: "That driver was ____." Got it? Now think of the last time you cut someone off. Be honest; it happens to all of us. Finish this one: "I did that because ____." Look at your two sentences side by side. What did you blame in the first one? What did you blame in the second?

If your first sentence ended with a word like "careless," "rude," or something I will not print in a book, welcome. You are in excellent company. And if your second sentence started with "because," followed by a reason (the sun, the merge, the kid in the back seat, the GPS that waited until the last second to mention your exit), you are in excellent company again.

Notice what happened. The other driver got a verdict about their character. You got an explanation about your circumstances. Same behavior. Different courtroom.

That is the lock in this chapter. And it matters more than ever, because there is a new driver on the road at work, and it is very easy to blame.

The Driver in the Next Lane

I live in Atlanta, which means I have had a lot of practice with this demonstration. Not on purpose.

One day a car kept drifting into my lane. Once, fine. Twice, I tightened my grip. The third time, I honked, and I did not need a full minute to decide who this person was. Careless. Aggressive. Probably on the phone. Definitely the kind of person who does not deserve a driver's license or a parking space.

Then I pulled alongside him.

He did not look aggressive. He looked like he was falling asleep, or possibly having some kind of health crisis. The verdict I had delivered with such confidence a few seconds earlier fell apart, and I went from honking at him to trying to call for help.

Here is the part that keeps me humble. Another time, I was the one drifting. A deer came out of nowhere, I swerved to avoid it, and I ended up in someone else's lane. That driver gave me exactly the look I had given the man in the car

beside me. I could see it on their face: careless, aggressive, who taught her to drive? They did not see the deer. They only saw me.

Neither of us was wrong about the behavior. A car really did come into a lane where it did not belong. We were wrong about the cause.

What Psychologists Call It

Psychologists have a name for this habit: the **fundamental attribution error**. When we explain other people's behavior, we tend to give too much weight to their character (their personality, their attitude) and too little weight to their situation. We see the action and jump straight to "that is the kind of person they are." When we explain our own behavior, the situation suddenly becomes very visible. I swerved because of the deer. You swerved because you are a menace.

It has a twin that we like much better, called the **self-serving bias**: our tendency to take credit for our successes and blame circumstances for our failures. When my project succeeds, it is because I am skilled and prepared. When my project fails, it is because the timeline was unrealistic, the data was late, and Mercury was in retrograde. The promotion was talent. The missed target was the economy.

Put the two together and you get a very comfortable way of seeing the world. Other people fail because of who they are. I fail because of what happened to me.

The urge to supply a story runs deep. In 1944, psychologists Fritz Heider and Marianne Simmel showed people a short animated film: two triangles and a circle moving in and around a rectangle. That is all it was. Shapes. Yet nearly everyone who watched described a story, with characters, motives, a bully, and a chase. The film had no intentions in it. The viewers brought them. Heider went on to become one of the founders of attribution research, the study of how we explain why things happen.

I do not say this to scold anyone. These habits save the brain time. Deciding "that person is careless" is fast and tidy. Gathering the facts about the deer, the sleepiness, and the sun in their eyes is slow and messy, and you are going highway speed with a coffee in the cup holder. Your brain picks tidy.

The trouble comes when we carry the fast, tidy verdict into places where the stakes are higher than a honk.

Kelley's Three Questions

So how do we get better at this? In 1973, the psychologist Harold Kelley offered a useful framework. When you are trying to figure out whether someone's behavior comes from them (an internal cause) or from the situation (an external cause), he suggested looking at three kinds of information instead of jumping to a verdict.

- **Consensus:** Do other people in the same situation behave the same way? If everyone is struggling, the situation may be the cause.
- **Consistency:** Does this person behave this way in this situation over time? If it happens every time, it may be something about them.
- **Distinctiveness:** Does this person behave this way across many different situations, or only this one? If it only happens here, the situation may be the cause.

Put them together and a pattern appears. When consensus is low, consistency is high, and distinctiveness is low, the cause is more likely inside the person. When lots of people struggle, the person is usually fine, and it happened only this once, the cause is more likely in the situation.

I wrote about this in *Stop, Look, and Listen* using two employees I called Nestie and Jack. Both had performance problems. On the surface, they looked the same. Underneath, they could not have been more different.

Jack's teammates were doing fine with the same work (low consensus). Jack's problem showed up again and again, week after week (high consistency). And it showed up across many different kinds of tasks, not just this one (low distinctiveness). That pattern points toward an internal cause. Something about how Jack works may need attention, and that calls for a direct, fair conversation with Jack about his performance.

Nestie was another story. Several of her teammates were also struggling (high consensus). She was usually a high performer, so this was not her pattern (low consistency). And the trouble showed up on one particular task, a one time assignment (high distinctiveness). That pattern points toward a situational cause. Something about the task, the resources, or the setup may need attention. Maybe the task was badly defined, the deadline unrealistic, the tools wrong. Coaching Nestie harder would miss the point. Fixing the task would not.

Same outcome on paper: a performance problem. Very different explanations once you look at the data instead of your first reaction. Two employees, two completely different right answers. If a manager had blamed both of them the

same way, one of those conversations would have been fair and the other would have been the man in the car all over again.

Hold on to Nestie and Jack. We are about to bring them to work with AI.

The New Driver on the Road

Now bring an AI tool into the picture.

You may have seen this one. Nadia is an account manager at a facilities services company. She is good at her job, well liked by her clients, and juggling eleven things at once. On a Thursday afternoon, she asks an AI tool to draft a follow up email to a client summarizing a renewal. She pastes in her notes, the client's questions, and a pricing file. The draft comes back clean, warm, and well organized. She reads it for tone, likes it, and sends it before her 3:05 meeting.

Friday morning, the client replies. Politely. The price in the email is last year's price.

Now the meeting starts. And you can almost hear the attributions lining up.

"The AI hallucinated."

"Nadia didn't check it."

"These tools aren't ready for client work."

"We need a policy."

Look closely and you will see the traffic demonstration at work. "The AI hallucinated" is a very convenient sentence. It puts the cause in the tool, the way "I swerved because of the deer" puts the cause in the deer. And "Nadia didn't check it" puts the cause in the person, the way "that driver is careless" does. Either can be true. Neither is a conclusion yet. They are verdicts that skipped the investigation.

To be clear, sometimes "the AI hallucinated" is exactly true. As we saw in Chapter 8, language models can produce fluent, confident answers that are simply wrong, and that is a normal failure mode we should plan for. I am not asking anyone to pretend otherwise. But listen to how the sentence works. It explains the failure entirely by the character of someone else (in this case, something else) and leaves our own situation, choices, and process completely out of frame. Flip it around and it becomes the new excuse: I did not make a mistake. The tool did.

I have a growing suspicion that "the AI hallucinated" is becoming the new "the other driver is a jerk." It is quick, it is satisfying, and it closes the case before anyone has opened it.

Was It the Tool, the Prompt, the Process, or the Person?

When AI work goes wrong, there are usually four places to look. Kelley's questions work beautifully on all four.

The tool. Does this tool make this kind of mistake for everyone who uses it this way (consensus)? Does it happen every time (consistency)? Does it happen only with numbers pulled from long files (distinctiveness)? If many people get the same error from the same kind of request, you may have learned something real about the tool.

The prompt. What did the tool actually receive? In Nadia's case, someone finally opens the pricing file she pasted in. It is the old one. The file name on the shared drive had not changed when the new prices were added in a different version. The tool did exactly what it was asked to do: it used the price it was given. The tool did not invent a number. It faithfully copied the wrong one. (The prompt can also invite trouble directly. A request to "include compelling statistics" with no sources attached is not only a hallucination problem. It is a question that invited one. Many "hallucinations" begin with a question that made one likely.)

The process. Is there a step where a figure in a client email gets checked against the current contract? Was there one before AI arrived? (Sometimes there was, and it quietly disappeared because the drafts started looking so finished.) Does the shared drive make it easy to grab the wrong file? If three people on the team could have made the same mistake with the same file, consensus is high, and the cause is mostly in the process.

The person. Is this Nadia's pattern? If she usually checks numbers carefully and this happened once, on one rushed afternoon, the evidence points away from her character. If she has sent several unchecked figures out the door this quarter, that is a different conversation, and a fair one to have.

Nestie and Jack Go Digital

Now run the full Kelley walkthrough on Nadia's email, the way I did with Nestie and Jack, because the same wrong figure can come from two very different patterns.

Suppose the team looks around and finds this. **Consensus:** two other account managers pulled prices from the same pricing file this month, and one of them had the same old figure in a draft that happened to get caught. Several people, same situation, same problem: consensus is high. **Consistency:** Nadia's recent client emails, month after month, had correct figures. This is not her pattern, so consistency is low. **Distinctiveness:** the error shows up only in renewal emails that pull from that one file, not in her reports, proposals, or summaries. Distinctiveness is high. That is a Nestie pattern. The cause is mostly situational: a stale file that looks current and a process with no step to check the figure. Coaching Nadia harder would not fix it. Banning the tool would not fix it. Archiving the file and adding a check would.

Now imagine a different set of facts. **Consensus:** colleagues using the same tool, the same file, and the same kind of request are sending clean emails. Only Nadia's go out wrong, so consensus is low. **Consistency:** this is one of several unchecked figures from her this quarter, so consistency is high.

Distinctiveness: the same thing turns up in her proposals and her reports too, not just renewal emails, so distinctiveness is low. That starts to look like Jack. The cause is more likely in how Nadia is working with AI, and the right conversation is about her habits: what she checks, what she skips, and what support or expectations need to change. That conversation is not cruel. It is fair, because the data earned it.

Same email. Same wrong price. Two different answers, and you only know which one you have after you look. Either way, notice what we did. We looked at the data before deciding who gets the blame.

Notice what this does. It does not let anybody off the hook. It puts the hook where it belongs.

And here is the line I want to hold firmly. Whatever the cause turns out to be, **accountability stays with the person who acted.** The tool does not sign the email. Nadia does. The researchers behind the mammography study I described back in Chapter 1 made the same point in a very different setting. You may remember that Dratsch and colleagues found that when an AI system suggested the wrong category for a mammogram, radiologists' accuracy dropped, including very experienced radiologists. That is automation bias: leaning on an automated recommendation more than we should. One of the safeguards the researchers recommended deserves attention here. They suggested that users should feel accountable for their own decisions. Not the system. The user.

That recommendation is simple, and it is powerful. When you know the decision is yours, you look harder. When you believe the tool is responsible, you look less. The sense of ownership is not just an ethical position. It appears to be a practical protection against the very errors we are worried about. Accountability is not punishment, and it is not about blame for its own sake. It keeps our attention on the part of the process we can actually change. It is the reason you read the key number before you click send.

So two things are true at once. Nadia owns the email. And the fix may have very little to do with Nadia.

What AI Does Well Here

Let me be fair to the new driver, because it deserves it.

AI is genuinely excellent at the work that sits around a mistake like this. Ask it to compare two versions of a pricing file and it will show you the differences in seconds, the kind of tedious line by line comparison that most of us do badly at 4:45 on a Friday. Ask it to list every figure in a draft along with where each one came from in the material you gave it, and it will hand you a tidy checklist. That one request alone would have caught Nadia's error.

It is also a strong partner for the investigation itself. Describe the situation and ask, "Give me several different explanations for how this happened, including ones involving the tool, the instructions, the process, and the person." You will get a range of causes, and that range is exactly the antidote to a fast verdict. It will not get tired, defensive, or embarrassed, which is more than most of us can say in a post mortem meeting.

And it can help with the repair. A well framed request will produce a clear, gracious correction to the client in a few minutes: owning the error, giving the right figure, and not groveling. Nadia still decides what goes out. But the tool can get her from "oh no" to "here's what we're sending" much faster than staring at a blank screen.

Most of the time, these drafts are right. That is why teams use them. The goal is not to stop drafting with AI. It is to know which piece of the draft you check before it leaves the building.

ASK It

Let's run Nadia's situation through ASK. Not the six month review. The quick version.

A: Assumptions. What am I assuming? The room was assuming one cause: either the tool or Nadia. The honest assumption to name is "I think I already know why this happened." That is the driver in the next lane talking. Write down at least two other possibilities before anyone decides.

S: Source. Where does this come from? For the figure itself, the source was the pricing file. Open it. That single step moved the whole conversation. Once the team saw the old file, "the AI hallucinated" was no longer an available explanation, and "Nadia is careless" looked a lot less likely too.

K: Key. What is the key piece, and does it unlock what I need? In a client email, the key is usually the number. Not the tone, not the greeting. The figure the client will act on.

So Nadia's manager does something simple for the correction. She asks the tool to list every figure in the original email with its source, then checks each one against the current contract. The renewal price was wrong, which they knew. The service dates, the number of locations, and the payment terms all match the contract. **Yes, this holds.** The rest of the email was accurate. They correct the one figure, send a clean note to the client, and move on.

Then the stopping rule. The team does not need a three week inquiry or a new AI policy for one email. They need two small changes: the old pricing file gets archived so nobody grabs it again, and anything with a price in it gets the "list every figure and its source" check before it goes out. That takes about a minute. If it holds, you are done. Close the laptop.

If this were bigger (a contract worth a lot of money, a client relationship on the line, a pattern of errors across several people), that is when you would ask agAI: ask the people closest to the work, look at several weeks of emails instead of one, and ask yourself whether you wanted a particular person or tool to be the cause.

Credit Goes Both Ways

Here is the part I did not expect when I started thinking about this chapter.

Teams that blame badly also tend to credit badly.

When an AI assisted project goes well, listen to how people describe it. "The AI did it." Or the opposite: "I did it; I just used AI for the typing." Both are attribution errors in nicer clothes. The self-serving bias is happy to take credit for wins ("I put together a great analysis") and hand the losses to the tool ("The AI got it wrong"). That is self-serving bias in a new costume, and it is costly. We learn nothing from failures, because they were the tool's fault, and we

overestimate our skill from successes, because they were ours. If you read the last chapter on borrowed confidence, you can see how these two patterns feed each other.

A more honest accounting sounds like this. When it goes well: "The tool gave me a strong draft, and I checked the numbers and sharpened the argument." When it goes badly: "The tool gave me a draft with an error, and I did not have a step that would have caught it." Both sentences include the tool. Both sentences include you.

Kelley's questions work just as well on success. Suppose Nadia's colleague uses an AI tool to build a renewal summary that a client loves so much they ask for one every quarter. What made it good? Maybe the tool organized the information beautifully (the tool). Maybe he gave it a precise request with the client's real priorities spelled out (the prompt). Maybe the team had a habit of pulling numbers only from the current contract (the process). Maybe he knew this client well enough to cut the paragraph they would never read (the person). Probably all four.

When you credit accurately, you learn what to repeat. If the win came from the prompt, share the prompt. If it came from the process, protect the process. If it came from the person, say so out loud, by name, in the meeting.

A healthy team looks at success and failure the same way. Same questions, same curiosity, same willingness to find out it was the deer.

PRESSURE TEST: COPY THIS

Use this when something went wrong, or right, and everyone wants a verdict.

"Before we assign blame or credit, help me separate what happened into four possible causes: the tool, the prompt, the process, and the person. What evidence would point to each one?"

BOTH THINGS CAN BE TRUE

Accountability is real. The person who sends the email, signs the report, or makes the call owns it, with or without AI. "The tool did it" is not a defense.

Causes are complicated. Owning a decision does not mean the person was the only cause. The tool, the prompt, and the process deserve the same honest look.

Credit and blame use the same lens. If you would investigate a failure, investigate a success too. That is how a team learns what to repeat and what to fix.

Back in Traffic

Go back to your two sentences from the start of the chapter. "That driver was ____." "I did that because ____."

Next time something goes wrong with an AI assisted piece of work (yours or someone else's), try giving it the second sentence instead of the first. Not "the AI is unreliable" or "she is careless," but "this happened because ____," and then go find out what goes in the blank. Was it the tool, the prompt, the process, or the person?

Sometimes it really is the person. Sometimes it really is the tool. And sometimes there is a deer you could not see from your lane.

The goal is not to find someone to blame. The goal is to find what you can change. And the part you can change most reliably is the part that was yours: the question you asked, the check you ran or skipped, the moment you decided an answer was good enough to act on. The tool can drift into your lane. You are still the one with your hands on the wheel.

THE KEY

*Before you decide who to blame, find out what happened;
before you take the credit, find out what worked.*

CHAPTER 14

Don't Take Away the Radio

TRY THIS FIRST

Grab a pen and do this one without AI or a calculator if you can. You are registering for a conference. Registration is \$240. There is a 15 percent early bird discount. You also want a workshop add-on for \$35. What is your total? Work it out and write the number down. Now open an AI tool and ask it a similar one: registration is \$360, the early bird discount is 20 percent, and the workshop add-on is \$45. Read its answer. Here is the real question: a week from now, which of those two problems will you still know how to solve? What did you do?

Be honest. Did you do the first one, or did you skim it and think, "I get the idea"? If you skimmed, welcome. That is the cognitive miser from Chapter 1 doing exactly what it was built to do: save effort whenever it can.

If you did the first one by hand, you probably did something your brain found mildly annoying. You figured out 15 percent of 240 (some of us went 10 percent plus half of 10 percent, some of us went 0.85 times 240, and some of us stared at the ceiling for a while). You landed on \$204. Then you added the \$35 and got \$239.

And somewhere in there, maybe without noticing, you made a decision: does the discount apply to the workshop too? I assumed it applied to the registration only, which is how I read it. If you applied it to everything, you got a different total. Neither of us is crazy. We just answered a question the problem never asked out loud. That little decision is part of the thinking, and it is exactly the part that tends to disappear when someone else does the math for us.

That little bit of friction is not a flaw in the exercise. It is the point of the exercise.

The AI answer, meanwhile, probably arrived in a second or two, neat and correct: \$360 less 20 percent is \$288, plus \$45 is \$333. Lovely. Fast. It looked right, and you nodded. Your brain watched the solution. It did not build it. And if you are like most people, it is already fading.

That is the lock in this chapter. Not that AI got it wrong. It got it right. The lock is that the effort we skipped was the part that would have stayed with us.

The Radio

Anyone who was ever a teenager knows how this goes. There were stations you were not allowed to listen to. A parent would say, "Not in this house," and that was the end of it.

Except it was not the end of it. It was the end of listening to it when the parent was around. Kids found a way. A friend's car. A friend's house. A radio turned down very low with an ear pressed against it. Removing the station never stopped anybody from listening. It just meant they listened without anybody to talk to about what they heard. Removing something rarely removes the desire for it. It usually just removes the adult from the conversation.

I think about that radio every time I hear a debate about banning AI in schools, and it applies just as much to our workplaces. When the answer to "people are using AI badly" is "then nobody uses AI," we have not built better judgment. We have built a hiding place. People still use the tool. They just use it on their phones, in a personal account, after hours, with no guidance and no one to ask, "How did you check that?" So the question is not really "radio or no radio." The question is whether anyone is teaching the listener how to listen.

I understand the instinct completely. If a tool can write the essay and solve the equation, the fear is that students will let it, and never learn to do either. That fear is not silly. As you will see in a moment, there is research that takes it seriously.

But taking away the radio was never the answer. Building a listener who could think about what they heard was the answer. Same here. Keep the tool. Build the thinker.

Handing Off the Thinking

Psychologists Evan Risko and Sam Gilbert, writing in 2016, gave us a useful phrase: **cognitive offloading**. It means using something outside your head to reduce the mental work inside it. A reminder note on the fridge. A calendar alert. Tilting your head to read sideways text instead of rotating it in your mind. A grocery list. A calculator. A sticky note on your monitor. We do it constantly, and offloading is not a moral failing. It is one of the smartest things human beings do. My brain has better things to do than remember the dry cleaning. I do not want to memorize every meeting on my calendar. I want my brain free for the meetings themselves.

The distinction that matters is what you offload. Offloading a narrow task while keeping the reasoning is different from surrendering the reasoning.

When I use a calculator to total a budget, I still know what the numbers mean, why the line items are there, and whether the total makes sense. If the calculator tells me a two day workshop costs more than a new building, I will notice. The calculator did the arithmetic. I kept the judgment.

When someone asks AI to "write the recommendation" and pastes it into the deck, something else has happened. The arithmetic, the reasoning, the framing, and the conclusion all left the building together. That is what Shaw and Nave call cognitive surrender, which we met in Chapter 1: adopting an AI's answer as your own with minimal scrutiny.

Same with the opening problem. If I ask AI to do the arithmetic after I have decided what the discount applies to, I have offloaded a calculation. If I hand it the whole problem and copy down the answer, I have offloaded the thinking, including the decision I did not know I was making. So the useful question is not "Did you use AI?" It is "What did you hand over, and what did you keep?"

There is a second idea that helps here, from Robert Bjork, a psychologist who has spent his career studying how people learn. He coined a phrase I love: **desirable difficulties**. Some kinds of struggle during learning feel slower and less pleasant in the moment but tend to produce stronger, longer lasting learning. Pulling an answer out of your own memory, working a problem before seeing the solution, wrestling with a concept a bit longer than feels comfortable: these feel harder, and that appears to be part of why they work.

This is deeply inconvenient news for anyone who enjoys efficiency. Which is all of us.

It also explains the exercise at the start of this chapter. The conference registration problem felt harder because your brain was doing the building. That effort is the part that tends to leave something behind.

Sir Ken Robinson, the education thinker whose talk "Do Schools Kill Creativity?" became one of the most watched TED talks ever, made a related point with a paperclip. He described a test of divergent thinking, the ability to see many possible answers to a question: how many uses can you come up with for a paperclip? Most of us come up with ten or fifteen. People who are very good at it come up with around two hundred, partly because they question the question. Does the paperclip have to be the size we picture? Does it have to be made of wire? If you tried the paperclip experiment from Chapter 10, you have already seen this test in action. He cited a study in which about 98 percent of kindergartners scored at what the researchers called genius level.* When the same children were tested again a few years later, the share had dropped sharply, and by their early teens it had dropped further still.

What happened in between? Robinson's answer: school. For years, children are told, in his words, that "there's one answer, it's at the back, and don't look, and don't copy, because that's cheating." Outside of school, he points out, that is called collaboration.

Look at what that teaches. The answer has already been decided by someone else, the teacher or the back of the book. Your job is to match it. Not to question it. Not to imagine other answers. Not to check it with the person next to you. Over the years, many of us learn that asking agAI can look like doubting the authority, and challenging the answer can look like breaking the rules.

Now put an AI tool in front of that habit. It produces a polished, confident answer in seconds, and it sounds a lot like the back of the book. The old training says: match it, accept it, move on. Asking agAI is the opposite habit. It means questioning the question, looking for the second and third possible answers, and checking with others, which is the very thing school once called cheating.

If Robinson is right, the capacity those kindergartners had does not disappear. It just gets less practice. AI can strengthen the one-answer habit, or it can help us rebuild the paperclip habit, if we ask it for more than one answer and do not treat the first one as the back of the book. Working with AI is collaboration too. The goal is not to copy its answer. It is to think alongside it, ask agAI, and work together until the answer is right.

Put those two ideas together and you can see the risk. AI is the most efficient remover of difficulty ever invented. That is much of why we love it, and most of the time it is a gift. The trick is learning to tell which difficulties are just friction (formatting a table, reorganizing notes, drafting a first version of a routine email) and which difficulties are the learning itself (understanding why the forecast changed, figuring out what the client actually needs, deciding which risk matters most). Remove the first kind freely. Protect the second kind on purpose.

What Happened in the Math Classroom

In 2025, Hamsa Bastani and colleagues published a study in *PNAS* that put this to a real test, and I think every leader thinking about AI should know it. They worked with nearly 1,000 high school math students, split into three groups. Some students practiced with unrestricted access to ChatGPT, a version the researchers called GPT Base. Some practiced with a version designed to act

like a tutor, giving hints rather than answers, called GPT Tutor. And some practiced without AI at all.

During practice, the students with unrestricted ChatGPT did better on the practice problems. That part is not surprising. When a tool can hand you the answer, you get more answers right. Good news, right?

Then came the exam, and the AI was removed.

The students who had practiced with unrestricted access performed worse on the exam than students who had never had access at all. Read that again, because it is the heart of this chapter. They did not simply lose their advantage. They ended up behind the students who had done the work the slow way. The help during practice appears to have replaced some of the learning instead of supporting it.

Now here is the part I want you to remember, because it is the hopeful part, and it is a big one. The tutor version told a different story. When the AI was designed to give hints and guide the student through the thinking rather than hand over the result, the harm was largely mitigated. Same technology. Different design. Very different outcome.

A few honest limits. This was one study, with high school students, in math. It does not tell us everything about every subject or every kind of learner, and workplace learning is not identical to a high school exam. Still, the pattern appears to point in a clear direction: the same underlying technology can help or hurt learning depending on how it is used. The problem was not that AI was in the room. The problem was AI doing the part of the work that was supposed to build the student.

Notice what the study did not find. It did not find that AI is bad for learning. The tutor version is the proof. It found that answers without effort may feel like progress while quietly skipping the learning.

You may remember the Harvard physics study from Chapter 6. Kestin and Miller ran a randomized study with 194 Harvard physics undergraduates across two lessons, and a custom AI tutor built with expert written guidance produced median learning gains more than double those of an active learning classroom. I will not repeat the whole study here. It has real limits, which we covered there: two lessons, a custom tutor rather than an ordinary chatbot, and measurement right after the lessons rather than long term retention. But notice what it has in common with GPT Tutor in the math study. In both cases, people who understood learning designed the AI to support thinking, not replace it. The tool kept the student doing the thinking.

That is the design lesson for leaders. Most AI tools, left to their defaults, behave like answer machines. You ask, they answer. If you want a tutor, you usually have to ask for one. Here is what that can sound like:

- "Do not give me the answer yet. Ask me questions that help me work through this."
- "I am going to explain my reasoning. Tell me where it is weak, but do not rewrite it for me."
- "Give me a hint about what I am missing, not the solution."
- "Quiz me on this report so I can see whether I actually understand it."
- "Here is my draft. What question would a skeptical reader ask first?"

Same tool. Completely different relationship with your own brain.

The School Board Meeting

Here is a story you may recognize. A school board is debating a proposal to block AI tools on every school device and network. The meeting runs long. A parent says her son finished his history essay in eleven minutes and cannot tell her what it is about. A teacher says many of her students are using AI at home anyway, ban or no ban. Another board member points out that these students will walk into jobs where AI is part of the work on the first day.

All of them are right. That is what makes it hard.

The ban is the radio solution. It will feel decisive. It will probably reduce AI use inside the building. And it will do very little about the kitchen table at 10:05 at night, except remove the one adult who could have helped a student use the tool well.

The approach the research points toward is harder and more hopeful. Ask what the tool is set up to do. Give hints, not answers, during practice. Have students attempt the problem first, then use AI to check their reasoning, then explain in their own words what changed. Let them use AI to generate practice questions, quiz them, and find the step where they went wrong. Keep some assessments that ask students to show their thinking without the tool, not as punishment but because that is how you know the thinking is there.

The district does not need to choose between AI and learning. It needs to decide whose brain is doing the work during practice.

The New Hire Problem

This is not just a school question. It lands on every company that trains people.

Here is another one you may know. A regional insurance firm rolls out an AI assistant for new claims staff. Ask it about any policy and it answers instantly, with the right section number. New hires love it. Managers love that questions stop piling up in their inbox. After the first month, the new group's accuracy on routine claims is excellent.

Then a claim comes in that does not fit the pattern. Something about it is slightly off. A seasoned adjuster would feel it before she could explain it. The new hire does not feel anything, because he has never had to work out why a policy says what it says. He has always had the answer handed to him. He asks the assistant, gets a confident answer to the question he asked, and does not realize it was the wrong question.

That is the math study in a business suit. Strong practice performance. Thin understanding underneath.

The fix is not to take the assistant away. The new hires will need it for their whole careers. The fix is to design the first few weeks differently.

What AI Does Well Here

This is where I get genuinely excited, because designed well, AI may be the most patient tutor most people will ever have.

It never sighs. It will explain the same idea a fifth way at 11:05 at night without making anyone feel slow. It can meet a learner exactly where they are: simpler for the beginner, deeper for the person who is ready, in examples drawn from their own job or their own interests.

It can be set up to ask instead of tell. Try a request like, "I'm learning how to calculate discounts. Don't give me the answer. Give me a problem, wait for my attempt, then give me a hint about where I went wrong." Most tools will do this happily. That one sentence turns an answer machine into something much closer to the tutor that helped in the math study.

It can generate endless practice. Twenty new problems on the exact skill you are weak on, each slightly different, in seconds. Teachers and trainers used to spend hours building that.

It can play the part of a difficult customer, a skeptical client, or a confused patient, so a new hire can practice the conversation before it is real.

And it can find the step. When a learner shows their work and asks, "Where did my reasoning break?", AI is remarkably good at pointing to the exact line. That is feedback that used to require a teacher sitting beside you.

None of this removes the difficulty that matters. It moves the difficulty to the right place and adds support around it. Keep the tool. Build the thinker.

ASK, Then Go

Let's run the insurance firm's training plan through ASK, quickly.

A: Assumptions. What am I assuming? The firm was assuming that fast, correct answers during training meant learning was happening. The math study suggests those can come apart. A second assumption worth naming: that new hires will figure out on their own when to think and when to ask. Most will not, because the shortcut is always right there.

S: Source. Where does this come from? The confidence came from first month accuracy numbers, measured while the assistant was available. That is like measuring the students during practice. The better source is how new hires perform on a few unusual claims worked through without the assistant, followed by a conversation about how they reasoned.

K: Key. What is the key piece, and does it unlock what I need? For this firm, the key is whether a new hire can explain why a policy applies, not just which policy applies. So the redesign is small. For the first few weeks, new hires write their own answer and reasoning before asking the assistant, then compare. The assistant is set to give hints first during training. Once a week, a manager works through one odd claim with each person and asks, "Why?"

You can use the same move on the opening problem. You asked AI the \$360 question and got \$333. What is the key piece? Whether it applied the discount to the registration only. Check it: 20 percent off \$360 is \$288, and \$288 plus \$45 is \$333. The discount was applied to the registration, not the add-on, which matches the reading you chose. **Yes, this holds.** Thirty seconds, and you are done. You did not need to redo the whole problem. You needed to check the decision inside it.

If it holds, close the laptop. And if you are a school board or a training director about to change how hundreds of people learn, that is when you ask agAI: ask the teachers and trainers closest to the learners, ask the learners themselves what they actually do with the tool when no one is watching, and ask yourself whether you are choosing a ban or a free for all because it is easier to announce.

PRESSURE TEST: COPY THIS

Use this when the goal is to learn, not just to finish.

"Act as a tutor, not an answer key. Do not give me the answer. Give me one hint at a time, and ask me what I think the next step is before you say more."

BOTH THINGS CAN BE TRUE

AI can weaken learning. When it hands over the answer during practice, the effort that builds understanding can quietly disappear, and the gap shows up later when the tool is gone.

AI can strengthen learning. Designed to give hints, ask questions, and let the learner do the work, it may be one of the best tutors ever built.

The design decides. Banning the tool does not build the thinker, and neither does handing it over with no plan. The question is what the tool is set up to do and whose brain is doing the work.

The Problem You Will Remember

Go back to the two conference problems. A week from now, I suspect you will still know how to do the first one, the one you wrestled with, including that small moment where you decided what the discount applied to. The second one, the one that arrived finished, will probably be gone.

That is not a reason to stop asking AI to do math. I ask it to do math all the time, and I am glad to. It is a reason to notice when you are trying to get an answer and when you are trying to become someone who can find one. For the first, let the tool drive. For the second, let it ride along and give hints.

Nobody ever stopped listening to the radio because a station was banned. What changed us was the people who listened with us and asked what we thought.

THE KEY

Keep the tool and build the thinker: when the goal is learning, let AI give the hint and let your brain do the work.

CHAPTER 15

Your Value, Plus

TRY THIS FIRST

This one needs company. Gather your team, or two or three colleagues, or even a friend over coffee. Pick one real question that matters to your work right now, something like "What is the biggest risk to our top priority over the next three months?" Each person answers it alone, in silence, on paper, before anyone speaks. No peeking, no "what did you put?" Fold the papers. Then ask an AI tool the same question and read its answer out loud. Now unfold the papers and go around the table. What did each person see that the others did not? What did the AI see that nobody wrote down? What did someone write down that the AI never mentioned? What did you do?

If you did it, I suspect a few things happened.

First, somebody groaned at the "write it down alone" part. Probably more than one somebody. Writing before talking feels slow, a little like being back in school, and it removes the comfortable option of nodding along with whoever speaks first.

Second, the papers did not match. One person wrote about the budget. Another wrote about a key client. Someone wrote a name, the person who is quietly holding three projects together and has mentioned twice that she is tired.

Third, the AI answer was good. Organized, thorough, sensible. It probably named a risk or two that nobody at the table had written, and someone said, "Oh, that's a good one." And it almost certainly missed the thing about the tired colleague, because the AI has never been in your hallway.

Fourth, and this is the one I watch for: there was probably a moment, right as the AI's answer was read aloud, when the room got a little quieter. A few people glanced at their folded papers and suddenly found them less interesting. Maybe someone was tempted to cross out what they wrote.

That quiet moment matters more than it looks. It is where the value of human judgment either shows up or disappears.

That is the whole chapter in one exercise. Every chapter before this one has been about one person and one screen. This one is about the room, because most decisions that shape our work are made by teams, with an AI answer on the shared screen and a clock on the wall.

The Surprising Average

Let me start with a study that surprised me, and that I have chosen to take seriously instead of hiding in a footnote.

In 2024, Michelle Vaccaro, Abdullah Almaatouq, and Thomas Malone published a meta-analysis in *Nature Human Behaviour*. A meta-analysis is a study of studies. They gathered 106 experiments, with 370 results, that compared three things: humans working alone, AI working alone, and humans and AI working together.

You would expect the combination to win. Two heads, one of them made of silicon. That is the whole sales pitch.

On average, it did not. The human and AI combinations performed worse than the better of the two working alone.

Pause on that. Not worse than everyone. Worse than whichever one was stronger. When AI was the better performer on its own, adding a human tended to drag the result down. When the human was the better performer on its own, adding AI tended to help.

But the average hides the most useful part. On creative tasks, such as content creation, the combinations did better. On decision tasks, such as classification or forecasting, they did worse.

That pattern makes a lot of sense once you see it. On creative work, AI can widen the field of options, and the human can choose, shape, and add what matters. On decision work, one likely reason is that people can defer to the AI when it is wrong and override it when it is right. We have met pieces of that throughout this book: the radiologists in Chapter 1, the borrowed confidence in Chapter 12.

I find that both humbling and hopeful. Humbling, because "put a person in the loop" is not a magic phrase. A person in the loop who nods at everything is a rubber stamp with a salary. Hopeful, because it tells me where the work is. When we create with AI, the partnership already appears to pay off. When we decide with AI, it seems to need habits of judgment: a moment where somebody looks at the key piece before the whole team walks through the door.

One honest limit: a meta-analysis averages across very different experiments, and most of them were not leadership teams making real decisions about their own people. Averages across very different tasks can hide a lot. So I hold it as a signal, not a verdict. And notice what the study does not say. It does not say humans and AI cannot work well together. It suggests something more useful:

combining them does not automatically produce something better. The combination has to be designed. It gets better when we get better at it.

That is the argument of this whole book in one finding. The tool is powerful. The results depend on the judgment around it.

Team Human Versus Team AI

There is a rivalry in a lot of workplaces right now. On one side, people who think AI will take everything over and are secretly delighted or secretly terrified. On the other, people who think it is overhyped and are waiting, arms crossed, for it to embarrass itself. Team AI versus Team Human.

I have seen this movie before, and so has social psychology.

In 1954, Muzafer Sherif and his colleagues ran a now famous study with boys at a summer camp near Robbers Cave State Park in Oklahoma. The boys were split into two groups, who named themselves the Eagles and the Rattlers. Once the groups competed for prizes, the rivalry turned ugly fast: name calling, flag burning, raided cabins. Simply putting the two groups together for meals and movies did not fix it. Mostly it gave them a new place to throw food.

Competition had created the hostility, and contact alone did not undo it. What helped was what Sherif called superordinate goals: shared goals that neither group could achieve alone and that required both to cooperate. The camp's water supply "broke." A truck bringing food "got stuck," and it took both groups pulling on the same rope to move it. Working toward a shared goal that mattered to everyone, the rivalry softened.

I think about Robbers Cave every time I hear the conversation framed as humans versus AI. Will AI take our jobs? Is AI smarter than us? Whose side are you on?

Team Human versus Team AI is a false rivalry. And false rivalries make people think badly. Some people defend the human side by rejecting every AI output on principle. Others defend the AI side by treating human doubt as resistance to progress. Both are loyal to a team. Neither is loyal to the decision.

Nobody at work needs a winner. Everyone needs the shared goal above both teams. The superordinate goal is not "prove humans are better" or "prove AI is better." It is "make the best decision we can with everything available to us": one that serves the client, protects the people, and holds up next quarter. That goal needs both. It needs the AI's range and speed, and it needs the human's context, values, and judgment. Neither side wins it alone. Then the question

stops being "Who is smarter?" and becomes "What does each of us bring to getting this right?"

I am a lifelong Atlanta Falcons fan, so I know something about loyalty to a side. I wrote about football loyalty in *Stop, Look, and Listen*, and I want to reframe it here. On Sunday, loyalty to a side is part of the fun, and I will defend my team past the point of reason. That is what Sundays are for. At work, loyalty to a side can quietly cost you. The better loyalty is not to a side. It is to a good decision. If the AI has the better answer, I want it. If the person in the corner has the better answer, I want that. My jersey stays at home.

The Person Who Says "Wait, Let's Test That"

Here is one you may know.

A regional distribution company's leadership team meets every Tuesday at 4:05. The COO, Theo, has put an AI analysis on the shared screen. The team had uploaded a year of customer feedback and asked the tool what was driving customers to leave. The answer is clear and confident: delivery time is the top complaint, and the analysis recommends investing in a faster delivery tier this spring.

It reads well. It is 4:40, and everyone can feel the pull of the parking lot. Then Nora, the director of customer accounts, says, "Wait. Let's test that."

There is a tiny pause. In some teams that pause is cold, and the unspoken message is, "Really? Now?" In those teams, Nora learns quickly to keep quiet, and the team loses the habit that catches the key piece before the money is spent.

This is where Amy Edmondson's work on psychological safety matters. Edmondson, a professor at Harvard Business School, has spent years studying teams, and her core idea is simple to say and hard to build: psychological safety is a shared belief that the team is a safe place to take interpersonal risks. To ask a question. To admit a mistake. To say, "I'm not sure that's right." Her research in hospital units found something that sounds backward at first: some of the better teams reported more errors, not fewer. They were not making more mistakes. They felt safe enough to talk about them.

With AI on the screen, psychological safety takes on a new job. It has always protected the person who disagrees with the boss. Now it also needs to protect the person who disagrees with the machine, especially when the boss already likes what the machine said.

Think about how that feels. A leader shares an AI analysis. The leader is clearly pleased with it. The analysis is polished, confident, and nicely formatted. (You may remember the halo from Chapter 6.) Now a team member has to say, "I think this is missing something." That person is disagreeing with the leader, with the tool, and with the general mood of the room, all at once. Questioning a confident recommendation can feel like questioning the future itself, and nobody wants to be the person who "doesn't get AI."

If that person does not feel safe, they stay quiet, and the thirty second question goes unasked. The AI's blind spot becomes the team's blind spot.

There is a film I think every team should watch at least once: *12 Angry Men*, from 1957. Twelve jurors go into a hot room to decide a murder case. The first vote is eleven to one for guilty. The one is Juror 8. He does not claim the young man is innocent. He simply is not sure, and he does not believe a decision this serious should be made without talking about it first. As they talk, one juror after another begins to see what they had skipped.

It is a story, not a study, but it captures something the research supports. Remember Asch's line experiment from Chapter 7: when even one other person gave the right answer, going along with the group dropped sharply. The psychologist Charlan Nemeth has spent decades studying people who disagree with the majority. Her research suggests that dissent can improve a group's thinking even when the person disagreeing turns out to be wrong, because it pushes everyone to look at more information and consider more possibilities. And in her studies, genuine dissent tended to do more good than a devil's advocate assigned to argue a side. A real "wait" carries weight.

With AI in the room, the vote can feel like twelve to nothing before anyone speaks. The leader likes the answer, the tool sounds sure, and the room is ready to go. Sometimes the room is right. But someone has to be willing to be Juror 8 long enough to find out.

So one of the most useful things a leader can do is go first: "Here is what AI gave me. Before we go further, tell me what it got wrong." Then, when someone answers, thank them. Out loud. Every time.

In a team with real psychological safety, Theo's response to Nora is, "Good. Thank you. What do you want to check?"

That is the whole culture change in four words: the person who says "wait, let's test that" gets thanked.

The Checks People Like Least

I wish I could tell you that good checks feel good. They often do not.

In Chapter 12 we met Zana Bućinca, Maja Malin, and Krzysztof Gajos (2021), who studied ways to reduce overreliance on AI recommendations. Their cognitive forcing functions, such as making your own decision before you see the AI's recommendation, reduced overreliance more than simply giving people explanations of how the AI reached its answer. And here is the detail that matters for leaders, the part I underline every time: people rated the most effective designs least favorably.

In other words, the checks that worked best were the ones people liked least.

Of course they did. Forcing yourself to decide first is slower. It is a little uncomfortable. It exposes you, because now there is a record of what you thought before the machine weighed in.

This is where leaders come in. If you leave it to preference, people will usually choose the smooth path. Most of us would. The cognitive miser votes every time. So if the checks that work are the checks people like least, someone has to decide that those checks are part of how the team works. That someone is usually the leader. This is not about making work painful. It is about deciding, on purpose, where a little friction earns its place.

Remember the groan when you asked everyone to write their answer alone first? That groan is the sound of a good check. It asks the brain to commit to its own view before an elegant answer arrives and quietly becomes everyone's view.

Leaders, if you decide which processes to keep by how they feel, you may cut exactly the habits that protect your decisions. Some friction wastes an afternoon. Some costs ninety seconds and saves a quarter. Learn to tell them apart.

Judgment Readiness

Many organizations measure AI readiness by tools, licenses, training hours, and adoption rates. Those matter. But the Vaccaro findings point to another kind, the one that determines whether people and AI actually decide well together. I call it judgment readiness.

Judgment readiness is whether your people can question, check, and decide well when they are working with a tool that is powerful, persuasive, and

agreeable. You can buy every license on the market and still not be ready. You can have a modest set of tools and a team that is very ready indeed.

It shows up in a few quiet places:

- **What leaders reward.** Is the person who ships the AI draft fastest the hero? Or is it the person who shipped it fast and caught the one number that would have embarrassed us? If the fastest deliverable gets the praise, people learn to deliver fast. If the person who says "I checked the source and the number was wrong" gets thanked in public, people learn to check. Teams tend to do what gets noticed. So ask yourself honestly: in the last month, did I notice speed or did I notice judgment? Did anyone on my team get credit for catching an AI error, or for changing a recommendation after examining it?
- **Time and permission.** Do people have room, even a few minutes, to question an AI recommendation? A recommendation that arrives at 4:45 on a Thursday for a decision due Friday morning leaves no real space to ask what the AI assumed or where else to check. Then the process has already decided that nobody will question the answer. That is not a character flaw in the team. It is the design. Permission matters just as much: can a junior analyst say, "I do not think this AI summary is right," to a senior leader who already liked the summary? If the honest answer is "not really," your organization has a readiness gap no license can close.
- **Rituals.** Does the team have a small, repeated habit that makes examining AI's work normal, instead of something one brave person has to start from scratch each time? A small, repeated question at the end of a meeting can do more than a thick policy document that nobody reads.

That last one is where I would start. Rituals are how teams make the right thing the easy thing. Here is the one I like best. At the end of any decision where AI played a role, someone asks:

"What did AI help us see, and what did we decide after examining it?"

The first half gives credit for what the tool genuinely contributed, which is often a lot. The second half puts the decision back with the people in the room and asks them to say out loud that they examined it. It takes about a minute and works in a meeting, a debrief, or a message thread.

If the team cannot answer the second half, that is useful information. It may mean the AI's answer was adopted rather than examined. It may mean there was no time. It may mean no one felt safe enough to question it. Each of those is fixable, and each is a readiness issue, not a tool issue.

Some teams write the answer into their meeting notes. Some ask it only for major decisions. Some rotate who answers it so everyone practices. Use the version that fits your culture. The point is that the question gets asked, again and again, until asking it is simply how your team works.

Other rituals a team can borrow:

- **Write first, then ask.** For the big questions, everyone writes a view before the AI answer goes on the screen. Yes, people will groan. See above.
- **Name the key piece.** Before approving any AI informed recommendation, someone names the one fact it depends on, and someone else says where that fact came from.
- **Rotate the tester.** Each week, a different person is assigned to say "wait, let's test that" at least once. When it is a role, nobody has to be brave alone.

What AI Does Well Here

Let me be clear about what the tool brings to a team, because it is a lot.

AI is a tireless synthesizer. It can read a year of customer feedback in the time it takes your team to find a conference room, and sort it into themes that would have taken an analyst a week. It gives everyone at the table the same starting picture, which alone can save half a meeting.

It is a fair witness to the quiet voices. Ask it to summarize everyone's written input before discussion, and the person who rarely speaks gets the same font size as the person who always does.

It is a devil's advocate on demand. Ask for the strongest case against the plan the whole room already loves, and it will give you one without worrying about its next performance review. Ask it to run a pre-mortem ("Imagine this decision failed a year from now; what went wrong?") and it will hand you a list your team can use.

It asks questions you had not thought to ask. In the opening exercise, the AI likely named at least one risk no one wrote down. That is exactly the kind of contribution the Vaccaro findings suggest combinations can make well: widening the field before the team narrows it.

The best way to use AI in a team is the way you would use a brilliant new colleague who has read everything and met no one. Let it widen the view. Let

the people who know the history, the clients, and each other decide what to do with it.

ASK, Together

Back to Tuesday at 4:42. The team could spend another hour digging through the analysis line by line. Nobody wants that, and nobody needs it. ASK as a team habit is quick. It is not a meeting tax.

Assumptions. Theo asks the room, "What are we assuming?" Nora says, "That the feedback file represents the customers we're actually losing." Owen from operations adds, "And that faster delivery is something we can offer this spring." Two assumptions, named in under a minute.

Source. Where does "delivery time is the top complaint" come from? Nora opens the feedback file and filters for customers who left in the past year. She reads the first dozen comments out loud. Late deliveries. Missed windows. Late again.

The check comes back: yes, this holds. The AI was right. The top complaint among customers who left really is delivery time. Nobody needs to apologize for asking, and nobody needs to apologize for the AI either. Checking is not the same as distrusting. Sometimes checking is how you earn the right to move with confidence.

Key. What is the key piece, and does it unlock what we need? Here is where the people in the room do something no tool could. Owen says, "Our main carrier contract renews in March. If we're going to offer faster delivery, that's where we negotiate it. If we miss that window, we're locked in for two more years."

The AI could not have known that; the contract was not in the feedback file. But it changes the timing. The analysis was right about the problem, and the team now knows the key piece: get faster delivery terms into the March renewal.

That is the stopping point. The source held. The key piece is named and it unlocks the decision. Theo assigns Owen the contract conversation, and the meeting ends at 4:51, which in some companies counts as a miracle.

And then someone asks the ritual question. "What did AI help us see, and what did we decide after examining it?" Nora answers: "It helped us see the pattern in a year of feedback, fast. We checked it against the customers who left, it held, and we decided to move on faster delivery through the March contract."

That sentence is going in the meeting notes, and it took less time than it took you to read it.

If the decision had been bigger, say, restructuring the whole delivery operation and changing people's jobs, this is where the team would ask agAI. Ask the drivers. Ask the customers. Ask for the evidence a different way. But this decision did not need that. The key held. They went.

PRESSURE TEST: COPY THIS

Use this before a team decision, after everyone has thought on their own first.

"Here are our team's ideas. Tell me what we agree on, where we disagree, and what none of us mentioned. Then give me the one question we should answer before we decide."

BOTH THINGS CAN BE TRUE

AI can make a team faster, and teams can make AI's work better. On creative work, the combination already appears to shine. On decisions, it seems to need habits that turn a person in the loop into a person with judgment.

The best checks can feel slow, and they can still be the fastest path to a decision you will not regret. The groan when people write their answer first is not a sign the check is wrong. It is often a sign it is working.

Loyalty is wonderful, and it belongs to the decision, not the side. Cheer for the Falcons on Sunday. On Tuesday, cheer for whoever, or whatever, got it right.

Your Value, Plus

I want to end this chapter, and this part of the book, with a question many people are asking right now, often quietly, after the meeting, when no one else is listening.

"If AI can do all this, what am I for?"

AI widens your options. It drafts faster than you can type. It catches patterns across more data than you could read in a month. It argues the other side without taking it personally. It asks questions you had not thought to ask.

Your value does not shrink when AI shows up. It moves.

It moves toward asking the right question in the first place. Toward noticing what is missing, the way Owen noticed the contract and the way someone in your opening exercise wrote down the name of the tired colleague. Toward

knowing the people and the history, the client who always says yes in the meeting and no in the email, the project that failed three years ago for reasons that never made the report. Toward deciding what matters, which no amount of data can do for you. And toward owning the decision, standing behind it, and learning from it when it goes well and when it does not.

That is not your value minus AI.

That is your value, plus.

The "plus" only works if you keep your value in the equation. The moment in the opening exercise when someone was tempted to cross out their own answer is the moment the plus turns into a replacement. The Vaccaro meta-analysis suggests that is exactly where combinations tend to go wrong.

Go back to the folded papers from the start of this chapter. Every paper held something the others did not. The AI's answer held something, too. The picture came together when a room full of people looked at all of it, compared, questioned, and decided.

That is what a team ready for AI looks like. Not one that trusts the tool blindly or refuses it on principle. One that thanks the person who says "wait, let's test that," checks the key piece, and then moves, together and with confidence.

AI holds a lot of keys. More every year. It can open doors you did not know were there, and it can do it faster than any of us ever could alone. But a key does not decide which door is worth opening. You do. Your team does.

THE KEY

AI holds a lot of keys; you and your team decide which doors to open.

CONCLUSION

The Master Key

We started with applause.

You described an idea you loved, and AI clapped. You described the opposite idea, and AI clapped again. Maybe you laughed. Maybe you felt a little uneasy. Either way, you learned something that many people using AI every day have not stopped to notice: an answer can feel like a verdict without being one.

And somewhere in those first pages, a T-shirt was waiting to be ordered.

I have good news about the T-shirt. You can order it. I mean that. Some of your ideas are excellent, and AI is going to help you make them better, faster, and more often than you could alone. The point of this book was never to cancel the T-shirt. It was to make sure that when you order it, the slogan is something you would still be proud to wear in six months.

The Locks You Felt

Think back over what your brain did across these pages.

You picked up a can of air freshener and shook it, because of course you did, and the directions said not to. I still shake mine. That can taught us that autopilot is not a failure of intelligence. It is how intelligence saves energy. The trick is knowing when a moment deserves more than autopilot.

You read "Paris in the the spring" and lost a word, and you met a gorilla that walked through the middle of a basketball game while about half of the viewers counting passes never saw it. That gorilla taught us that what goes missing rarely announces itself, which is why a revised document deserves one simple question: what changed, and what is no longer there?

You looked at two tables that could not possibly be the same size, measured them, confirmed they were identical, and kept seeing them as different anyway. That one still makes me smile. Knowing about the illusion did not make it go away. What helped was a ruler. The original document. The actual math. The people closest to the work. A ruler does not argue with your eyes. It just tells you the size.

You tried numbers in the 2, 4, 6 puzzle, and if you were like most people, you tested sequences that fit your guess. Then someone tried a descending set like 6, 4, 2, heard "no," and suddenly the real rule came into view. One test

designed to prove you wrong taught you more than ten tests designed to prove you right.

And you stared at a painting that made no sense until someone flipped the canvas. Nothing on the canvas changed. Only the angle did. That is what a good second question can do, and it is one of the things AI does brilliantly when you invite it to.

None of that means you are careless. It means you are human. Your brain is built to be fast, efficient, and confident, and most of the time that serves you beautifully. Every one of those locks exists for a good reason. The trouble starts when a fast, efficient, confident brain meets a fast, efficient, confident machine, and neither one pauses.

The pause is yours. And so are the keys.

ASK, in One Breath

Here is the whole everyday method, in one breath:

A: Assumptions. What am I assuming? **S: Source.** Where does this come from? **K: Key.** What's the key piece, and does it unlock what I need?

That is it. Most days, ASK is enough.

I want to underline that, because I know some of you are overthinkers like me, and you may have closed several chapters thinking, "So I should check everything, always, forever." Please do not. That is not judgment. That is a new way to be stuck.

If the key piece holds, go. If the tone of the birthday message is right, send it. If the source for that number checks out, keep the number and move on. If it does not, take it out and move on. Close the laptop. Speed is part of good judgment. AI saved you an hour; a two minute check should not cost you that hour back.

And remember how often the check comes back with good news. Throughout this book, questioning AI frequently confirmed it was right. The quotes matched the contract. The numbers added up. The summary was faithful. That is not wasted effort. That is earned confidence, which is a much better kind than the borrowed kind.

When the Stakes Are High: One Possible Path

Some moments are different. A recommendation to restructure a team. A decision about who gets promoted, who gets let go, where the budget goes, or

what you tell a client about a problem. People's livelihoods are inside those answers. There is not one key piece. There are several.

This is when you ask agAI.

I promised you no long checklists, and I meant it. So what follows is not a rule. It is one possible path through a high-stakes moment, a set of questions I have found useful over the years. Think of it as a buffet, not a set menu. Most people will use two or three. Pick what fits the moment and leave the rest on the table.

- **What did my question assume?**
- **What kind of answer is this: an idea, a checkable fact, or a calculation?**
- **What is missing, and what did it echo back to me?**
- **What would change my mind?**
- **Where else should I check?**
- **What did I want to hear?**

Here is how a leader might use a few of them, in a story you may recognize.

Simone runs a division of about forty people. Her budget is tight, and she asks an AI tool to help her think through restructuring two teams into one. The answer comes back thoughtful and organized: merge the teams, eliminate a layer of management, consolidate reporting, and expect faster decisions and lower costs. It is persuasive. It is also, she notices, almost exactly what she had been leaning toward before she asked.

So she starts with **What did I want to hear?** She is honest with herself. She wanted permission. That does not make the recommendation wrong. It does mean the applause deserves a second look.

Then **What did my question assume?** She rereads her prompt. She had written, "How should I combine these two teams to reduce cost?" The question assumed combining was the answer and cost was the only goal. She opens a new chat and asks differently: "Here is the situation and the constraints. What are three or four different ways to address this, including options that do not involve merging the teams, and what does each one cost and risk?" The new answer includes the merge, which is reassuring. It also includes two options she had not considered, one of which keeps a team lead she would hate to lose.

Then **What kind of answer is this?** She sorts. The cost savings are a calculation, so she has her finance partner recalculate them with real salary and transition figures. Most of the math holds. One assumption about how quickly the merged team would be fully productive looks optimistic, and she

adjusts it. The claim that decisions will get faster is an idea, not a fact, so she judges it against what she knows about these people and this work.

And then **Where else should I check?** This is the one that matters most. She goes to the people closest to the work. She talks with both team leads, quietly and separately, and with a couple of long-tenured team members she trusts. They do not hand her an answer. They hand her something the AI could not: the history of a client relationship that depends on one team's specific way of working, and a handoff process that already breaks down every quarter and would break down worse in a merged team.

She did not use all six questions. She did not need to. She used four, and they took a few days, not a few months, because the stakes were high and the people deserved it.

Then she did the most important thing on the path.

She decided.

When you are done, decide, and own it. Do not keep asking until the anxiety goes away; it will not, and the asking will turn into avoidance. Asking agAI is meant to lead somewhere. Simone chose a modified version of one of the options she had not originally considered. AI helped her see more of the field. Her people helped her see what the field was standing on. The decision was hers, and she could explain every part of it without the screen in front of her.

That is the master key. Not more questions. The right few, then a decision you can stand behind.

Both Things Were True All Along

Throughout this book, I showed you research that pulled in more than one direction. AI raised people's performance and made them worse at judging how well they had done. Models caught real mistakes when they had something specific to check against, and stumbled when they were simply asked to try again. Several models debating improved accuracy, and several models sometimes shared the same blind spot. A well designed AI tutor helped students learn, and unrestricted answers helped them practice while leaving them less ready for the test.

I did not resolve those contradictions for you, and I am not going to start now. That is not a gap in the book. That is the book.

The world you lead in will keep handing you evidence that points two ways at once. The tools will keep getting more capable, more fluent, and more

persuasive. Some of what we know today about how they behave will have changed by the time you read this sentence. The three questions will not. What am I assuming? Where does this come from? What's the key piece, and does it unlock what I need? Those questions worked before AI arrived. They will work after the next model, and the one after that.

The Elephant, the Rider, and the Path

Chip Heath and Dan Heath, in their book *Switch*, borrow an image from the psychologist Jonathan Haidt: a Rider sitting on top of an Elephant. The Rider is our reasoning side. It plans, analyzes, and knows what it should do. The Elephant is our emotional, automatic side. It is enormous and strong, and it wants comfort, and it wants it now. The Rider may hold the reins, but when the two disagree, the Elephant usually wins.

If you have read this far, your Rider knows ASK. It knows to look at the assumptions, check the source, and find the key. Your Elephant, meanwhile, enjoys the applause, likes the polished answer, and would very much like to be done before lunch. That is not a flaw. It is being human.

The Heaths add a third piece: the Path, the situation around the Rider and the Elephant. Change gets easier, they argue, when you shape the path so the right move is also the easy move. For asking agAI, that might mean keeping your favorite pressure-test questions saved right where you write your prompts (Appendix C has a set to start with), building one quick check into your team's routine, or agreeing as a group that "Let's test that" is a welcome sentence, not a challenge.

Direct the Rider. Motivate the Elephant. Shape the Path. That is how asking agAI stops being a matter of willpower and becomes a habit.

The Cheer, One Last Time

When I was a little girl, we chanted, "Stop, Look, and Listen to what we have to say!" I still catch myself chanting it while getting ready in the morning or waiting for a Falcons game to start. Some habits are for life.

That cheer became a book about the power of the pause. And the pause turned out to be the doorway to this one. Stop, Look, and Listen became Ask agAI. The pause is where you notice the lock. ASK is how you choose the key. And agAI is what you reach for when the door really matters.

I find something hopeful in that. The skills that protect our judgment with AI are not brand new. They are the oldest skills in the room. Noticing what we

assumed. Asking where something came from. Finding the one thing that matters most and checking it. People were doing that long before anyone typed a prompt. AI did not make those skills obsolete. It made them valuable again.

The Promise

I want to leave you with the same balance we started with.

Catch what AI gets wrong. Use what it gets right. Move faster because you know the difference.

AI is not your replacement, and it is not your rival. It is a remarkable thought partner. It widens your options, drafts faster than you can type, catches patterns you would have missed, argues the other side without sulking, and asks questions you had not thought to ask. Your value does not shrink when it shows up. It moves toward the questions only you can ask, the people only you know, and the decisions only you can own.

Not your value minus AI. Your value, plus.

So use it. Explore it. Let it surprise you. Enjoy the applause when it comes; it is genuinely pleasant, and some of it will be deserved. Then ASK. Find the key piece. If it holds, go. If the stakes are high, ask agAI, decide, and own it.

A key does not break a door down. It opens it.

AI holds a lot of keys. You decide which doors to open.

THE KEY

Most days, ASK and go. When the door really matters, ask agAI, then decide.

APPENDIX A

ASK at a Glance

Keep this page nearby. Print it, photograph it, or tape it next to your screen. The prompts in quotation marks are ready to paste into any AI tool.

A: Assumptions

What am I assuming?

- Before you send a question, circle the words that carry a judgment (best, failing, resistant, obviously, why did). Try the question again with a neutral word.
- Paste this: "What in my question might be an assumption?"
- Paste this: "Answer this as if you did not know which option I prefer. What are the strongest arguments on each side?"

S: Source

Where does this come from?

- Sort the answer into ideas (judge them), checkable facts (find the source), and calculations (run the math).
- Paste this: "Quote the exact sentence from the document that supports each point."
- Paste this: "For every number in your answer, tell me where it comes from. If you are not sure, say so."

K: Key

What's the key piece, and does it unlock what I need?

- Find the one or two pieces everything else depends on. Check those, not everything.
- Paste this: "Which single claim does this recommendation depend on most?"
- Paste this: "If that claim turned out to be wrong, what would change in your recommendation?"

Match the depth to the stakes

- **A birthday message for a colleague.** The key is the tone. Read it. If it sounds like you and it sounds kind, send it.
- **A client proposal that says "improves productivity by 25 percent."** The key is that number. Source it. If there is a real source, keep it. If there is no source, take the number out and go.
- **A recommendation to restructure a team.** There are several key pieces, and people's livelihoods are involved. This is when you ask agAIIn.

Ask agAIIn: the master key

For high-stakes moments, choose the ways of asking agAIIn that fit. You will rarely need all five.

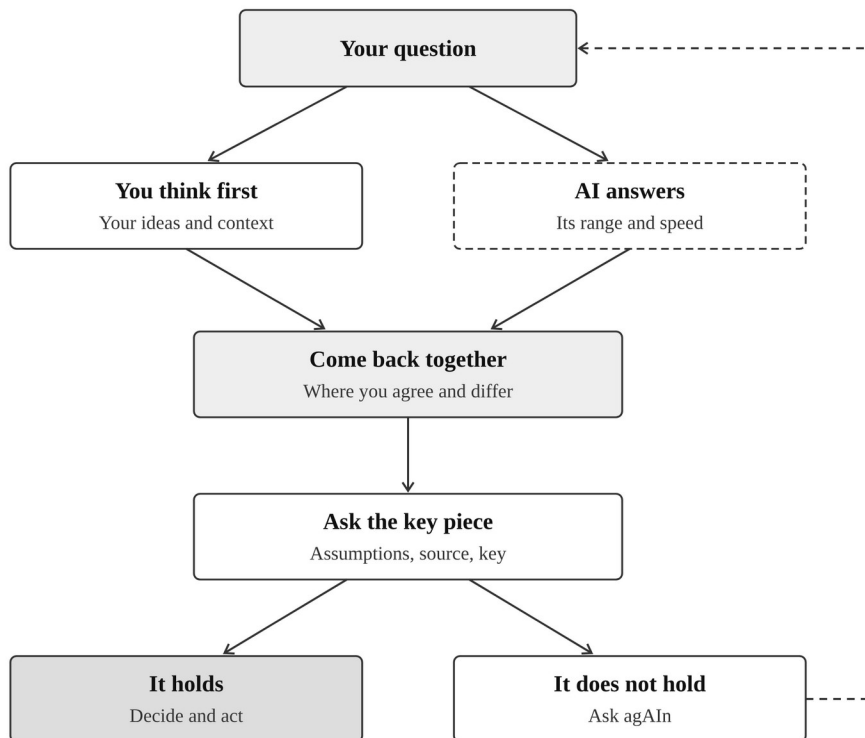
- **Ask differently.** Change the question, the angle, or the verb. Ask for options instead of approval.
- **Ask for the evidence.** What is this based on, and could I check it myself?
- **Ask somewhere else.** The original document. A second model, asked separately, without seeing the first answer.
- **Ask the people closest to the work.** They know the history, the exceptions, and what will actually happen on Monday.
- **Ask yourself.** What did I want to hear before I typed the question?

Asking agAIIn never means asking the same thing until you get the answer you like.

Questions to bring

- What am I taking for granted?
- What would a thoughtful person who disagrees with me notice?
- What information would change this recommendation?
- What could I test before investing more time or money?
- What else could explain this?

Go apart, come back together



When to stop

- If the key piece holds and it unlocks what you need, go.
- If a check comes back "yes, this holds," that is a good outcome, not wasted time.
- On a high-stakes decision, when you have asked agAI, decide, and own it.

APPENDIX B

The Judgment Readiness Reflection

Most AI readiness plans cover tools, data, security, and training. This reflection covers the part that rarely makes the roadmap: whether people have the habits, the time, and the permission to use their judgment with AI. There are no scores here and no right answers, only honest ones. Use the individual questions on your own, perhaps once a quarter. Use the team questions to open a conversation, not to grade anyone. If a question makes you pause, that pause is the point.

Habits

For me

- When an AI answer arrives that I like, what do I usually do next?
- How often do I notice what my question assumed before I send it?
- When I use an AI answer at work, could I say where the key piece came from?
- Which kinds of tasks do I check too little? Which do I check too much?
- When did a quick check last confirm that an AI answer was right? How did that feel?
- Do I know when to stop checking, or do I tend to keep going?

For our team

- What does our team usually do with an AI answer before it goes to a client, a leader, or a decision?
- Do we have a shared sense of which tasks are low stakes and which deserve asking agAI?
- When someone uses AI well, do they share how, so others can learn from it?

Permission and time

For me

- Do I feel I have permission to question an AI recommendation that a senior leader already likes?

- When I am busy, which check is the first one I skip?
- Do I give myself the few minutes a high-stakes answer deserves, or do I borrow that time from my judgment?

For our team

- When someone says "wait, let's test that," what happens next?
- Do our deadlines leave room to check the key piece on important work?
- Who on our team is most likely to ask the uncomfortable question? How do we treat them?
- Is it safe here to say, "I used AI on this, and I am not sure the number holds"?

What we reward

For me

- What gets praised in my work: speed, polish, accuracy, or good questions?
- Have I ever been rewarded for an answer that looked good but had not been checked?
- Have I ever thanked someone for catching something I missed?

For our team

- Do we celebrate the person who caught the problem as much as the person who shipped the work?
- Do we reward fast answers more than right answers? Do we reward right answers more than useful ones?
- If someone slowed a decision by a day and saved us from a costly mistake, would anyone notice?
- If someone moved quickly because they checked the key piece and it held, would we recognize that as good judgment too?

How we learn from misses and successes

For me

- When an AI-assisted piece of my work went wrong, what part was the tool, the prompt, the process, and the person?
- When an AI-assisted piece of my work went well, what did I do that made it work?

- What is one thing I now check that I did not check a year ago?

For our team

- When an AI-assisted decision goes wrong, do we look for someone to blame or for something to learn?
- When an AI-assisted decision goes well, do we look at it with the same care we give a mistake?
- Do we give credit accurately, to the tool, the prompt, the process, and the people?
- What did AI help us see, and what did we decide after examining it?

A closing question

- If we asked one question more often, starting this week, which one would make the biggest difference?

APPENDIX C

Pressure-Test Prompts

These prompts help you investigate. They cannot guarantee a correct answer. A tool can still agree too easily, miss something, or sound sure when it should not. For anything important, open the source yourself, and ask agAI.

Chapter 1: Autopilot

PRESSURE TEST: COPY THIS

Use this when the answer feels finished and you are ready to move on.

"Pause before you recommend anything. What assumptions in my question or in your answer could change the conclusion? Separate what you know from what you inferred."

Chapter 2: You Missed the Gorilla

PRESSURE TEST: COPY THIS

Use this when a revised draft comes back and you are about to approve it.

"Compare this version with the previous one. List every change you made, including small wording changes, anything you removed, and any number that is different. Do not summarize. List each change."

Chapter 3: The Tables Are the Same Size

PRESSURE TEST: COPY THIS

Use this when one claim is carrying the decision.

"Go get the research. Find the original source for the claim that matters most here and give me its title, date, and link. If you cannot verify it, say so. Do not invent a citation."

Chapter 4: What a Fool Believes, AI Agrees

PRESSURE TEST: COPY THIS

Use this when an AI answer agrees with you a little too nicely.

"Pressure test this plan as if it belonged to a competitor. What are the strongest arguments against it, what risks am I not seeing,

and what would need to be true for it to work? Do not soften your answer to be agreeable."

Chapter 5: Smashed or Hit?

PRESSURE TEST: COPY THIS

Use this when an AI summary is about to become the official record.

"List every statement in this summary that describes a decision. For each one, quote the exact words from my notes that support it. Mark anything that was discussed but not decided."

Chapter 6: The Halo Around the Answer

PRESSURE TEST: COPY THIS

Use this when an answer looks so polished that you want to trust it on sight.

"Strip the formatting and the confident tone. Rewrite your answer as a plain list of claims. Next to each claim, tell me whether it is a fact you can support, an estimate, or an opinion."

Chapter 7: Persuaded by a Machine

PRESSURE TEST: COPY THIS

Use this when one option suddenly sounds like the obvious choice.

"Make the strongest honest case for each option, including the one you did not recommend. Then tell me what evidence would actually settle the choice, separate from how persuasive each case sounds."

Chapter 8: Same Question, Different Answer

PRESSURE TEST: COPY THIS

Use this when you are about to act on an answer with facts or numbers in it.

"Sort your answer into three groups: ideas and suggestions, factual claims, and calculations. For each factual claim, tell me how I could verify it. For each calculation, show your work step by step."

Chapter 9: Second Opinions

PRESSURE TEST: COPY THIS

Use this when two tools, or two people, agree and you feel done.

"Two sources agree on this answer. Tell me whether they could be relying on the same underlying information. What is the original evidence, and where would I go to check it independently?"

Chapter 10: Flip the Canvas

PRESSURE TEST: COPY THIS

Use this after you have written your own ideas down first.

"Here are my ideas. Do not improve them yet. Give me five approaches that are completely different from mine, including at least one that borrows from a different field or industry."

Chapter 11: The Ladder of Inference at Machine Speed

PRESSURE TEST: COPY THIS

Use this when a summary hands you a conclusion.

"Walk me down the ladder. What is the raw data, what did you select from it, what did you assume, and where did the data turn into a conclusion? What other conclusions could the same data support?"

Chapter 12: Borrowed Confidence

PRESSURE TEST: COPY THIS

Use this before you present or defend an AI answer as your own.

"Ask me three questions a skeptical colleague would ask about this answer. Wait for my response to each one before you ask the next. Tell me where my explanation was thin."

Chapter 13: Who Gets the Blame?

PRESSURE TEST: COPY THIS

Use this when something went wrong, or right, and everyone wants a verdict.

"Before we assign blame or credit, help me separate what happened into four possible causes: the tool, the prompt, the process, and the person. What evidence would point to each one?"

Chapter 14: Don't Take Away the Radio

PRESSURE TEST: COPY THIS

Use this when the goal is to learn, not just to finish.

"Act as a tutor, not an answer key. Do not give me the answer. Give me one hint at a time, and ask me what I think the next step is before you say more."

Chapter 15: Your Value, Plus

PRESSURE TEST: COPY THIS

Use this before a team decision, after everyone has thought on their own first.

"Here are our team's ideas. Tell me what we agree on, where we disagree, and what none of us mentioned. Then give me the one question we should answer before we decide."

Notes

Chapter 7: Persuaded by a Machine

- Milgram's studies have been debated for decades, both for their ethics and for what exactly they show. Later researchers, including Gina Perry in *Behind the Shock Machine*, have questioned how some sessions were run and how participants understood the situation.

Chapter 14: Don't Take Away the Radio

- The study was described by George Land and Beth Jarman in *Breakpoint and Beyond* (1992) and was not published in a research journal, so it is best read as an illustration of the pattern rather than as proof. Robinson develops these ideas further in his books, including *Out of Our Minds*, *The Element*, and *Creative Schools*.

About the Author

Dr. Cherry Collier-Davis is a social and organizational psychologist, educator, and executive coach who has spent more than two decades helping leaders and teams think, decide, and work better together. She is the founder of Personality Matters, Inc., where she serves as Fractional Chief AI Enablement Officer, helping organizations bring AI into their work in ways that strengthen human judgment. She is known to many as The Whisperher®, for helping people hear what is not being said.

She holds a PhD from the University of Georgia, earned her BA magna cum laude from Spelman College, where she now serves as Executive in Residence, and is an ICF Master Certified Coach. She is the author of more than 22 books. She lives in Atlanta, where she still asks agAI, about everything.

Research and Further Reading

Before We Begin

- Thompson, J. (2000, June 18). I was certain, but I was wrong. *New York Times*.
- Thompson-Cannino, J., Cotton, R., and Torneo, E. (2009). *Picking Cotton*.
- Wixted, J., and Wells, G. (2017). The relationship between eyewitness confidence and identification accuracy. *Psychological Science in the Public Interest*.
- Wells, G., and Bradfield, A. (1998). "Good, you identified the suspect." *Journal of Applied Psychology*.
- National Academies of Sciences (2014). *Identifying the Culprit*.
- Pew Research Center (2025). How Americans view AI and its impact on people and society.
- KPMG and the University of Melbourne (2025). Trust, attitudes and use of AI: a global study.

The Introduction and Part One

- McKinsey and Company (2026). The state of AI.
- Loewenstein, G. (1994). The psychology of curiosity: a review and reinterpretation. *Psychological Bulletin*.
- Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*.
- Healy, A. F. (1976). Detection errors on the word "the." *Journal of Experimental Psychology: Human Perception and Performance*.
- Greenberg, S. N., and Koriati, A. (1991). The missing-letter effect for common function words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*.
- Roediger, H., and McDermott, K. (1995). Creating false memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.
- Asch, S. (1956). Studies of independence and conformity. *Psychological Monographs*.

- Milgram, S. (1963). Behavioral study of obedience. *Journal of Abnormal and Social Psychology*.
- Perry, G. (2013). *Behind the Shock Machine*.
- Lee, H.-P., and colleagues (2025). Survey of 319 knowledge workers on generative AI and critical thinking. CHI Conference on Human Factors in Computing Systems.
- Shaw, S., and Nave, G. (2026). Cognitive surrender and Tri-System Theory. Working paper, The Wharton School.
- Kahneman, D. (2011). *Thinking, Fast and Slow*.
- Langer, E., Blank, A., and Chanowitz, B. (1978). The copier study on mindlessness.
- Stroop, J. R. (1935). Interference in serial verbal reactions.
- Fiske, S., and Taylor, S. The cognitive miser.
- Kool, W., and colleagues (2010). Decision making and the avoidance of cognitive demand.
- Kruglanski, A., and Webster, D. (1996). Need for cognitive closure: seizing and freezing.
- Parasuraman, R., and Manzey, D. (2010). Complacency and bias in human use of automation.
- Dratsch, T., and colleagues (2023). Automation bias in mammography. *Radiology*.
- Simons, D., and Chabris, C. (1999). Gorillas in our midst.
- Drew, T., Võ, M., and Wolfe, J. (2013). The invisible gorilla strikes again.
- Simons, D., and Levin, D. (1998). Failure to detect changes to people during a real-world interaction.
- Johansson, P., and Hall, L. (2005). Choice blindness. *Science*.
- Pronin, E., Lin, D., and Ross, L. (2002). The bias blind spot.
- Shepard, R. (1990). *Mind Sights* (the Shepard tables).
- Adelson, E. (1995). The checker shadow illusion.
- Jastrow, J. (1899). The duck and rabbit figure. Hill, W. E. (1915). My wife and my mother-in-law.
- McGurk, H., and MacDonald, J. (1976). Hearing lips and seeing voices.
- Steyvers, M., and colleagues (2025). What large language models know and what people think they know. *Nature Machine Intelligence*.
- Wason, P. (1960). The 2-4-6 task.

- Cheng, M., and colleagues (2026). Sycophantic AI decreases prosocial intentions and promotes dependence. *Science*.
- Loftus, E., and Palmer, J. (1974). Reconstruction of automobile destruction.
- Loftus, E., Miller, D., and Burns, H. (1978). The misinformation effect.
- MIT Media Lab with Elizabeth Loftus (2024). Conversational AI and false memories. Preprint.
- Costello, T., Pennycook, G., and Rand, D. (2024). Durably reducing conspiracy beliefs through dialogues with AI. *Science*.
- Thorndike, E. (1920). A constant error in psychological ratings (the halo effect).
- Zajonc, R. (1968). Attitudinal effects of mere exposure.
- Hasher, L., Goldstein, D., and Toppino, T. (1977). Frequency and the conference of referential validity (the illusory truth effect).
- Kestin, G., and Miller, K. (2025). AI tutoring outperforms active learning. *Scientific Reports*.
- Cialdini, R. (2021). *Influence*, new and expanded edition.
- Salvi, F., and colleagues (2025). On the conversational persuasiveness of large language models. *Nature Human Behaviour*.

Part Two

- Farquhar, S., and colleagues (2024). Detecting hallucinations in large language models using semantic entropy. *Nature*.
- Du, Y., Li, S., Torralba, A., Tenenbaum, J., and Mordatch, I. (2024). Improving factuality and reasoning through multiagent debate. ICML.
- Kim, and colleagues (2025). Correlated errors in large language models. ICML.
- Lorenz, J., Rauhut, H., Schweitzer, F., and Helbing, D. (2011). How social influence can undermine the wisdom of crowd effect. *PNAS*.
- Madaan, A., and colleagues (2023). Self-Refine.
- Huang, J., and colleagues (2024). Large language models cannot self-correct reasoning yet. ICLR.
- Sharma, M., and colleagues (2023). Towards understanding sycophancy in language models.
- Panickssery, A., Bowman, S., and Feng, S. (2024). LLM evaluators recognize and favor their own generations. NeurIPS.

- Fazio, L., Brashier, N., Payne, B., and Marsh, E. (2015). Knowledge does not protect against illusory truth. *Journal of Experimental Psychology: General*.
- Popper, K. *The Logic of Scientific Discovery*.
- Taleb, N. N. (2007). *The Black Swan*.

Part Three

- Duncker, K. (1945). On problem solving (the candle problem).
- Glucksberg, S. (1962). The influence of strength of drive on functional fixedness.
- Maier, N. (1931). The two-string problem.
- German, T., and Defeyter, M. (2000). Immunity to functional fixedness in young children.
- Luchins, A. (1942). Mechanization in problem solving (the Einstellung effect).
- Bilalić, M., and colleagues (2008). Why good thoughts block better ones.
- Jansson, D., and Smith, S. (1991). Design fixation.
- Wadinambiarachchi, S., and colleagues (2024). The effects of generative AI on design fixation. CHI.
- Thompson, P. (1980). The Thatcher illusion.
- McCaffrey, T., and Pearson, J. (2015). Find innovation where you least expect it. *Harvard Business Review*.
- Guilford, J. P. (1967). *The Nature of Human Intelligence*.
- Argyris, C. The ladder of inference. Senge, P. (1990). *The Fifth Discipline*.
- Rozenblit, L., and Keil, F. (2002). The illusion of explanatory depth.
- Fernandes, D., and colleagues (2025). AI makes you smarter but none the wiser. *Computers in Human Behavior*.
- Dell'Acqua, F., and colleagues (2023). Navigating the jagged technological frontier. Harvard Business School and BCG.
- Bućinca, Z., Malin, M., and Gajos, K. (2021). To trust or to think: cognitive forcing functions.
- Heider, F., and Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology*.
- Kelley, H. (1973). The processes of causal attribution.

- Plato. *Apology*.
- Bastani, H., and colleagues (2025). Generative AI without guardrails can harm learning. *PNAS*.
- Risko, E., and Gilbert, S. (2016). Cognitive offloading.
- Bjork, R. Desirable difficulties.
- Robinson, K. (2010). Changing education paradigms. RSA Animate.
- Land, G., and Jarman, B. (1992). *Breakpoint and Beyond*.
- Vaccaro, M., Almaatoug, A., and Malone, T. (2024). When combinations of humans and AI are useful. *Nature Human Behaviour*.
- Edmondson, A. Psychological safety in teams.
- Nemeth, C., Brown, K., and Rogers, J. (2001). Devil's advocate versus authentic dissent. *European Journal of Social Psychology*.
- Heath, C., and Heath, D. (2010). *Switch*.
- Sherif, M. (1954). The Robbers Cave experiment and superordinate goals.
- Collier-Davis, C. (2019). *Stop, Look, and Listen*.