

Data Literacy for Human Capital Management

Aug 10-11, 2026

Kevin Hanegan

Turning Data into Wisdom

kevin@turningdataintowisdom.com

Class Resources

<https://www.turningdataintowisdom.com/navair>

COURSE RESOURCES

Data Literacy for Workforce Decision-Making

Everything from the two days, in one place. The workbook and the cards are yours to keep and use at your desk. Come back to this page whenever you need them.



The Slides

The full two-day deck as a PDF, in the order we ran it. Useful for retracing a module or borrowing a framework for your own work.

PDF Download



The Reference Cards

All nine cards in one booklet: UIE, the analytical moves, reading red flags, which chart, the shared language, and both canvases. This is the one to keep at your desk.

PDF Download



The Participant Workbook

Your own copy, including the sections we did not get to in the room: the troubleshooting index for when you are stuck at work, and a brief on all ten modules.

WORD Download



The Team Workspace

The tabbed workspace we used for every activity, with the case exhibits built in. Open it any time to rerun an exercise on a decision of your own.

OPEN IN BROWSER Launch



Resources For This Class

Monday August 10

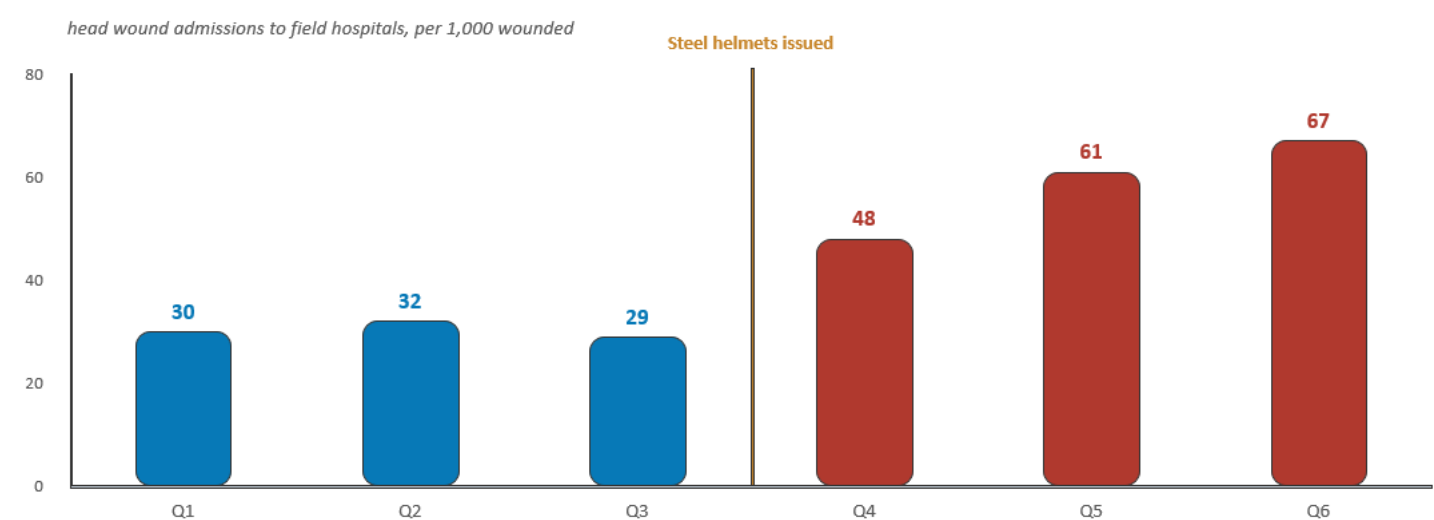
Time	Session	Mins
08:00	Opening and Course Foundations	30
08:45	Module 1 · Data Literacy Today	45
09:30	<i>Break</i>	<i>15</i>
09:45	Module 2 · Frame Better Questions	80
11:05	Module 3 · Data as Evidence	40
11:45	<i>Lunch</i>	<i>60</i>
12:45	Module 3 · Activity and debrief	25
13:10	Module 4 · Reading Data Correctly	65
14:15	<i>Break</i>	<i>15</i>
14:30	Module 5 · Exploring the Data	65
15:35	Day One reflection	25
16:00	<i>End of Day One</i>	

Tuesday August 11

Time	Session	Mins
08:00	Day One recap	15
08:15	Module 6 · From Pattern to Cause	75
09:30	<i>Break</i>	<i>15</i>
09:45	Module 7 · Visual Thinking	60
10:45	Module 8 · Designing Better Visualizations	65
11:50	<i>Lunch</i>	<i>60</i>
12:50	Module 9 · Driving Decisions	75
14:05	Module 10 · Telling the Story	40
14:45	<i>Break</i>	<i>15</i>
15:00	Capstone	40
15:40	Course Wrap Up	20
16:00	<i>End of Day Two</i>	

Question # 1 - A Decision From 1915

At the beginning of World War I, the uniform of the British soldiers included a brown cloth cap. As the war went on, the British officials became more and more concerned about the growing number of soldiers suffering from head injuries. They decided to replace the cloth cap with a metal helmet. However, to their surprise, the number of soldiers with head injuries grew even more, although the intensity of the battles was the same before and after the change.

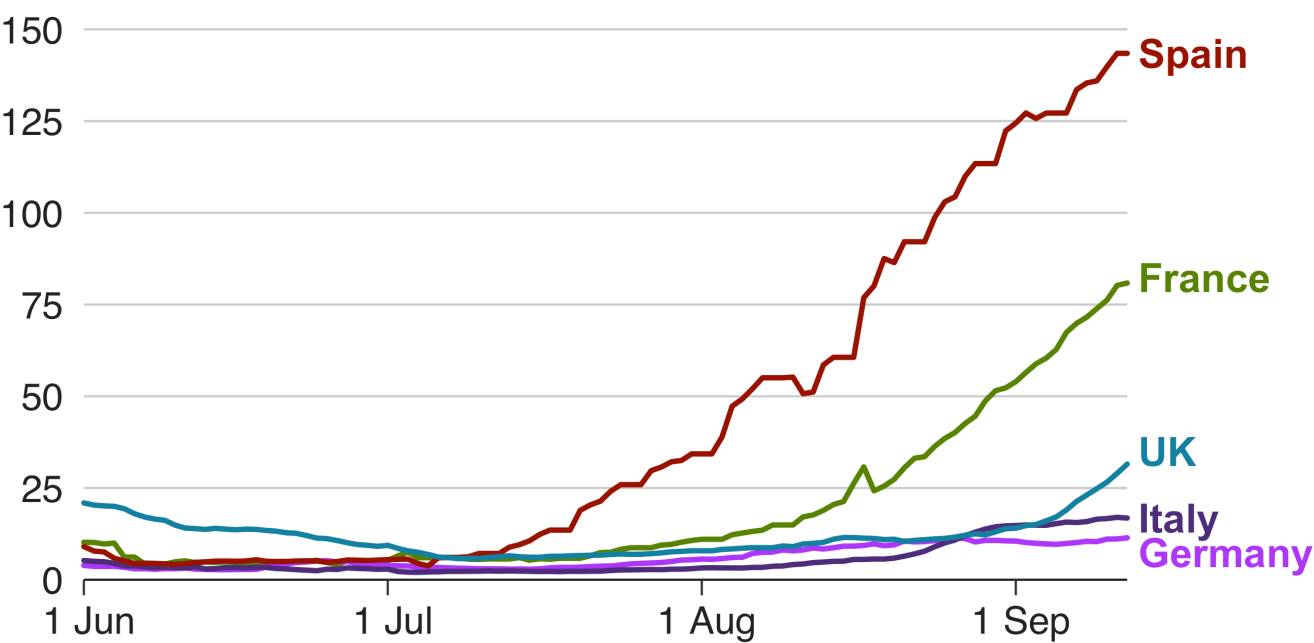


Were metal helmets hurting soldiers? Do we go back to cloth caps?

Question #2 - A Chart From 2020

Coronavirus cases increasing in European countries in recent weeks

Total cases per 100,000 people by week up to 12 September



Note: Countries do not always release figures every day, which may explain some of the sharp changes in the trendlines

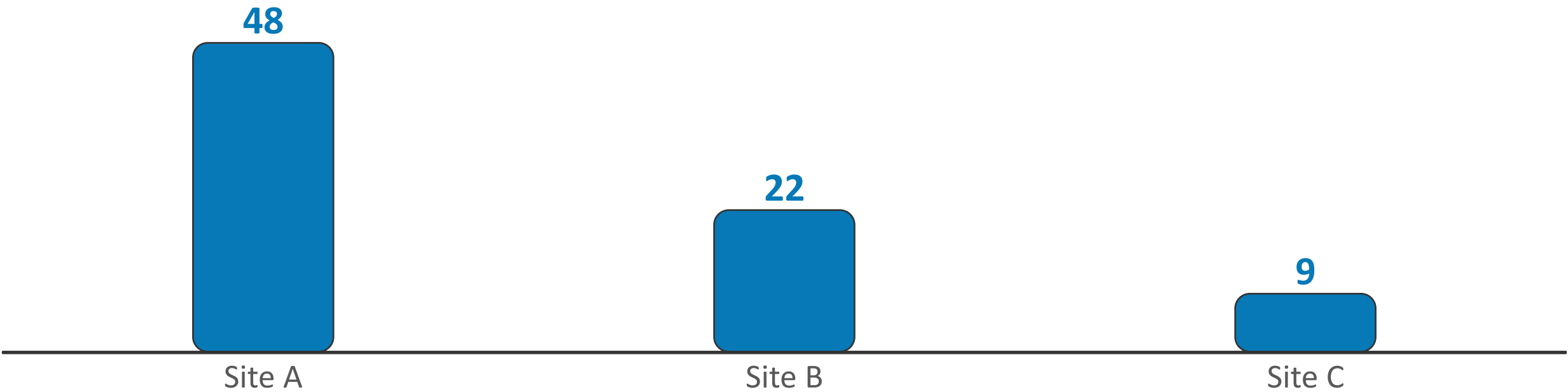
Source: ECDC, data to 12 Sep

BBC

How much worse did the outbreak get over those two months?

Question #3 - Complaints Filed by Site

Which site has the culture problem?



Question #4 - Internal Candidates vs. External Hires

Last year, across two departments

Overall Hire Rate

Internal applicants: **33%**

500 internal applicants, 165 hired

External applicants: **52%**

500 external applicants, 260 hired

An internal candidate is far less likely to get the job than someone from outside

A finding like this usually ends with: “we are overlooking our own people. Fix the internal mobility process.”

Are we passing over internal candidates?

Question # 1 - A Decision From 1915

The number was true, the conclusion was wrong

What The Data Counted

Soldiers who arrived at a field hospital with a head wound

Who Was Never In The Data

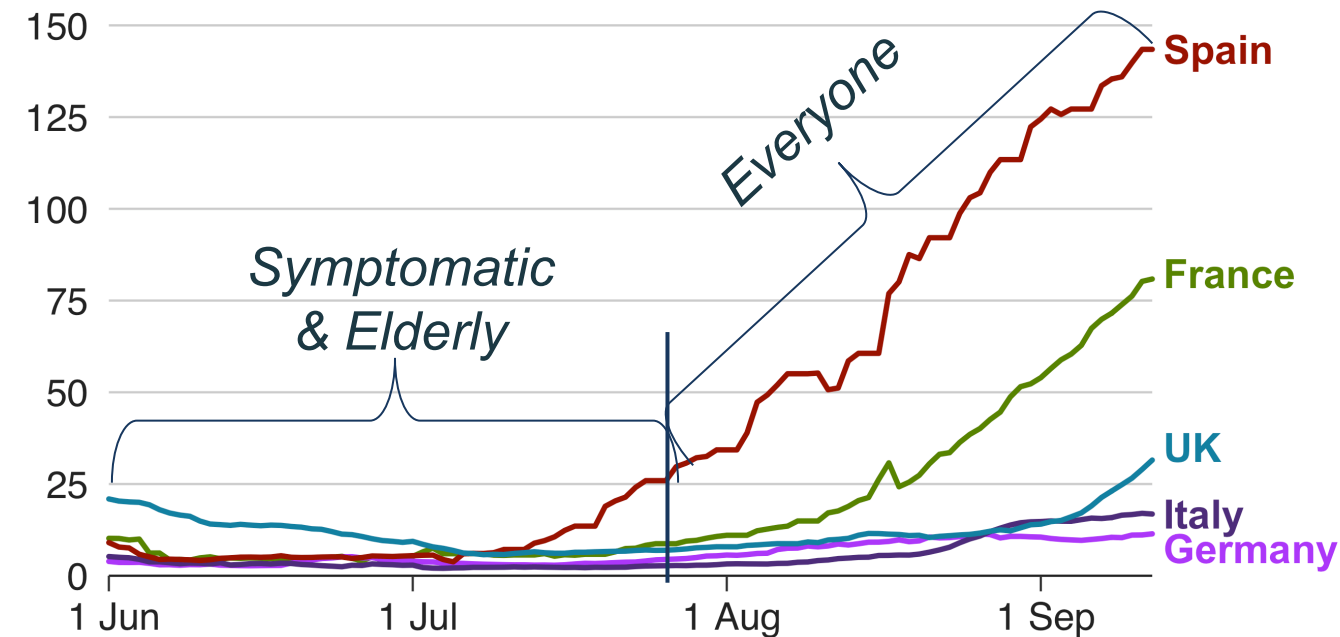
Soldiers who died of head wounds before reaching a hospital

The helmets were working. Soldiers who would have been counted as dead were now counted as wounded.

Question #2 - The Ruler Changed Underneath the Line

Coronavirus cases increasing in European countries in recent weeks

Total cases per 100,000 people by week up to 12 September



Note: Countries do not always release figures every day, which may explain some of the sharp changes in the trendlines

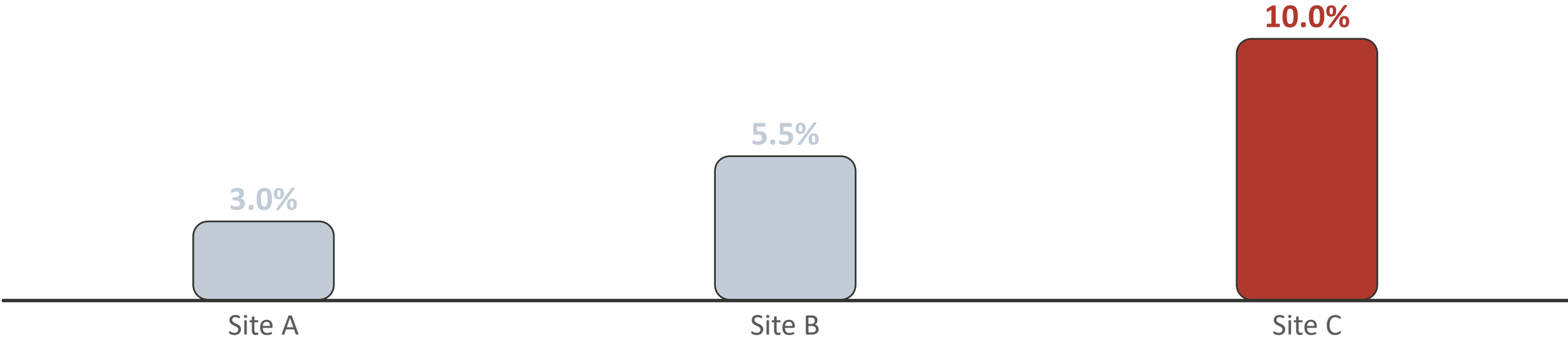
Source: ECDC, data to 12 Sep

BBC

Part of the rise is more looking, not more disease
A case count cannot tell you prevalence when the testing regime moves underneath it

Question #3 - The Same Complaints Per 100 Staff

Headcount: Site A - 1,600 Site B - 400 Site C - 90



Site C has the worst rate by far. Site A has the best. It is also the smallest site, so it never looked like a problem.

Question #4 - Split by Department, It Reverses

Internal candidates do better in every department, and worse overall

Department	Group	Applied	Hired	Rate
Technical roles	Internal	400 applied	100 hired	25%
	External	100 applied	20 hired	20%
Administrative	Internal	100 applied	65 hired	65%
	External	400 applied	240 hired	60%

Internal candidates are hired at a HIGHER rate in both departments. 25% vs 20%, and 65% vs 60%.

The gap is where they applied. Internal candidates concentrated in the competitive technical roles.
The mobility process was not the problem.

The Number Was True In All 4 Scenarios

The missing people

The data described who showed up, not who was absent

The moving ruler

When the counting changes, a line can rise without the world changing

The missing denominator

A count tracks size. A rate tracks the problem

The missing split

An overall number can reverse when you look inside it

Not one of these was a tool problem. Better software would have produced the same wrong answer, faster

One of them was not even a number. It was a visualization, and it still could not answer the question we asked of it.

What These Two Days Are, And Are Not

We will not teach software tools. We will teach the judgment that decides what to build in them, and what the output means. This includes framing better questions, correlation versus causation, getting to root cause, communicating insights and findings, and more.

This Course Is About One Habit

Don't get ahead of the evidence

Start with the question, let the evidence earn the answer

The Journey

QUESTION

EVIDENCE

READ

EXPLORE

EXPLAIN

VISUALIZE

RECOMMEND

TELL

Here Is What That Looks Like

Stage	The Jump We Refuse To Make
Question	<i>“We already know what to analyze”</i>
Evidence	<i>“The data is the truth”</i>
Read	<i>“The number means what it appears to mean”</i>
Explore	<i>“The first pattern is the important one”</i>
Explain	<i>“The pattern proves the cause”</i>
Visualize	<i>“A clearer chart is automatically an honest chart”</i>
Recommend	<i>“A finding tells us what to do”</i>
Tell	<i>“A compelling story is necessarily a truthful one”</i>

What You'll Be Able To Do

By the end of these two days, you will be able to:

	Ask Better Question	<i>Turn broad requests and assumptions into questions you can actually investigate</i>
	Judge the Evidence	<i>Understand what data can tell you, what it can't, and where you should be cautious</i>
	Read Data Correctly	<i>Interpret numbers in context and recognize when comparisons, averages, or other summaries can mislead</i>
	Find and Explain Patterns	<i>Explore data without confusing an interesting pattern with an explanation</i>
	Communicate What Matters	<i>Use visualization and storytelling to make evidence understandable without distorting it</i>
	Make Better Recommendations	<i>Move from evidence to action while keeping judgment, uncertainty, and context visible</i>

Evidence strength ≠ Confidence ≠ Action

How We Talk About Evidence, Confidence, and Action - One shared vocabulary

Evidence strength

How much the evidence supports the claim



Confidence

How sure we are, given the evidence



Action

What confidence and consequence justify



Strong evidence does not automatically mean “act.” Weak evidence does not automatically mean “wait.”

Should I Bring an Umbrella?

Three different questions, and only the third one is a decision.

Evidence strength

what the information supports

The app says 70% chance of rain. Dark clouds outside. Two forecasts agree

Moderate to strong

Confidence

how sure you are, given the evidence

Fairly high, but forecasts are forecasts and conditions change

Moderate to high

Action · the evidence and the confidence are now fixed. Only the consequence changes

Walking twenty minutes to an outdoor event

Being wrong costs you a soaking. Carrying it costs almost nothing.

Bring it

Driving straight into a parking garage

Being wrong costs you nothing. Carrying a golf umbrella all day costs you something.

Leave it

Same evidence. Same confidence. Different action. The evidence tells you what to believe. The consequence tells you what to do.

Shared Language Reference Card



When To Use It

Any time you say how strong the evidence is, or how sure you are

What’s On It

Three scales: evidence strength, confidence, action

Module 1

Why Data Literacy

Understanding What It Really Means and Why It Matters

What is Data Literacy?

The ability to **read, work with, argue** with data

The ability to **read, understand, create, and communicate** data as **information**

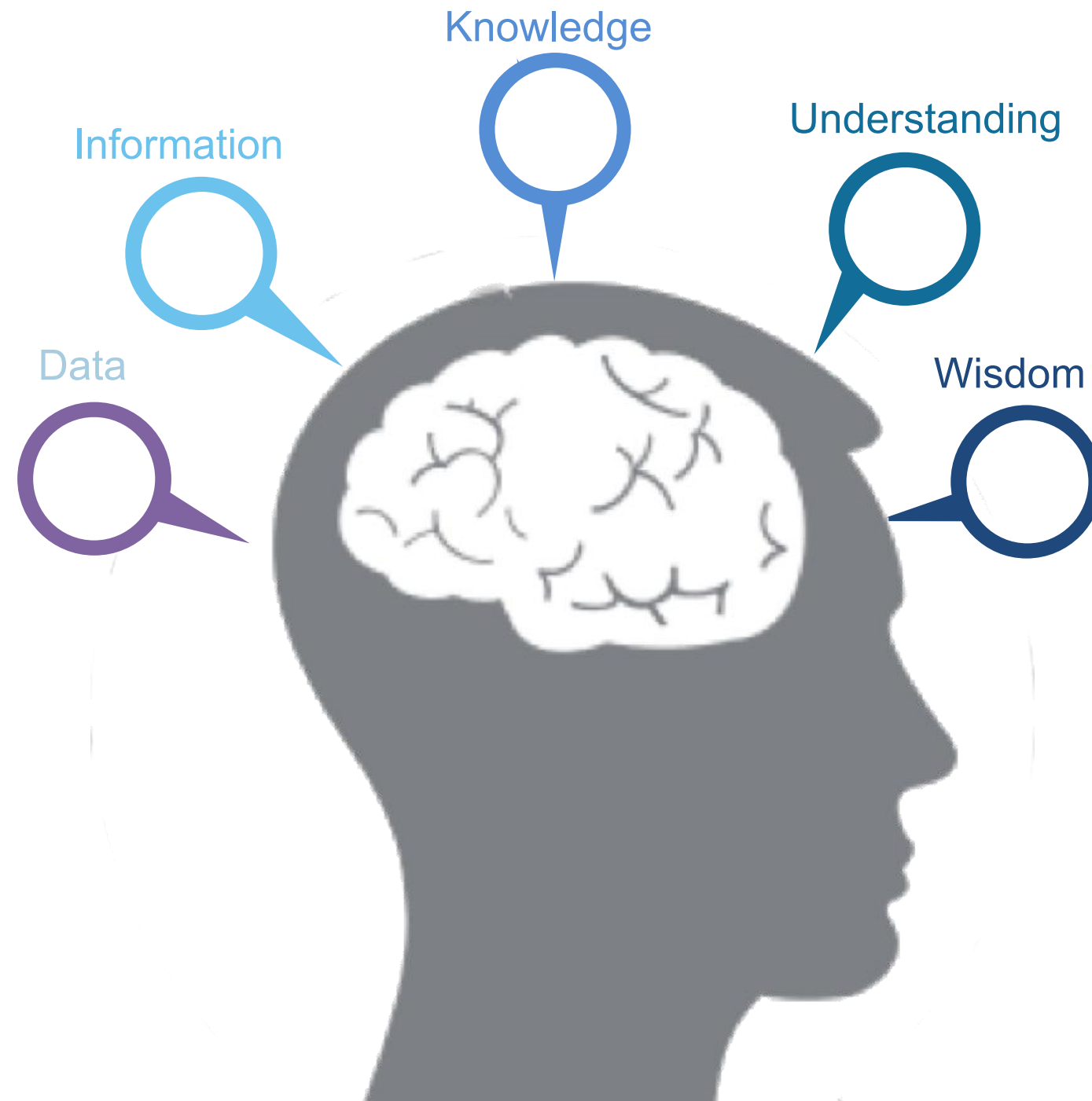
The ability to **formulate and answer questions using data** as part of evidence-based thinking; **use appropriate data, tools, and representations** to support this thinking; **interpret information** from data; **develop and evaluate data-based inferences and explanations**; and **use data to solve real problems and communicate their solutions**

The ability to understand and use data effectively to inform decisions

The ability to **collect, manage, evaluate, and apply data** in critical manner

The set of abilities around the use of data as part of **everyday thinking and reasoning for solving real-world problems**

What Is Data?



Data – Raw facts and figures in various forms, including numbers, text, images, etc... That are collected through observation, measurement, and sensory input. *Raw facts and observations*

Information – Organized or summarized data so that we can see what happened. *Processed data*

Knowledge – Patterns, relationships, or insights we identify from the information and its context. *Contextualizing insights*

Understanding – Interpret what the patterns mean in the context of the question, using evidence, context, and relevant experience. *Comprehending and interpreting insights*

Wisdom – Evaluated understanding. You applied the decision from the knowledge and reflected upon it. *Applied and evaluated understanding.*

What Is Data?

Making the decision isn't wisdom. We don't get to call it wisdom just because we acted. Wisdom comes from applying what we understood, seeing what happened, and carrying that learning into the next decision.

WHAT IS MY QUESTION?

Are we hiring in a way that gets us the right people before we lose them to someone else?

Knowledge

- Roles with longer interview-to-offer gaps also have higher decline rates
- Panel debriefs account for most of the delay

Information

- 32% of offers were declined last quarter, up from 21%
- Average time from final interview to offer rose to 11 days, from 6

Understanding

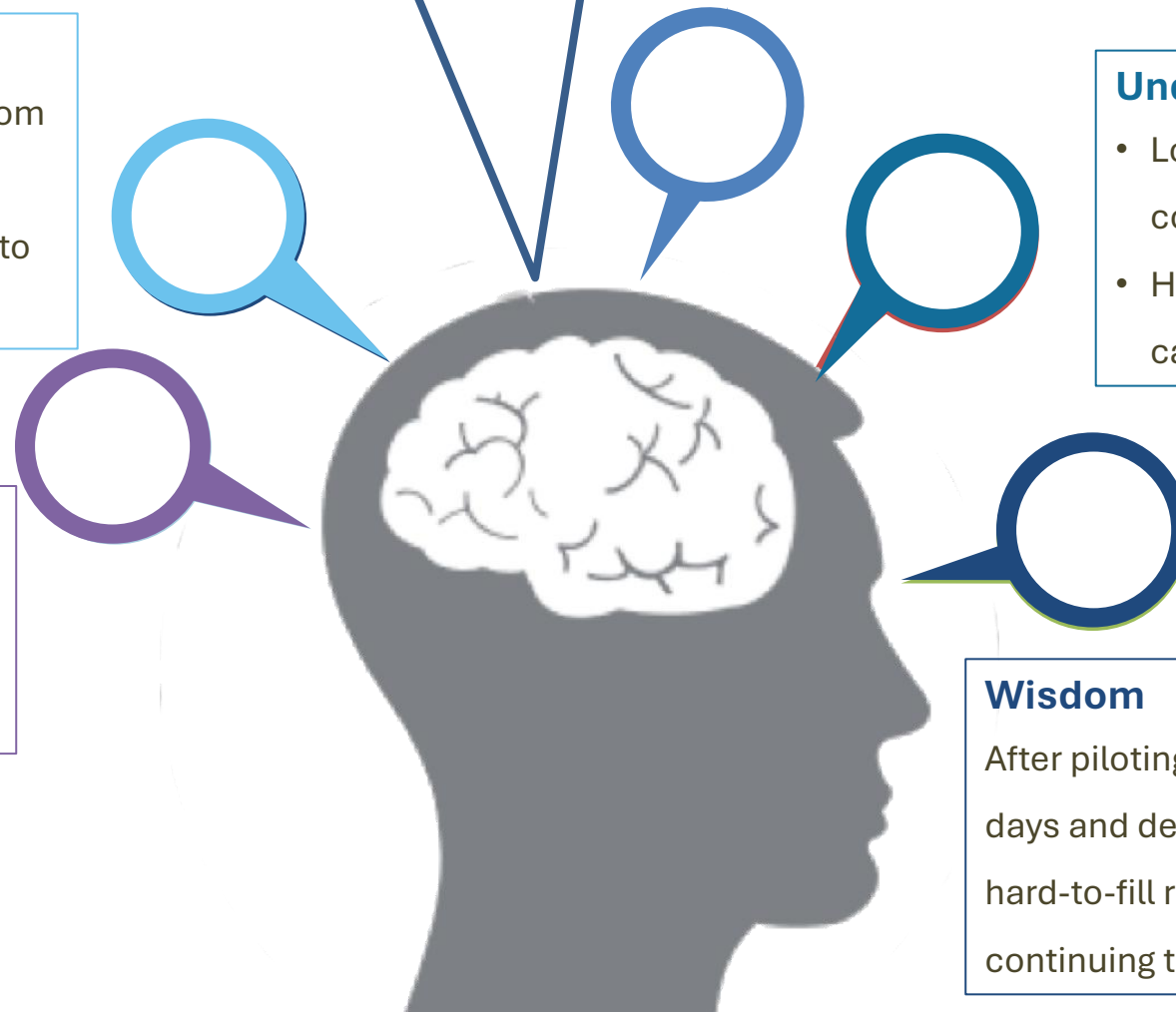
- Longer delays may give candidates more time to accept competing offers
- Hard-to-fill roles appear especially vulnerable because candidates are more likely to have alternatives

Data

- Applicant tracking records
- Interview and debrief scheduling logs
- Offer dates, and accept or decline decisions

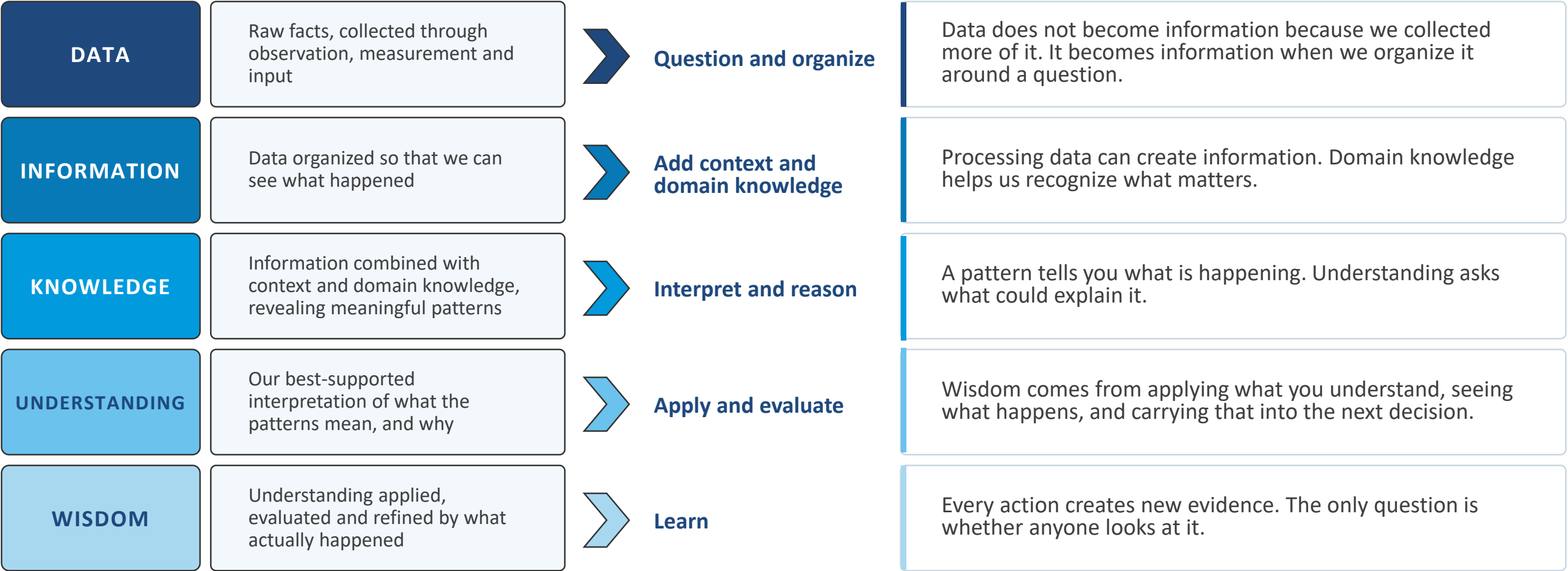
Wisdom

After piloting five-day decision windows, offer time fell from 11 to 7 days and decline rates decreased, with the largest improvement in hard-to-fill roles. We therefore expanded the practice while continuing to monitor results.



The Value Is in the Connections

Each step is a state and each move between them is a judgment



↻ What you do next produces NEW DATA, and the loop starts again

Technology can help at every step. The moves between the steps are where judgment lives. The value is not in climbing the ladder, it is in knowing what it takes to move up one rung.

What is Data? Training Completion and Readiness

Statement	Where It Fits	How to Move It Up?
450 sailors did not complete mandatory cyber readiness training last year	Data	Organize it around the question: How many were eligible? What percentage completed? How does that compare with the target, prior years, or units?
Completion rate = 82% Below the Navy target of 90%	Information	Add context and domain knowledge: Where is non-completion concentrated? What do we know about deployment schedules, access, role requirements, or other conditions that might matter?
Survey responses and operational context show non-completion is concentrated during deployment periods, when sailors report limited time and access to training.	Knowledge	Interpret and reason: Does the broader evidence support time during deployment as an important explanation? What other explanations could account for the pattern?
Deployment demands and limited training access appear to be contributing to lower completion, particularly in units where sailors have fewer opportunities to complete training during normal workflows.	Understanding	Apply and evaluate: Take a proportionate action based on this understanding and measure what happens
After piloting just-in-time micro-learning during two deployment cycles, completion increased from 82% to 91%, with the largest gains in units previously reporting scheduling barriers. The approach was expanded, with continued monitoring.	Wisdom	Learn and generate new evidence: What did we learn from implementation? Where did it work or not work? What should we change next?

A Dashboard Is Not a Decision

Reporting shows what happened. Deciding takes judgment about what it means and what to do next

What The Dashboard Shows

TURNOVER

14%



TIME TO FILL

62 days



ENGAGEMENT

3.8 / 5



OPEN ROLES

47



The judgment gap



What Leadership Still Has To Work Through

Context

Is 14% high? Compared to what?

Knowledge

Which roles and units are driving it?

Understanding

What might explain the pattern?

Apply + Evaluate

What should we try, and how will we know if it worked?

The metric is where the conversation starts, not where it ends

What Happens When We Make A Decision Before The Evidence Has Earned It

Find the miss in each

1

The attrition alarm

One unit logged 12 resignations last quarter, the most of any unit

Decision: Leadership shifted retention bonuses to that unit

2

The engagement dip

Engagement slipped from 3.9 to 3.7 year over year

Decision: A wellness pilot was cut as not working

3

The training win

A command reached 95% mandatory training completion

Decision: The command reported that readiness had improved

4

The exit interview

Exit interviews named better pay elsewhere as the top reason for leaving

Decision: A raise package was proposed to slow attrition

Eight Failures

Notice where they cluster



all eight broke in here

The Four We Started Today With

The Helmets

EVIDENCE

The data described who reached a hospital

The Coronavirus Chart

READ

The measurement moved underneath the line

Complaints by Site

READ

A count where a rate belonged

Internal vs. External hiring

EXPLORE

Nobody split it by department

The Four Decisions

The Attrition Alarm

READ

Twelve resignations, and no headcount to divide by

The Engagement Dip

READ

0.2 of a point, with no sense of what is normal

The Training Win

EVIDENCE

Completion is activity. Readiness is the outcome

The Exit Interview

EXPLORE

A stated reason taken as the cause

So, What Is Data Literacy?

Data literacy is the ability to **use data as evidence**, **combine it with context and domain knowledge**, and **apply judgment** to make better-informed decisions

Use data as evidence

The four questions this morning. Every number was true, and every conclusion was wrong. A number is something you weigh, not something you obey

Combine it with context and domain knowledge

The ladder. Data does not become information because you collected more of it. What you know about the work is what tells you which of it matters

Apply judgment

The dashboard. Software climbs the first two rungs and stops. Eight failures today, and every one sat in a move a person had to make

Informed, not driven. The data does not make the decision. It tells you what you have reason to believe, and you are still accountable for the rest.

Module 1 - What We've Earned

You can see that judgment, not tools, is what makes data useful

What's Next

Learning how to frame the question

THE JOURNEY SO FAR

QUESTION

EVIDENCE

READ

EXPLORE

EXPLAIN

VISUALIZE

RECOMMEND

TELL

Module 2

Frame Better Questions

Judgment before analysis

A Clear Request Can Hide the Wrong Question

This lands in your inbox one morning. It is clear, reasonable, and specific

“We are building the leadership pipeline for the next three years. Can you analyze performance ratings and give us a shortlist of our high-potential people?”

What would you pull?

How would you rank?

What would you deliver?

So what are you actually about to answer?

A Clear Request Can Hide the Wrong Question

This lands in your inbox one morning. It is clear, reasonable, and specific

“We are building the leadership pipeline for the next three years. Can you analyze performance ratings and give us a shortlist of our high-potential people?”

What would you pull?

Ratings, tenure, role, maybe promotion history

How would you rank?

Top decile? Two years of consistent scores? A weighting?

What would you deliver?

A shortlist of names, on their desk Friday

So what are you actually about to answer?

You Would Have Answered It Perfectly

And still answer the wrong question

What The Request Led You To Answer

Who are our highest performers today?

A ranking. Backward-looking, and easy to produce

What They Actually Had To Decide

Who should we invest in for roles that do not exist yet?

A bet on the future. Different question entirely

Why The Perfect Answer Is The Wrong Answer

- 1

Performance is not potential

A high rating says someone is excellent in the job they hold now, under their current manager, with today's demands. Leadership roles ask for different things
- 2

You would be ranking managers as much as people

Ratings carry the generosity of whoever wrote them. Part of that shortlist is just people with a lenient boss
- 3

The list could shape the wrong decisions

You may promote strong individual contributors out of roles where they excel while overlooking people whose leadership potential isn't reflected in current performance ratings.

Nobody asked you the real question. That is normal, and it is not their fault. Finding it is the work

Don't Answer the Request. Find the Decision

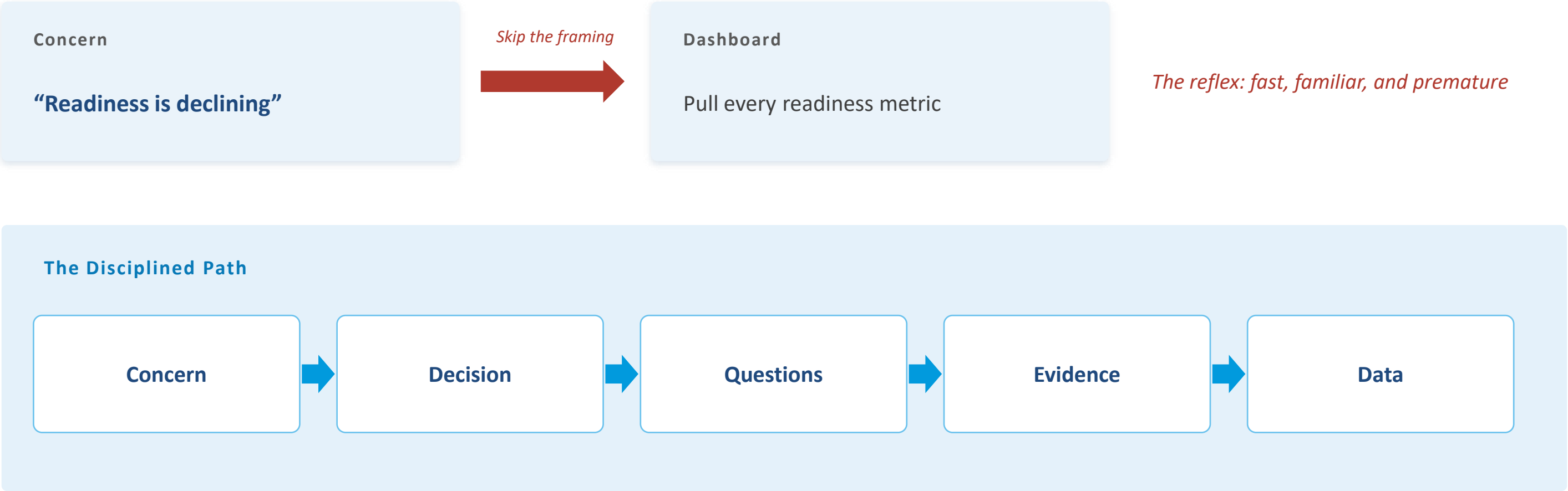
Generalize the principle

Request-first	Decision-first
Answer what was asked	Clarify the decision to be made
Use the data available	Identify the evidence needed
Produce the requested output	Analyze what would reduce uncertainty
Deliver the answer	Support a better-informed choice

Analysis cannot rescue a poorly framed question

A Concern Is Not Yet a Decision

A concern tells you something feels wrong, it does not tell you what to decide



The Framing Funnel

The decision comes before the question and the question comes before the data



Leadership Development Program Use Case

Leadership Request

“We've run the leadership development program for two years and the results look good. Graduates lead teams that rate 4.1 against 3.7 for everyone else. The proposal on the table is to expand it to every mid-level manager next year, and the budget decision lands this quarter. Before we commit, tell us whether it's working”

Framing Funnel #2 - What Decision Actually Needs to Be Made?

Separate the concern from the choice. A decision has options and consequences. A concern does not.

The Motive

Why decide now? What happens if we do nothing?

The Choice

What are we actually choosing between?

The Levers

What can we realistically change or fund?

The Boundaries

What is off the table, fixed, or out of scope?

In The Program Case

The decision is not “is the program working.” That is a question in service of a choice

The real decision: expand the leadership program to all mid-level managers next year, hold the current cohort, or redirect the budget. Do nothing and it auto-renews at current scope.

What is the first question this decision demands?

Are graduates better because of the program, or were they already better before they attended?

Framing Funnel #3 - What Does Success Look Like?

If you cannot say what better means, you cannot tell whether any recommendation works

Success measures

What does success look like, stated as something you could actually observe or measure? Name it before you start.

Boundary conditions

What is in scope and out. What authority, budget, and time you have. What is off the table before you begin.

In The Program Case

Success: Managers who complete the program lead measurably more effective teams within a year. Not attendance, not satisfaction scores. **Boundaries:** the L&D budget is fixed, and we cannot add headcount to deliver it

Boundaries: Within current headcount budget, no change to hiring authority, and a recommendation is needed before the next funding round

Framing Funnel #4 - What Are We Assuming?

Every framing hides assumptions. Name them so they can be tested



The program causes the improvement



Graduates and non-graduates are comparable



Performance ratings measure manager effectiveness



A year is long enough for the effect to show

Assumptions are not the enemy. Unexamined ones are. The skill is surfacing them early, then deciding which ones the evidence needs to test

Framing Funnel #5 - Turn the Business Question Into Analytical Questions

The business question is broad and human. Analytical questions are ones evidence can actually answer

1

Specific

Names the who, what, and when. No vague terms left standing

2

Scoped

Bounded to a population and a period you can actually look at

3

Answerable

Evidence for it could plausibly exist and be found

4

Decision-linked

The answer would change what you recommend or do

Framing Funnel #5 - From Concern to Answerable Questions

One business question, broken into questions evidence can settle

Business question: *Is the leadership program producing the manager improvement we expect?*

1

Do teams led by graduates outperform comparable teams led by non-graduates?

2

Were graduates already rated higher before they attended?

3

Does the effect hold across departments, or does it sit in one or two?

4

Were participants nominated, or did they select themselves?

Each one is scoped, answerable, and tied to a possible action. Compare that with “is the program working?”

Framing Funnel #6 and #7 - State a Hypothesis, And Try To Break It

A hypothesis is a testable claim. Rigor is naming what would prove it wrong before you look

Hypothesis

The program improves manager effectiveness, and that shows up in team performance within twelve months.

What Would Prove Us Wrong

- Graduates were already stronger performers before they attended
- The gap disappears once we account for who volunteered
- Teams led by graduates show no difference from comparable teams

If you only look for evidence that supports you, you will always find it. Name the disconfirmer first.

Decision Classification Reference Card

Size the decision before you size the evidence.

QUESTIONEVIDENCEREADEXPLOREEXPLAINVISUALIZERECOMMENDTELL

THE DECISION CLASSIFICATION

Size the decision before you size the evidence.
Turning Data Into Wisdom

Uncertainty

How much we do not yet know.
Earns more evidence when: unknowns are high.

Reversibility

Can we undo the decision?
Earns more evidence when: it is a one-way door.

Impact

How much rides on getting it right.
Earns more evidence when: the stakes are high.

Frequency

One-off, or a recurring call?
Earns more evidence when: it recurs (build a repeatable method).

Stakeholders

How many are affected, and must agree.
Earns more evidence when: many must agree.

Complexity

How many moving parts interact.
Earns more evidence when: the parts are tightly coupled.

Risk level

The cost of being wrong.
Earns more evidence when: failure is costly.

Time horizon

How long the consequences last.
Earns more evidence when: they are long-lived.

Size the decision, then size the evidence.

When To Use It

At the start of any analysis, to set how much rigor it deserves

What’s On It

Eight dimensions that tell you how hard to look

What Evidence Will This Decision Require?

Set the bar before you collect anything. The decision decides how rigorous the evidence needs to be.



Reversibility

Can we undo it cheaply?
Reversible calls need less evidence



Stakes

What is the cost of being wrong?
Higher stakes demand more



Uncertainty

How much do we not know?
More unknowns, more evidence

Low stakes and reversible: act on lighter evidence and adjust as you learn

High stakes and hard to undo: demand stronger evidence before you commit

You will not get certainty. You can get enough confidence to act.

Same Evidence, Different Action

How much evidence is enough depends partly on what you are about to do with it.

What We Have

Graduates of the leadership program lead teams that rate 0.4 higher. Most of that gap was there before they attended

Evidence: moderate

Confidence: low to moderate

Is this enough to expand the program?

Low Consequence · Reversible

Run one cohort with managers assigned rather than volunteering for two quarters

If we are wrong

If we are wrong, we have spent one cohort and learned something.

Act narrowly, and gather

High Consequence · Hard to Reverse

Expand the program to every mid-level manager next year, at full budget

If we are wrong

If we are wrong, we scale a selection effect for a year and cannot tell later whether it worked.

Gather. Defer the large move.

Evidence and confidence don't decide the action, the cost of being wrong does

Framing Funnel #8 - Choose the Evidence Before You Choose the Data

Decide what would actually convince you, then go find the data that provides it

Evidence-First

Name what would convince you and what would change your mind. Then collect the data that speaks to it. Honest, and it leads straight into weighing the metrics.

Data-First

Pull whatever is handy, then build a story around it. Invites cherry-picking, and you rarely notice the evidence you skipped.

This is where Module 3 begins

We take the questions and the hypothesis we just framed, and start weighing the workforce metrics as evidence

Workforce Readiness Use Case Reveal #1 - The Investigation Begins

Leadership Request

“We’re hearing concerns that workforce readiness is declining. Hiring is slower, critical positions remain vacant longer, certifications appear delayed, and employee turnover has increased in several areas. Before we invest in new initiatives, help us understand what is actually happening and what we should do.”

Leadership drops this on your desk. What metrics do you go pull?

Team Activity - Frame the Investigation (~20 min)

Run the readiness case down the funnel. Bring back a framed investigation

As a team produce

- The decision leadership actually needs to make
- What success looks like, and the boundaries
- Three assumptions hiding in the request
- Three analytical questions that pass the four tests
- One hypothesis worth testing
- The evidence that would prove your hypothesis wrong

At the debrief

Each team shares two things:

- 1. Your hypothesis**
- 2. The evidence that would break it**

Module 2 Team Activity Debrief

Run It Through the Funnel

Read us your analytical question.

Out loud, as written.

What decision does it support?

If none, it is a curiosity, not a question.

What would change your mind?

Say it before you look at anything.

Reflect

How much of your own analysis this year would survive that last question?

Module 2 - What We've Earned

You can turn a vague concern into a clear, answerable question with a testable hypothesis.

What's Next

What evidence would actually answer it

THE JOURNEY SO FAR

QUESTION

EVIDENCE

READ

EXPLORE

EXPLAIN

VISUALIZE

RECOMMEND

TELL

Module 3

Data as Evidence

Data is not a fact, it is evidence

Data Is Not Reality

The one idea that changes how you read every dashboard, report, and model

Reality

- Workforce readiness
- Morale
- Burnout
- Engagement
- Trust in leadership

Real, and impossible to observe directly

*we never see
reality directly*



Data

- Turnover = 14%
- Engagement score = 3.8 / 5
- Time to fill = 62 days
- Sick leave = up 9%

What we can actually collect

Data is evidence about reality, not reality itself. The metric is never the thing it stands for

“But I Sat in the Exit Interview”

Direct observation is evidence too. It is not unfiltered truth.

“I heard it myself. She told me she was leaving because of her manager. That is not a proxy, that is the truth.”

So What Did You Actually Hear?

Who agreed to talk

You heard the people who showed up and opened up. The angriest and the most guarded both stay silent.

What they would say to HR

People manage their exit. Few burn a reference by naming a manager outright.

What they knew

People report a reason. It is not always the reason. They may not know what actually tipped them.

What you were listening for

You already suspected something. Attention finds what it expects.

What It Is Good For

Depth, mechanism, and the words people actually use. It tells you what to go look for at scale.

What It Cannot Do

Tell you how common it is. Four vivid conversations feel like a pattern. They are four conversations.

Direct observation is not the opposite of evidence. It is evidence, with its own filters. Name them, and it gets stronger.

A Number Is Not Evidence Until There Is a Claim

Evidence is support for a claim. With no claim, a metric just sits there.

A Number

Turnover = 14%

True, and it supports nothing yet. Evidence of what?

Claim

“New supervisors are contributing to increased turnover”

*Now the number has a job. Does it support this claim? Contradict it?
How far does it actually reach?*

This is where Module 2 pays off. Your hypothesis is the claim. The metrics become evidence for it, or against it.

Evidence Is Not Explanation

Evidence is what you observe. Explanation is the account of why

Evidence

what you can observe

Turnover in two units is up. Exit interviews mention managers

observations, measures, records, testimony

Explanation

the why you infer

“Supervision is driving people out”

an account you have to test, not read off

The same evidence can support several explanations. Choosing among them is Module 6, done carefully, not now.

Evidence you can point to. Explanation you have to earn

Every Metric Is a Proxy

Nobody measures readiness. We measure things that stand in for it, and each has a gap

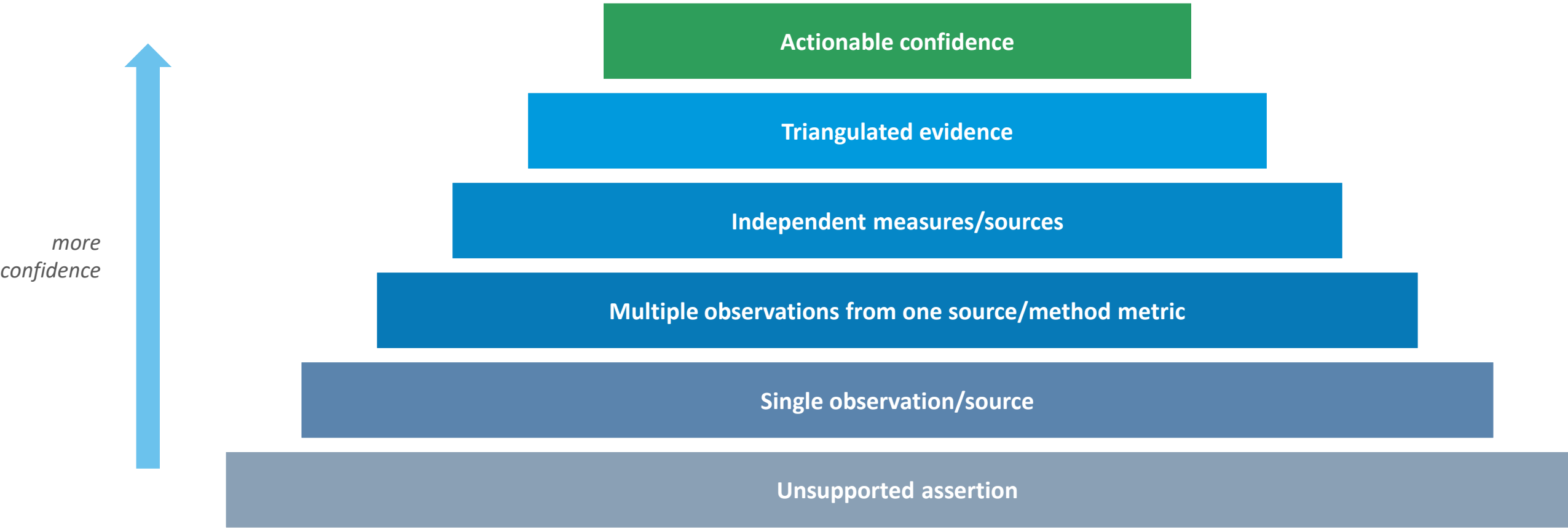
Proxy	Stands in for	The Gap it leaves
Vacancy time	Staffing of critical roles	Says nothing about the capability of who remains.
Training completion	Readiness to perform	Finishing a course is not being able to do the job.
Turnover	Stability of the workforce	Who left matters far more than how many.
Engagement score	Morale and commitment	A survey moment, shaped by who chose to answer.

Name what your proxy stands for, the gap it leaves, and whether a change is real or just a change in how it was counted.

Remember Module 1: training completion was not readiness. That was not just the wrong metric. It was the wrong proxy.

The Data as Evidence Pyramid

Evidence is not true or false. It is weak or strong. The climb is how much confidence it earns.



An interview can move up this pyramid. So can a metric. What moves it upward isn't whether it's a number. It's rigor, independence, and convergence

How Much Evidence Is Enough?

The bar is not fixed. It rises with what the decision puts at risk

Earlier we set the bar, before we collected anything. Now the evidence is in front of us. The question has flipped: does what we have clear it?

Six Checks On The Evidence You Now Have

Strength

How far up the pyramid does it sit? One metric is not several that agree.

Independence

Are the sources genuinely separate, or one source wearing four hats?

Convergence

Do different methods point the same way, or only the ones we favour?

Contradiction

What argues against it, and have we actually looked for that?

Absence

Whose data is missing, and would it change the answer if we had it?

Uncertainty

What remains unknown, and can we name it out loud?

Clears a moderate bar

Three independent sources converge, we looked for contrary evidence and named what is missing

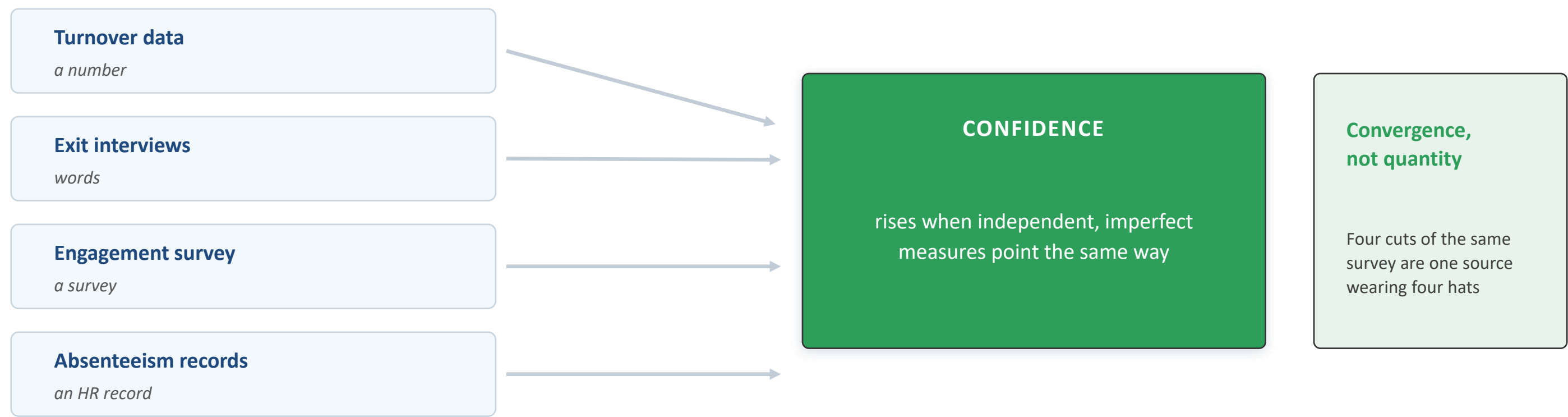
Clears no bar at all

One metric, no comparison, and nobody checked what would argue the other way

Enough is not a quantity. It is a relationship between what you have and what you are about to do with it.

One Proxy Is Weak. Several That Agree Are Strong

Confidence comes from independent, imperfect measures telling the same story.



Different methods matter more than more numbers. A number, an interview and an HR record together are triangulation. Four slices of one survey are not

And what is absent counts too: the voices not surveyed, and the people who already left

When the Evidence Disagrees

Triangulation does not promise agreement. When sources conflict, that is information too.

The survey says pay. The exit interviews say managers. The turnover data tracks supervisor tenure. They do not line up.

Don't ask: "Which one is right?"
Ask: "Why are they different?"

Why might they disagree?

Are they measuring different things?

a source may be answering a different question

Does each source have different filters?

Selection, self-report, timing, incentives, missing voices

Is the explanation more specific?

The answer may differ by role, unit, population, or circumstance

Agreement can increase confidence. Disagreement should increase curiosity.

Worked Example - The Number Arrives. Now What?

The claim we are testing: the leadership development program improves manager effectiveness, enough to justify expanding it

What Arrived

Teams led by program graduates score 12% higher on performance ratings.

It Is A Proxy, For What?

“Performance rating” stands in for team effectiveness.
The gap: ratings are judgments, made by people, on different scales.

What It Can Tell Us

- Graduates' teams rate higher than non-graduates' teams
- The difference is large enough to notice, not noise
- Something distinguishes these two groups

What It Cannot Tell Us

- Whether the program caused the difference
- Who these managers were before they attended
- Whether raters knew who had attended

Evidence strength so far: weak **Confidence the program works:** low

This is an important data literacy skill: being able to learn something from a number without claiming that it tells you more than it does

Worked Example - What Happens as Evidence Accumulates?

Evidence accumulates. It rarely proves

1	Performance ratings SUPPORTS	Graduates' teams score 12% higher	<i>Single metric. Points toward the program</i>
2	Prior-year ratings UNDERCUTS	Graduates were already rated higher before they attended	<i>This is the disconfirmer we named earlier</i>
3	Participant survey MIXED	Graduates say the program changed how they lead	<i>Self-report, from people invested in it</i>
Evidence strength: mixed Confidence: low to moderate Action: test further before expanding <i>Not “the program works.” Not “the program fails.” Run one comparable cohort, and check whether raters knew.</i>			
<i>When sources disagree, that is information. Ask why they differ before you average them into mush.</i>			

Team Activity Reveal #2 - The Dashboard Arrives



Team Activity - Three Sources That Do Not Agree (~20 minutes)

Your team has three sources on the compensation claim. They point different ways. Reconcile them.

As a team, produce

Source 1: the dashboard. Turnover is up 4 points

Source 2: a senior manager. “It is pay. I hear it constantly.”

Source 3: recruiting. Candidates are accepting our offers at the usual rate

Where does each sit on the pyramid, and why do they differ?

.

At the debrief

Each team reports:

1. Why the sources disagree
2. Its one sentence: strength, confidence, action

Disagreement is information. Do not let a team average it away.

Module 3 Team Activity Debrief

Three Sources. They Do Not Agree.

Where did each source sit on the pyramid?

And why there

Why do they point different ways?

Do not average. Explain

Give us your sentence.

Strength, confidence, action

Reflect

Is your confidence in “it’s compensation” lower than it felt this morning?

What We’ve Earned

You can treat data as evidence, with a strength and a confidence, not as truth

What’s Next

Reading Data Correctly

THE JOURNEY SO FAR

QUESTION

EVIDENCE

READ

EXPLORE

EXPLAIN

VISUALIZE

RECOMMEND

TELL

Module 4

Reading Data Correctly

Slow down before making meaning

The Most Common Mistake

A dashboard says

Engagement is 71%

Someone in the room says “morale is fine” or “that is low”

You are already skipping a step

Before you believe it, what would you want to know?

Before You Act on Data

There are three different questions to work through.



Understand

See clearly

What the data actually is, before any meaning



Interpret

Make meaning

What it might be saying, while holding more than one explanation



Evaluate

Judge wisely

Whether to trust it, and whether it justifies action

Slow down. Understand it. Interpret it. Then decide whether to act.

Understand

U

I

E

Before any meaning: what is this data, really?



What exactly am I looking at?



Who or what is included, and who is left out?



Compared to what?



Is this the full picture, or a slice?

Skip this and you act on surface impressions and misread the chart in front of you

Understand - Run the Checks

Those are the questions you probably already ask instinctively. This is a more systematic way to run that check

Definition what exactly is being counted?	Unit & timeframe per what, and over what period?	Population who is included, and who is left out?
Denominator a raw count, or a share of something?	Missing values what is not in the data at all?	Source & method who produced it, and how?
Denominator, In One Line Overtime hours rose 12%. But headcount rose 15%. Per person, overtime fell. The count went up; the rate went down		

Interpret

U

I

E

Now, what might the data be saying? Hold more than one answer



What is this actually showing me?



What else could explain it?



How certain am I?



Where does the data stop and my assumption start?

Interpret - Where Meaning Creeps In

Notice the exact point where the data stops and your inference begins



Watch the jump

“Engagement fell because people dislike leadership.” “Sales rose because of the campaign.”

Each sentence leaps from a pattern to a cause

The data can show a pattern. It cannot hand you the reason.

Evaluate

U

I

E

Is this insight trustworthy, and does it justify action?



Is this trustworthy and valid?



Is there enough evidence to act?



What decision does this actually support?



What am I still assuming?

Confidence, not certainty. Say low, moderate, or high. This is the pyramid from Module 3

Worked Example - Reading a Readiness Number

Vacancy time in critical roles is up 30%



Understand

- Where did this come from?
- Which roles count as critical?
- Up 30% compared to when?
- A few roles, or many?



Interpret

- A steady climb, or a recent spike?
- Could a retirement wave or a policy change explain it?
- How sure are we it is worsening?



Evaluate

- Is the definition of critical stable enough to trust?
- Enough to act, or do we need role-level detail?
- What decision does this support?

What we can say: Vacancy time appears to be increasing

What we cannot say yet: Why, how widespread it is, or what action it warrants

Worked Example #2 - Reading a Comparison

“Teams led by program graduates rate 4.1. Teams led by non-graduates rate 3.7.”

UNDERSTAND

What are we actually looking at?

A rating on what scale, given by whom? Are the two groups the same size? Rated when, relative to attending? Nothing here says how many people sit behind either number.

INTERPRET

What could a 0.4 gap mean?

The program made them better. Or better managers volunteered. Or raters knew who attended. Or one strong cohort is carrying the average. Four readings, all consistent with 4.1 and 3.7.

EVALUATE

How far would we trust it?

Enough to look at what graduates scored before they attended. Not enough to expand the program. Confidence: low to moderate.

An important question: what did graduates score BEFORE they attended? That is the question Evaluate produced

Reading Red Flags

Eight things UIE helps you notice

No baseline a number with nothing to compare	No denominator a count posing as a rate	Wrong comparison measured against the wrong thing	Changed definition the metric quietly redefined
Incomplete population the ones who matter left out	Average hides the spread one mean over many stories	Correlation language “because” doing a cause's job	False precision decimals the data cannot support

Every one of these is a jump the data did not earn

You don't need to memorize eight red flags. If you get into the habit of Understand, Interpret, Evaluate, these are the kinds of problems you'll start noticing

Team Activity - Read It Before You Believe It (~15 minutes)

Take one metric from the readiness dashboard and run it through the three steps (U-I-E).

As a team, produce

- Two things you must **understand** before reading anything into it
- Two **explanations** other than the obvious one
- One statement of how much you would **trust** it to act

At the debrief

Each team shares:

1. Its trust level for the metric
2. The one question it most wants answered

Those open questions are where Module 5 picks up.

Module 5

Exploring the Data

Explore before you explain

Explore Before You Explain

Module 4

Can I trust what I am looking at?

UIE

Module 5

What patterns exist?

Six Analytical Moves

Module 6

Why do those patterns exist?

Causal reasoning

*We framed the question, weighed the evidence, and made sure we knew what we were looking at
Now we explore it from different angles to see what patterns emerge, before we try to explain why*

Analytics Answers vs. Analytical Thinking

The tool shows you something. The thinking determines where to look next.

Analytics

Produces an output

“Turnover is up 14%”

Analytical Thinking

Decides what to examine next

Compared to what? Where is the increase? Is it widespread or concentrated? What else moves with it?

The Six Analytical Thinking Moves

Six ways to look at the data before you explain it

1

Compare before you conclude

“Compared to what?”

2

Segment before you generalize

“Does the pattern hold in the groups?”

3

Check the rate, not just the count

“How big is it relative to the whole?”

4

Separate level from movement

“Is it high, or is it rising?”

5

Look for what moves together

“What else changes when this changes?”

6

Chase the outlier before you trust the average

“What is different about the exceptions?”

Worked Example - Applying the Moves on the Leadership Program Question

“We've run the leadership development program for two years and the results look good. Graduates lead teams that rate 4.1 against 3.7 for everyone else. The proposal on the table is to expand it to every mid-level manager next year, and the budget decision lands this quarter. Before we commit, tell us whether it's working”

Compare before you conclude	Graduates' teams versus non-graduates', and versus the same managers before they attended
Check the rate, not just the count	40 graduates sounds like a lot. It is 12% of managers. The denominator changes how we read the number
Separate level from movement	Were they better after the program, or already better before it? Level, not change
Segment before you generalize	The difference is concentrated in two departments. Everywhere else it is flat
Look for what moves together	The gap tracks with prior rating, not with time since the program
Chase the outlier before you trust the average	One cohort of eight accounts for most of the difference

Do Not Explore Only Where You Expect to Find the Answer

You can use all six moves correctly and still only look where you already believe the answer is

If You Already Think		You May Only Look At	
Turnover is caused by pay	➤	Pay differences, and never the supervisor data	
One department is the problem	➤	Department comparisons, and never the role families	
The program worked	➤	Evidence of improvement, and never who volunteered	
Younger employees are leaving	➤	Turnover by age, and never turnover by tenure	
Your assumptions decide where you look and where you look decides what you find			
Use your hypothesis to start the exploration, not to limit it			

How to Think While You Explore

The Six Moves tell you where to look, these principles help you stay open to what you find

1 Look beyond your first explanation

2 Try more than one comparison

3 Look for patterns that challenge your hunch

4 Pay attention to exceptions

5 Stop at what the data shows

The goal of exploration is not to prove your first idea. It is to see what else might be true

Team Activity Reveal #3 - The Detail Arrives

New Data

Leadership sends the segmented data. Turnover, vacancies, and tenure, broken out by department, grade, supervisor, and location

And They Repeat

“Looks like compensation to us”

Your job now is not to agree or disagree, it is to explore

Team Activity - Perform the Investigation (~20 min)

Take the segmented data and put the moves to work on the compensation claim

As a team, produce

- Apply at least four of the six moves on the claim
- Two patterns you can defend with the data
- The one question those patterns raise
- Do not explain the patterns. That is Module 6

At the debrief

Each team names:

- 1. Its strongest pattern**
- 2. Its sharpest question**

We carry both into Module 6, where we start explaining.

We Want a Pattern Here

“Turnover concentrates among first-year supervisors”

Not An Explanation (That is Module 6)

“First-year supervisors cause the turnover”

Module 5 Team Activity Debrief

Run the Investigation

Which moves did you run?

And what did each one show?

Two patterns you can defend.

Patterns. Not causes.

What question do they raise?

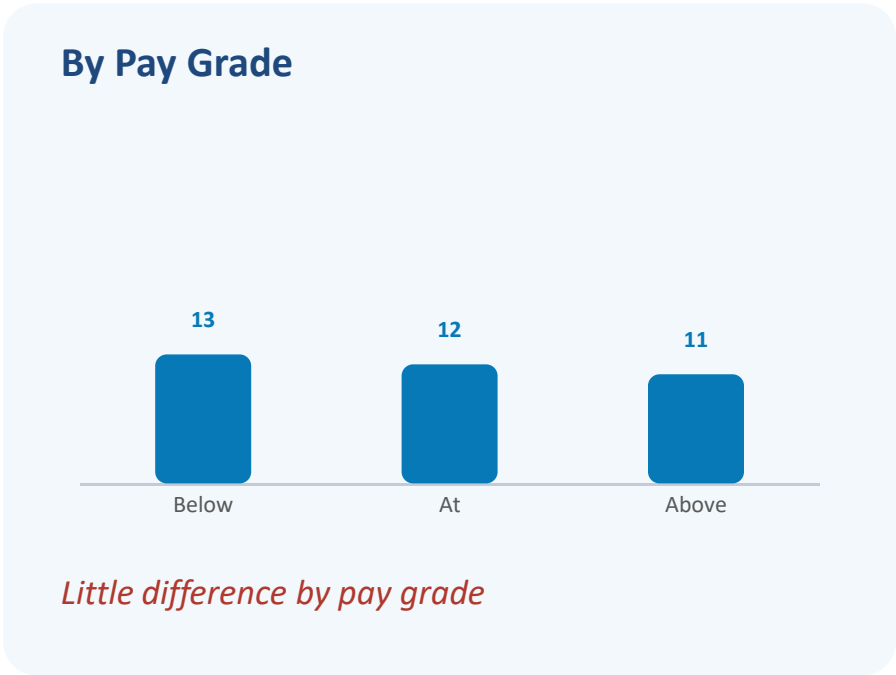
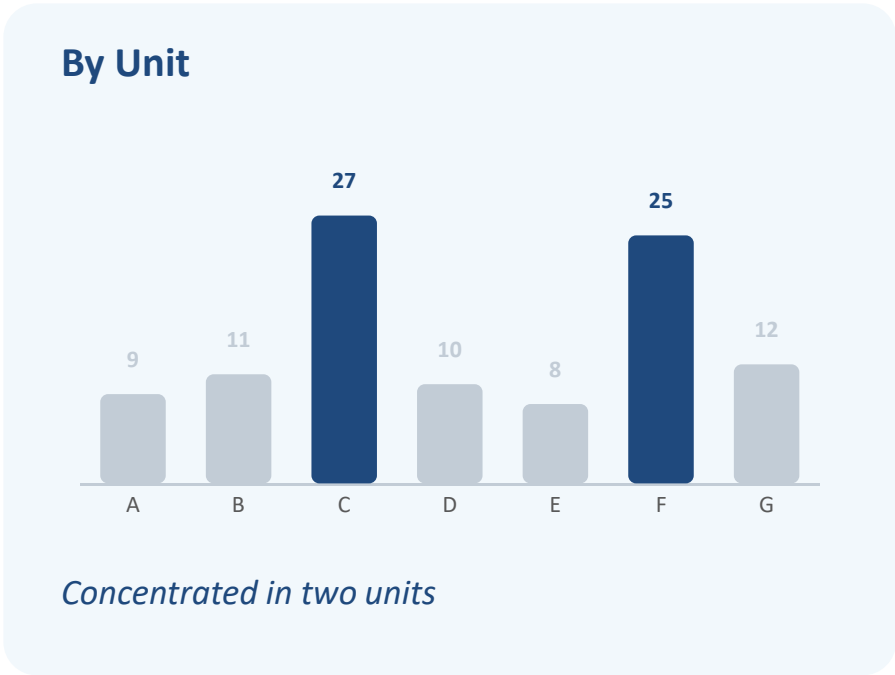
This is the output. Not an answer.

Reflect

Where did you feel the pull to explain, and did you resist it?

Segment Before You Generalize

Three different cuts reveal three very different patterns



Different cuts reveal different patterns. The strongest signal is supervisor tenure, so that's where we investigate next.

Module 5 - What We've Earned

You can surface patterns you can defend

What's Next

Patterns create questions, not causes

THE JOURNEY SO FAR

QUESTION

EVIDENCE

READ

EXPLORE

EXPLAIN

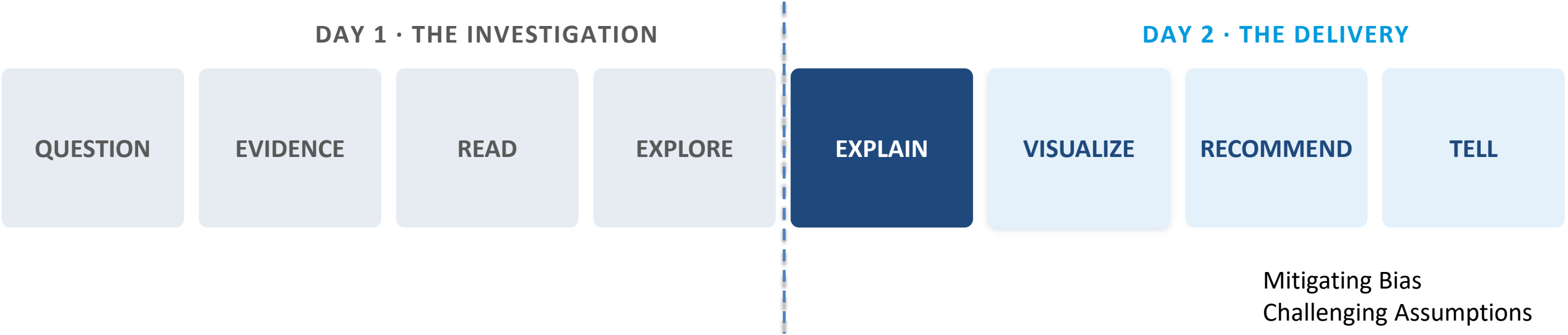
VISUALIZE

RECOMMEND

TELL

Where We Are – Day 1 Recap

The investigation is done. Day 2 is the delivery, and it starts with seeing it clearly



A NAVAIR installation at China Lake conducts a survey to determine workforce satisfaction with range facilities, test infrastructure, and support resources. They find that 70% of personnel report being satisfied, with a margin of error of $\pm 5\%$. Leadership concludes that the majority of the workforce is satisfied.

Do you accept this insight? Why or why not?

Do these results actually mean “most personnel are satisfied”? What if the survey was conducted right after a new lab upgrade, expanded telework policy, or a high-profile test program milestone? Could timing bias the results and misrepresent typical satisfaction levels?

Sample Representativeness

- Who was surveyed (engineers, flight test crews, range operators, civilian staff, contractors)?
- Do they represent all divisions (EW, weapons integration, telemetry, range ops)?

Response Rate

- What percentage of personnel responded?
- Could non-responders (like flight line crews with less time) hold systematically different views?

Margin of Error

- 70% $\pm 5\%$ means satisfaction could range from 65% to 75%.
- Does that swing matter for decision-making (e.g., calling satisfaction “high” vs. “mixed”)?

Context and Benchmarks

- How does this score compare to prior surveys at China Lake or similar sites like Point Mugu?

Survey Bias or Limitations

- Were the questions phrased neutrally, or did they emphasize recent improvements?

Actionability

- Does this result warrant leadership action (e.g., further investment in facilities, morale initiatives)?
- Or is the variation within expected norms for a large technical installation?

A report shows that 90% of personnel who completed a new logistics training program remained in their roles after 12 months, compared to 70% of personnel who did not complete the training.

Leadership concludes the training program is clearly driving higher retention. Do you agree? Why or why not?

Is the Training Truly Causing Higher Retention?

Causation or Correlation?

- Could personnel who choose to complete training already be more motivated or engaged?
- Is training driving retention, or just correlated with it?

Self-Selection Bias

Are those completing training different in motivation, need, or commitment compared to those who didn't?

Control Factors

Were other factors controlled? (e.g., rank, assignment type, years of service)

Alternative Explanations

Could other initiatives (mentorship programs, improved work conditions, leadership engagement) be contributing to the difference?

Next Steps Before Acting

- Consider testing with a randomized or phased rollout
- Analyze additional personnel attributes to isolate the training's impact
- Investigate whether training completion is acting as a proxy for already highly engaged individuals, rather than being the driver itself.

Day One Recap

Where We Got To Yesterday

1

The connections are the work

Data does not become a decision on its own. Every step up the ladder is a judgment somebody makes

2

The question decides everything

A request is not a decision. Frame it, size it, and name what would prove you wrong before you look

3

Evidence is weak or strong, not true or false

Every metric is a proxy standing in for something. Ask what it can and cannot establish

4

Read the number before you believe it

Understand what it is. Interpret carefully. Evaluate how far it will carry you

5

Explore before you explain

Six moves to find patterns. The output is patterns and sharper questions, never causes

Yesterday was about not concluding too early. Today is about deciding and communicating

Same discipline, opposite pressure. Yesterday we slowed you down. Today we have to get you to commit

The Case, So Far

Three reveals in. Here is everything you have been given, and what it has done to your answer

The request	Leadership believes readiness is slipping and it is a pay problem.	<i>You started here, and so did we.</i>
Reveal 1 and 2	The dashboard. Readiness down 7 points, turnover up from 12% to 16%, time to fill up 30%.	<i>Something is happening. Nothing about why.</i>
Reveal 3	The segmented data. Turnover by unit, by supervisor tenure, by pay position.	<i>This is where it got interesting.</i>

What you Found During the Activities		
• Turnover is concentrated, not general	Two units at 25 and 27 percent. Everywhere else, 8 to 12.	
• It travels with supervisor tenure	24% under a first-year supervisor. 8% under one with three years.	
• It does not travel with pay position	13, 12, 11 percent. Almost flat across the whole pay range.	

You have three patterns and no cause. That is exactly where Module 5 was supposed to leave you.

Four Questions From Yesterday You Will Use Today

“What decision does this actually serve?”	Module 2	Before you pull anything. If nobody is deciding anything, you are producing a report, not analysis.
“What would change my mind?”	Module 2	Written down before you look. It is the single cheapest protection against finding what you expected.
“Compared to what?”	Modules 4 and 5	No number means anything on its own. This is the question that killed most of yesterday's demos.
“Who is missing from this data?”	Modules 3 and 5	The helmets. The people the dataset cannot describe are usually the ones who matter.

What Today Does

6

Explain

You have patterns. Today you make explanations compete, and find the evidence that tells them apart.

8

Show

Design the picture so the finding survives being looked at. Clarity is a safety property.

9

Recommend

Turn a finding into a decision, proportioned to what you actually know.

10

Tell

Make it land with people who were not in the room, without changing what is true.

Then the capstone: your team presents the whole case, start to finish

Module 6

From Pattern to Explanation

Build the best-supported explanation

One Pattern Can Tell Many Stories

Turnover is concentrated in two units, and falls steeply as supervisor tenure rises

Supervision quality

Workload and overtime

Compensation

A retirement wave

Training and readiness gaps

A recent policy change

Who gets assigned where

Something we have not thought of

Every one of these is consistent with the pattern. The data cannot rule any of them out yet

Turn Each Explanation Into a Prediction

If this were true, what else should we see? Write it down before you look at the evidence that would test it.

If This Were True

We Should Also See

Supervision quality



Turnover clustered under particular supervisors, and engagement lower under them

Workload and overtime



Turnover rising with overtime hours, and highest where staffing is thinnest

Compensation



More exits where pay sits furthest below market, regardless of unit

A retirement wave



Exits concentrated among retirement-eligible employees

Training and readiness gaps



Turnover concentrated among newer or less-trained employees, and training exposure differing between the affected units and the rest

NOT AN EXPLANATION

Supervision is an artifact of who gets assigned where



New supervisors concentrated in the hardest units. Then the unit explains BOTH the turnover and the tenure pattern.

Not All Evidence Helps You Choose

Some evidence fits every explanation. That kind cannot decide anything.

Turnover rose from 12% to 16%	Every explanation on the list predicts this	NOT DISCRIMINATING
Turnover falls as supervisor tenure rises	Fits supervision. It also fits the assignment account, if new supervisors are placed in the hardest units. Two explanations, same prediction	WEAKLY DISCRIMINATING
Turnover is flat across pay position	Compensation predicts a slope. There is none. This one source rules an explanation out	STRONGLY DISCRIMINATING
Training completion is 82% and flat across units	The training-gap explanation predicts that the affected units are less trained. They are not. Exposure is the same everywhere, so it cannot explain the difference	STRONGLY DISCRIMINATING
Ask of every piece of evidence: which explanation does this rule OUT? Evidence that rules nothing out is background, however solid it is.		

The Explanation Canvas

OBSERVED PATTERN	what the exploration found. This stays fixed while the explanations change.
------------------	---

Explanation	If true, we would expect	What we see	What does not fit	What we still need
one candidate per row	the prediction, written first	does the prediction hold?	what argues against it	what would decide it

Best-supported explanation, and at what confidence	it may be a combination, not a winner
--	---------------------------------------

What we set aside, and what ruled it out	name the prediction that failed
--	---------------------------------

Next step	act / act narrowly and gather / gather / reject
-----------	---

Worked Example - Predictions First, on a Different Case

The pattern: graduates of the leadership program lead teams that rate 0.4 higher.

So We Wrote The Predictions, Then Went and Looked. Prior ratings for both groups · how people came to be on the course · whether raters knew who attended

Explanation	If true, we would expect	What we see	Does not fit	Verdict
The program develops managers	Graduates should be similar to others BEFORE attending	They were not. Graduates rated 3.9 before, others 3.6	Most of the gap predates the course	Ruled out
The program selects strong managers	A gap should exist before attendance	It does. 3.9 against 3.6	Does not explain the small gain after	Strong
Better managers volunteer	Self-selected cohorts should outperform assigned ones	Attendance was self-nominated throughout	Two cohorts were manager-assigned, and we have not checked them	Moderate
Raters know who attended	Blind raters should show a smaller gap	No blind ratings exist	No evidence either way	Untested

Say It as Confidence, Not Certainty

The strength of the evidence sets the strength of the claim.

	Strong evidence	Several independent sources converge, predictions held, we looked for what would break it	High confidence	<i>“The evidence supports X”</i>
	Moderate evidence	Sources agree but share a weakness, or one prediction is untested	Moderate confidence	<i>“X is the best-supported explanation, at moderate confidence”</i>
	Mixed evidence	Sources point different ways and we cannot yet say why	Low confidence	<i>“We cannot separate these yet. Here is what would”</i>
	Weak evidence	One source, no prediction tested, nothing ruled out	Low confidence	<i>“This is a hypothesis, not a finding”</i>
	Never say “the data shows.” Say what it supports, how strongly, and what you set aside.			

Team Activity Reveals 4, 5, and 6 - The Evidence Arrives

Reveal 4 • Exit interviews

Managers and workload are cited most often, and mostly in the affected units

Reveal 5 • Training and certification

Lapses are rising, but they start after the departures, not before

Reveal 6 • Supervisor and engagement

Engagement is lowest under the newest supervisors. Turnover tracks supervisor tenure

Team Activity - Build the Case (~25 minutes)

Use the Explanation Canvas and the evidence from Reveals 4 through 6

As a team, produce

- The field of explanations, not just one
- A filled Canvas for your top three
- A confidence rating for each, with the reason
- One next step: act, gather more, or reject

How you present

Not “we solved it.”

Instead: “Our strongest-supported explanation is X, at Y confidence, because Z.”

We carry your explanation and confidence into Module 7.

The Answer Is Usually a Combination, Not a Winner

Five explanations go in. One does not necessarily come out.

The main contributor

What is doing most of the work

Supervision. Turnover falls steeply as supervisor tenure rises, and three sources agree.

What amplifies it

A condition that makes the effect worse where it is present

Workload. In two small units, every departure raises the load on whoever remains.

What triggers it

What turns dissatisfaction into a resignation on a particular week

We do not know. Nothing in our data tells us why people left when they did.

What we ruled out

An explanation whose prediction failed

Compensation. It predicted a slope across pay position. There is none.

“Supervision is the strongest contributor, amplified by workload in the two understaffed units. Pay does not explain the unit differences.”

That sentence is longer than “the cause is supervision,” and it is the one you can defend.

Module 6 Team Activity Debrief

Strong Responses

- ✓ “Supervision, at moderate confidence, because three independent sources converge. We set compensation aside: turnover is concentrated and does not track pay grade.”
- ✓ A team that catches the timing on certification lapses. They rise after the departures, so the backlog is a symptom.

Weak Responses, And Why They Are Weak

- A canvas with only the “evidence for” column filled in.

Why it is weak: That is advocacy with a canvas on top. The against and missing columns are where the honesty lives.

- “We are certain it is supervision.”

Why it is weak: Three converging sources earn moderate, not certain. You still do not know why those supervisors struggle.

Five Ways to Get This Wrong

Check your canvas against these before you take it anywhere.

THE TRAP	WHAT IT LOOKS LIKE	ASK YOUR TABLE
1 Plausible is not supported	It sounds right, and nobody asked what evidence would separate it from the alternatives.	<i>“Did we write a prediction, or just a story?”</i>
2 Travelling together is not causing	Two things move together. A third thing may be moving both.	<i>“What else could produce this pattern?”</i>
3 Looking only where you expect	Every column filled with evidence for, and nothing under what does not fit.	<i>“What did we go looking for that could have proved us wrong?”</i>
4 Stopping at the first fit	The first explanation that fits gets adopted, and the others never get a prediction.	<i>“Did every explanation get the same test?”</i>
5 Ignoring what is missing	The strongest evidence is often the data we do not have.	<i>“Who is absent, and would they change the answer?”</i>

These are not five more things to remember. They are five checks to run against the canvas you just filled in.

Changing Your Mind Is the Win

You came in this morning with an answer. Where is it now?

YESTERDAY MORNING

“It is compensation.”

High confidence. One source. No prediction written down, and nothing ruled out.

NOW

“Supervision, at moderate confidence.”

Three sources converge. A prediction failed and eliminated pay. We can name what we still cannot say.

Your confidence went DOWN, and the analysis got BETTER. Those are not in tension.

Nobody was wrong yesterday

Compensation was a reasonable first explanation. It was the one leadership offered and it fit the pattern.

What changed was the method

We wrote down what it would predict, and then checked. The prediction failed.

This is the part that is hard at work

Saying “I thought it was pay and the evidence moved me” costs something. It is also the most credible thing you can say.

Take a show of hands: who has changed their mind since yesterday morning?

Module 7

Visual Thinking

Charts by purpose, not software

London, 1854

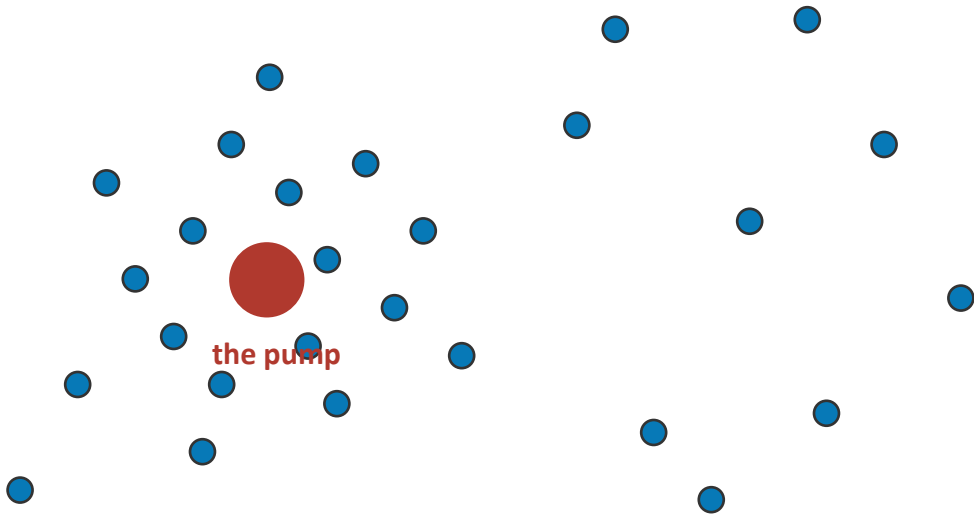
Same deaths. Two ways of holding them. Only one answers the question.

The Table They Had

Broad Street	18
Cambridge Street	12
Marshall Street	9
Poland Street	7
Berwick Street	6
Silver Street	5
Dufours Place	4
Bentinck Street	3

Deaths by street. Accurate, complete, and it tells you nothing about where the cause is

The Map He Drew



Each dot is a death and the cluster is the answer.

The question was “where?” A table cannot answer where. The question picked the chart.

Where We Are

The investigation is done. Day 2 is the delivery, and it starts with seeing it clearly



The Question Picks the Chart

The Wrong Question

“What chart should I use?”

Open the tool, grab the default bar, decorate. The chart leads and the meaning follows, if it shows up at all

The Right Question

“What question am I answering?”

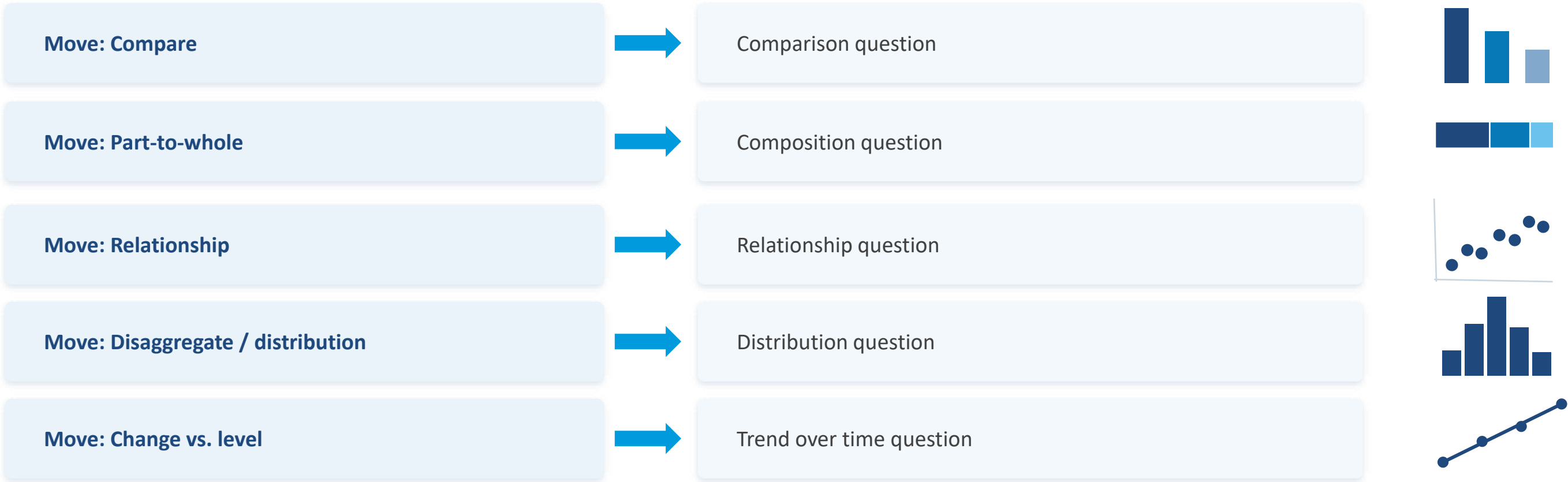
Name the question, the variables, and the counts. The chart follows almost on its own

Start with the question, not the data. Charts by purpose, not software

Charts do not create insight. They either preserve it or lose it

Yesterday's Move Picks Today's Chart

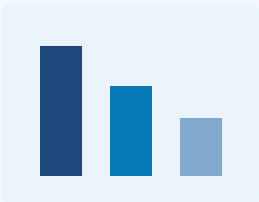
The analytical move you used to explore is the same axis that chooses the chart



The move chose the question. The question chooses the chart. One continuous line from exploring to showing.

Six Questions, Six Chart Families

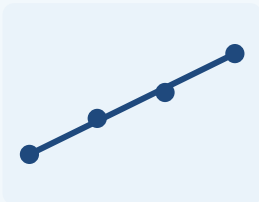
Name the question type, and the family follows



Comparison

“How do they rank or differ?”

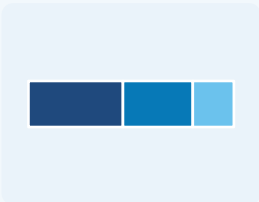
Bar, sorted bar, dumbbell



Trend over time

“How is it changing?”

Line, slope, area



Part-to-whole

“What is it made of?”

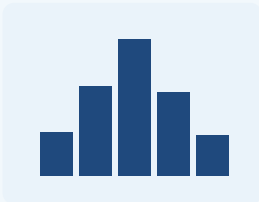
100% stacked, sorted share



Relationship

“Do they move together?”

Scatter, bubble



Distribution

“What is the average hiding?”

Histogram, box plot



Flow through a process

“Where does it drop off?”

Funnel, Sankey, step-rate

This Is Science, Not Art

Question type, then variable types, then counts

1

Question type

Comparison, trend, part-to-whole, relationship, distribution, flow.

2

Variable types

Qualitative or quantitative. And does the measure have a true zero?

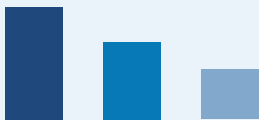
3

The counts

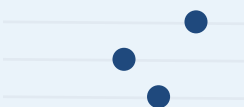
How many categories, groups, series, or points.

The One Rule That Separates Correct From Wrong

Does the measure have a true zero?



Yes, ratio → bar. Length needs a zero

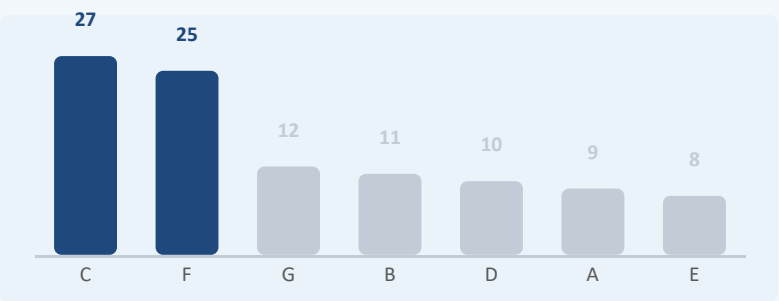


No, interval → dot plot. Position can start anywhere

Worked Example - The Explanation Becomes a Picture

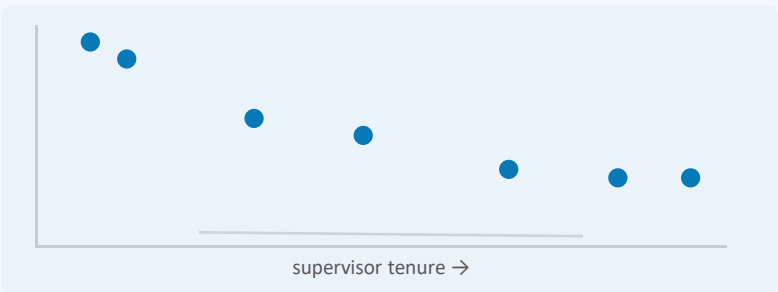
Three questions from the readiness case

“Is the turnover concentrated in a few units?”



Sorted bar, or Pareto if concentration is the point

“Does turnover track supervisor tenure?”



Scatter. Two measures, one point per unit

“Is it uniform with pay across units?”



Dumbbell. The gap between two values, per unit

Four Ways Chart Choice Goes Wrong

The traps that happen before the design work even starts

1

Chart first, question never

Opening the software before naming the question you are answering

2

A pie doing an ordinal job

Circles lose order and hide close shares. Stages and ratings belong on a bar

3

Bars without a true zero

Length encodes value, so a cut baseline lies with it

4

Too many lines at once

Five or more crossing series is spaghetti. Switch to small multiples

Team Activity - Pick the Chart (~15 minutes)

For each finding from your investigation, name the question type and the chart

As a team, for three findings

- Write the question in words first
- Name the question type: comparison, trend, part-to-whole, relationship, distribution, flow
- Name the variables, and whether the measure has a true zero
- Pick the chart, and say why it, not the alternative

Use the card

The Which Chart card at your table takes you question to family in one step.

If you cannot write the question in words, that is the finding: no chart yet

Module 7 Team Activity Debrief

Pick the Chart

Read the question in words.

Before you name any chart.

What type, and is there a true zero?

That pair picks the family.

Why that chart, and not the alternative?

Reflect

How often do you pick the chart before you have written the question?

The Which Chart Reference Card

Let the question pick the chart, not the software



When To Use It

Before you build any visual

What's On It

Question types mapped to chart families

Module 7 - What We've Earned

You can choose the right chart for the question

What's Next

Making that chart clear and honest

THE JOURNEY SO FAR

QUESTION

EVIDENCE

READ

EXPLORE

EXPLAIN

VISUALIZE

RECOMMEND

TELL

Module 8

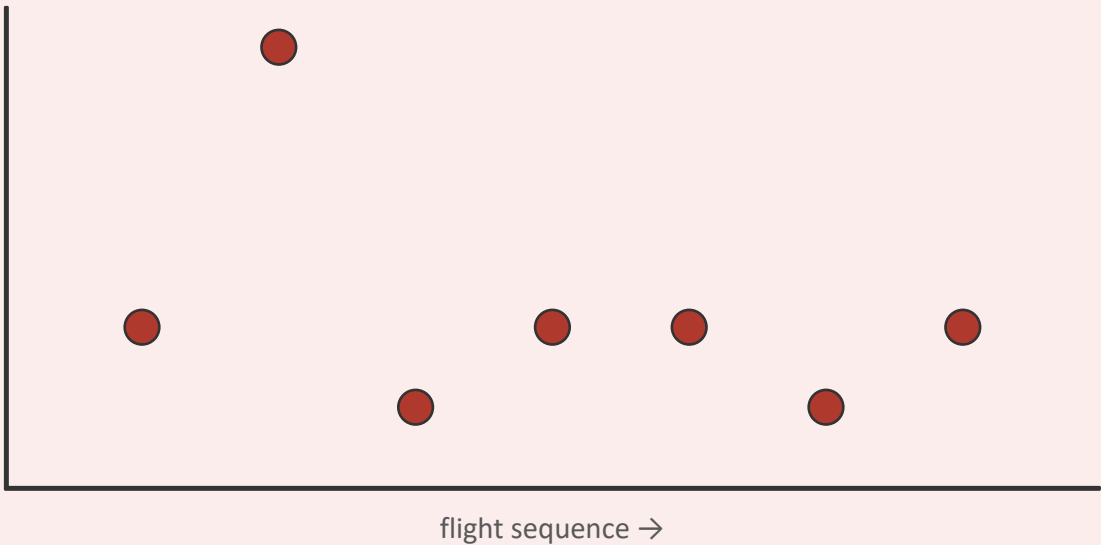
Designing Better Visualizations

Make the right chart honest and clear

The Night Before a Launch, 1986

The data existed but the chart did not let anyone see it

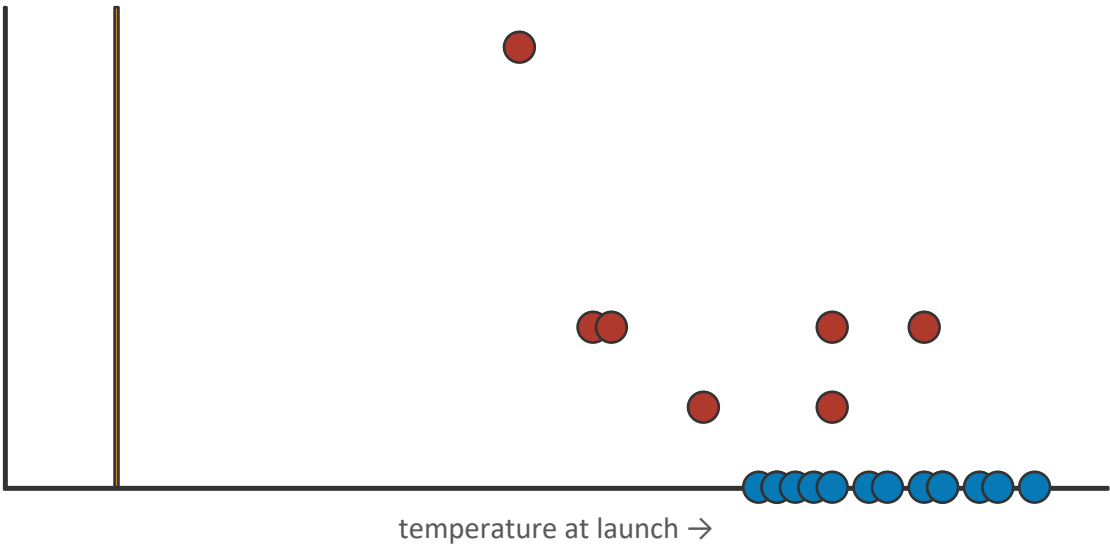
What The Charts Showed: Damaged Flights Only



No pattern. Nothing to argue with.

All 24 Flights, Plotted Against Temperature

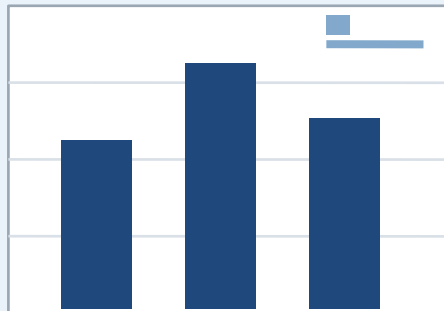
launch: 31°F



Damage concentrates in the cold. The coldest prior launch was 53°F.

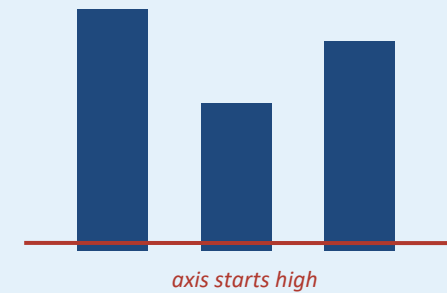
The Right Chart Built Wrong Still Lies

Confusing



The point is in there, buried under gridlines, borders, and a legend. The reader gives up before finding it.

Misleading



The chart argues something the data does not. A cut axis, a missing denominator, a verdict in the color.

Design has two jobs: clarity and honesty

Yesterday we protected ourselves from being fooled. Today we protect the person reading our chart.

Design Has Two Jobs

Everything that follows is one of these two

Clarity

Help the eye find the point fast. Remove what does not carry meaning, and guide attention straight to the finding.

Honesty

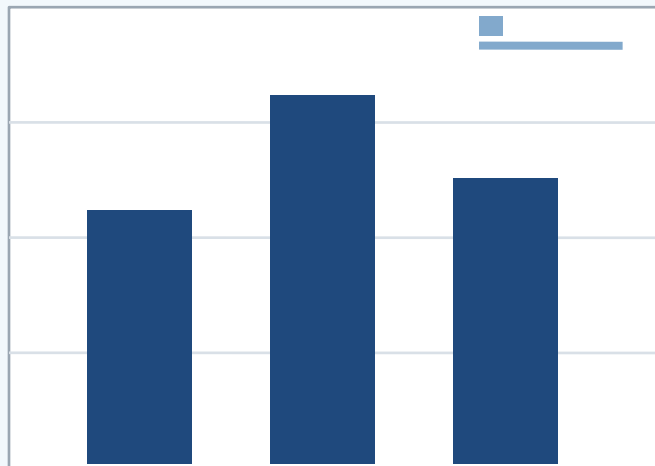
Never let the chart argue what the data does not. The baseline, the denominator, the distribution, the color.

Clear but dishonest persuades of the wrong thing. Honest but cluttered persuades of nothing. You need both.

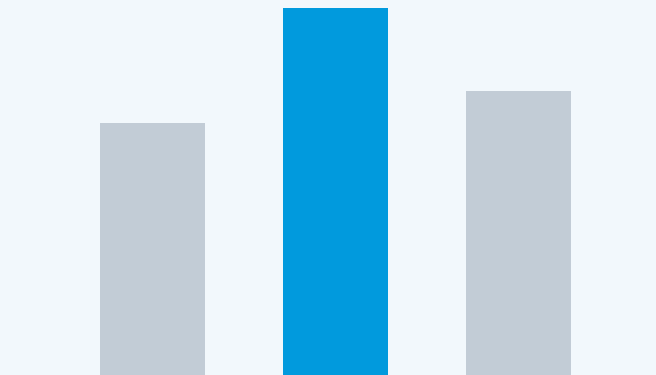
Clarity #1 - Remove Everything That Isn't Carrying Meaning

Every element the eye must process is a tax. Non-data ink is clutter.

Before



After

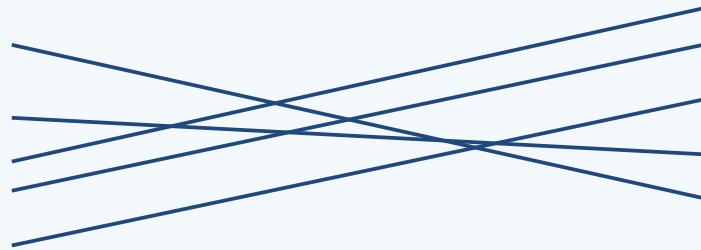


Cut gridlines, borders, and backgrounds. Kill the legend when direct labels will do. If it does not encode data, it competes with it.

Clarity #2 - The Eye Sees Before It Reads

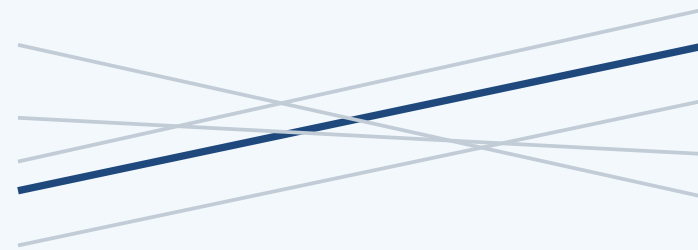
Color, position, and size register in milliseconds, before any label. Use that.

Before



everything highlighted

After



one line carries the story

One emphasis color. Mute the rest to grey. Direct-label the finding. If everything is highlighted, nothing is.

Every design choice makes something easier to see, and something else harder to see.

Clarity #3 - Let the Title Say the Finding

Title the chart with the takeaway it earned, not the variables it plots.

Not The Variable List

“Turnover rate by unit, 2025”

The Takeaway It Earned

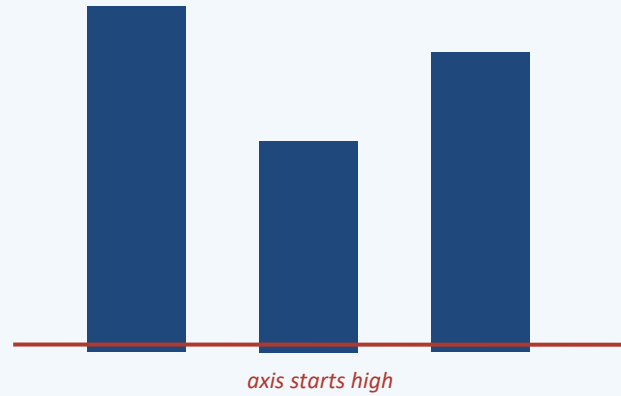
“Turnover is concentrated in three units”

Annotate the finding where it happens. An arrow and six words on the chart beat a paragraph below it. The reader should know why they are looking before they start.

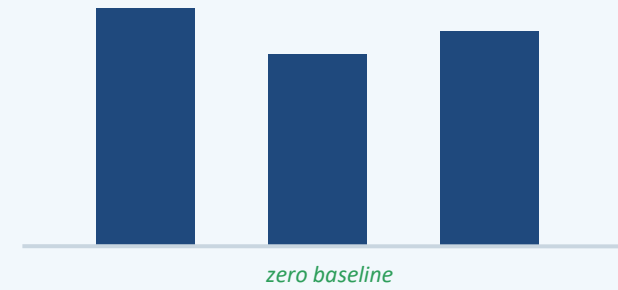
Honesty #1 - The Baseline and the Scale Make the First Claim

Before a single number is read, the axes have already argued something.

Misleads



Honest



Bars need a true zero, or length lies. Keep the aspect ratio neutral. One scale, never a dual axis.

Honesty #2 - Show the Denominator and Show the Spread

Two omissions that turn a true chart into a false impression.

1

Rate, not raw count

A count flatters the biggest group. Divide, and print the base, n = ...

2

The lonely average

An average alone hides its distribution. Two very different realities can share one mean.

3

Percent without its base

When groups differ in size, a share with no denominator misleads. Carry the n.

Every one of these is a callback. You met them as failures in Modules 1, 3, and 5. Design is where you stop repeating them.

Honesty #3 - Color Judges and Time Misleads

Two channels that argue on their own if you let them.

1

Color is an encoding, not decor

Red reads as bad, green as good, before any label. Do not pass a verdict you do not mean.

2

Match the ramp to the data

Sequential light-to-dark for magnitude. Diverging only when there is a true midpoint.

3

Design for color blindness

Roughly 8% of men cannot separate red from green. Vary lightness, never encode by color alone.

4

Separate actuals from projections

Dashed beyond today, shaded bands for ranges. Never let hope look like measurement.

Would Every Choice Survive Being Said Out Loud?

The closing test for honesty. State each design choice and its reason to a skeptic.

Survives

“Inverted axis, clearly labeled, because this audience thinks in splits”

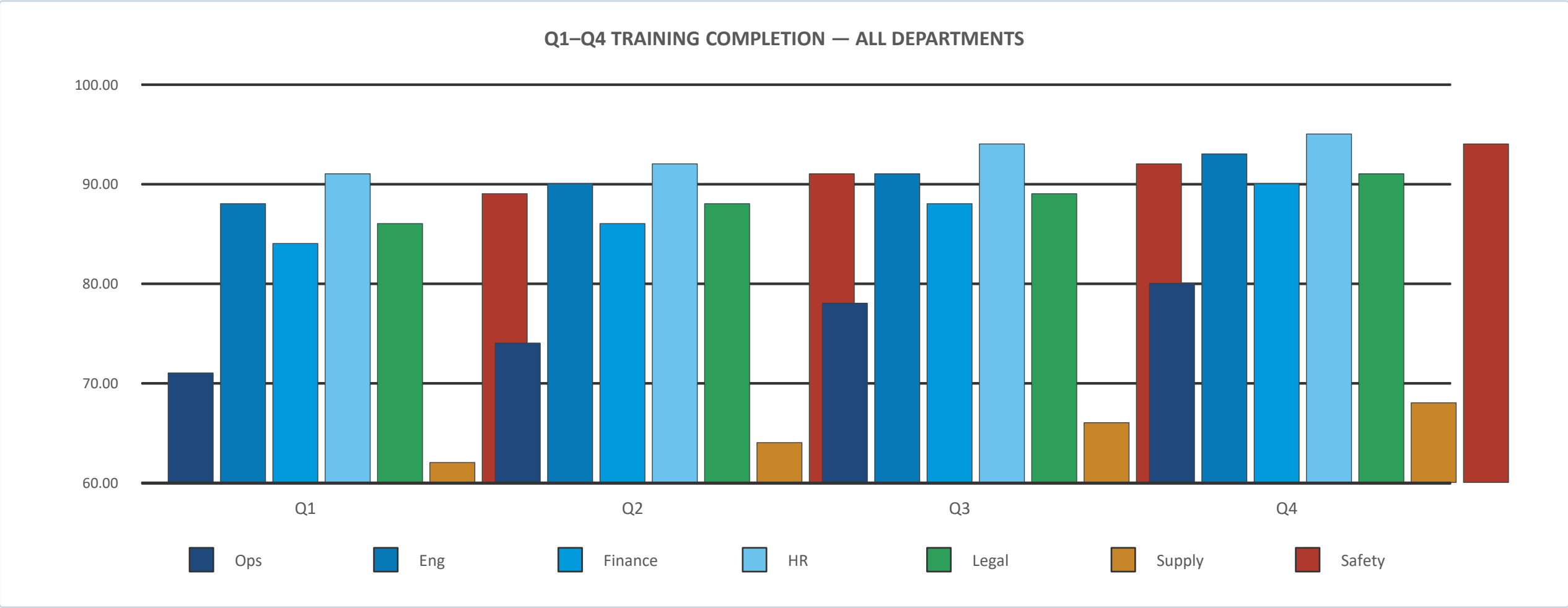
Does Not

“I started the axis at 40 to make the growth look bigger”

Design is not decoration. Clarity is a safety property

Worked Example - Before

Mandatory training completion. Every number is accurate.

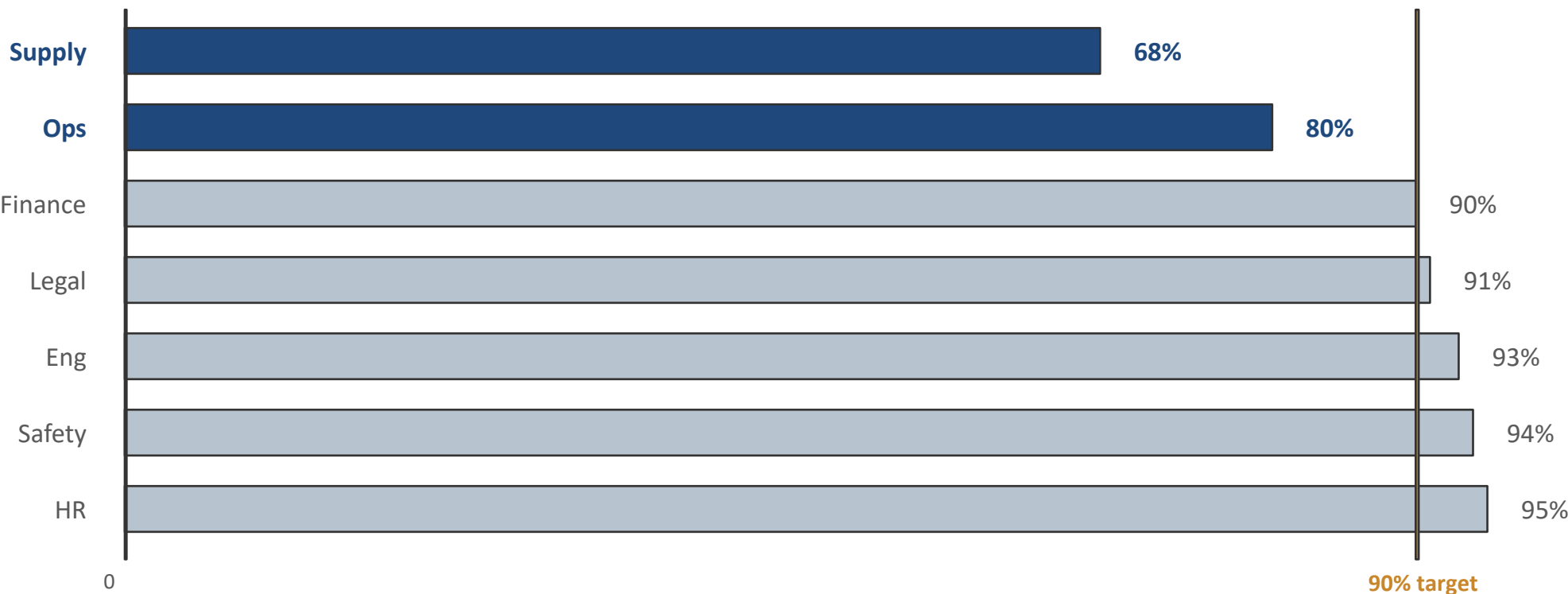


Worked Example - After

Same data but one question answered

Two departments are below the 90% training target

Q4 completion, all departments



Nothing was added. Three quarters, seven colors and a legend were removed, and the title now says the finding.

Worked Example - Which Moves Did That Use?

Clarity #1	Remove what is not carrying meaning	Dropped three quarters of history and the seven-color legend. The finding lives in Q4.
Clarity #2	The eye sees before it reads	Two departments in navy, the other five in grey. Attention goes where the problem is.
Clarity #3	Let the title say the finding	“Q1–Q4 Training Completion” became “Two departments are below the 90% target.”
Honesty #1	The baseline makes the first claim	The before chart started at 60, which exaggerated the gaps. The after starts at zero.
Honesty #2	Show the spread	All seven departments stayed in. Showing only the two would have hidden that five are fine.
Honesty #3	Color judges	Grey and navy, not red and green. The chart points; it does not scold.

Two questions, six moves. That is the whole module, and it is what the activity asks of you next.

Team Activity - Redesign the Dashboard (~20 minutes)

Take the cluttered readiness dashboard and fix one view for clarity and honesty

As a team, on one view, produce

- Three things you removed for clarity
- One emphasis choice: what the eye should hit first
- One honesty fix: baseline, denominator, distribution, or color
- A new title that states the finding

Present

The before, the after, and the one sentence your title now says.

That sentence is what carries into the recommendation and the story.

Module 8 Team Activity Debrief

Redesign the Dashboard

What did you remove, and why?

What should the eye hit first?

One emphasis choice.

Your honesty fix, and your new title.

The title states the finding.

Reflect

Which of your changes would alter what someone decides?

Module 8 - What We've Earned

You can show a finding clearly and honestly.

What's Next

Turning a finding into a decision

THE JOURNEY SO FAR

QUESTION

EVIDENCE

READ

EXPLORE

EXPLAIN

VISUALIZE

RECOMMEND

TELL

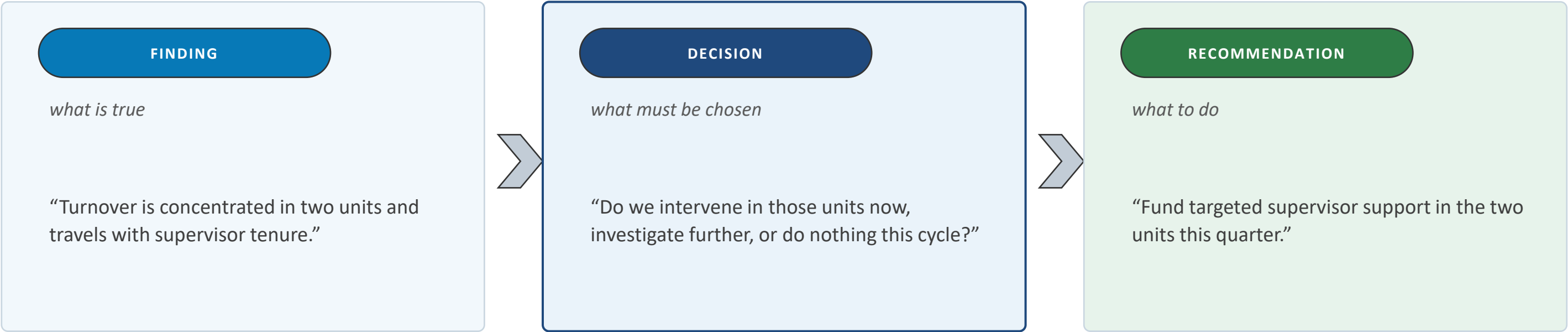
MODULE 9

Driving Decisions

Turn the finding into a decision someone can act on

A Finding Is Not a Recommendation

And there is a step between them that most analysis skips.



Judgment sits between all three. The data does not carry you from the first box to the last.

Skip the middle box	and you produce recommendations attached to no choice anyone is actually facing.
The decision is not yours to invent	Find the one already on somebody's desk, with a date on it. That is Module 2, arriving at the finish line.
Naming it changes the answer	“What should we do about turnover?” has no shape. “Intervene now, investigate, or wait?” can be answered.

The Recommendation Canvas

Eight fields. The first four build the case. The second four make it defensible.

BUILD THE CASE

Situation what is happening	Tension why it cannot wait	Evidence what supports it	Confidence how sure we are
---	--	---	--

MAKE IT DEFENSIBLE

Recommendation what to do	Consequences if we act, and if we do not	Assumptions what must be true for it to work	Limitations what this does not tell us
---	--	--	--

Assumptions are the bridge between evidence and action. “If we do X, we assume Y will happen” is rarely in the original analysis, and it is what the recommendation rests on.

Advocate, Within the Evidence

A recommendation must be proportioned to what you actually know.

No stronger, no broader, and no more irreversible than the evidence allows

NO STRONGER

How assertively should we act?

Moderate evidence earns a moderate recommendation, not a mandate.

“We recommend” not “the data requires”

NO BROADER

How widely should we act?

Evidence from two units justifies acting in two units, not company-wide.

Two units, not the whole organization

NO MORE IRREVERSIBLE

How much should we commit?

Uncertain evidence earns a reversible, testable move.

A quarter of supervisor support, not a restructure

Advocacy and honesty are not opposites. You are allowed to argue for something. You are not allowed to argue for more than you can support.

Say the Canvas Out Loud

The question in front of you is whether we intervene in those two units now, investigate further, or do nothing this cycle.

Situation	<i>“Workforce readiness has slipped for four quarters and turnover is up from 12% to 16%.”</i>
Tension	<i>“The budget decision lands this quarter, and it funds two units or a small step across everyone.”</i>
Evidence + Confidence	<i>“Turnover is concentrated in two units and tracks supervisor tenure, not pay. Three sources agree. Moderate confidence.”</i>
Recommendation	<i>“Fund targeted supervisor support in those two units this quarter.”</i>
Consequences	<i>“We learn within two quarters at a fraction of the cost. Spreading it thin helps nobody measurably.”</i>
Assumptions + Limitations	<i>“We are assuming tenure stands in for quality. We cannot say supervision causes it, and we have not checked who gets assigned where.”</i>

Confidence, assumptions and limitations get woven in as you go. Nothing is left out, and nothing new has to be memorized.

When Leadership Disagrees

First ask what kind of disagreement this is. Not all of them are about the data.

	WHAT WE DISAGREE ABOUT	WHAT IT SOUNDS LIKE	WHAT RESOLVES IT
EVIDENCE	We disagree about what is true	<i>"I do not think the data says that."</i>	More or better evidence can settle it.
ASSUMPTIONS	We disagree about what will happen	<i>"Coaching will not change anything."</i>	Name the assumption and design a test.
VALUES	We disagree about what matters	<i>"Retention matters less than the delivery date."</i>	No amount of data resolves this. It is a leadership call.
CONSTRAINTS	We disagree about what is feasible	<i>"We cannot release supervisors for training."</i>	Not a dispute about your analysis. It changes the options.

End With What It Does Not Tell You

Two habits that separate a trustworthy recommendation from a confident-sounding one.

THE LANGUAGE	THE LIMITATIONS
“The data says so”	“We recommend, because”
“We proved”	“The evidence suggests”
“This shows”	“At moderate confidence, this supports”
“Obviously”	“The strongest reading is”
NAME THE LIMITATION OUT LOUD, LAST “This says nothing about the units we did not segment. Confidence is moderate, not certain. And we have not checked whether the newest supervisors were assigned to the hardest units, which would change the answer.”	
A leader who hears the limitation from you trusts the rest. A leader who finds it later trusts none of it.	

Worked Example: The Program Question, Decided

The Canvas

Decision	Expand the program next year, hold, or redirect the budget.
Situation	We have run it for two years and expansion is on the table.
Tension	The budget decision lands this quarter.
Evidence	Graduates' teams rate 4.1 against 3.7. But graduates rated 3.9 before attending.
Confidence	Low to moderate.
Recommendation	Do not expand yet. Run one cohort with managers assigned.
Consequences	Two quarters and a fraction of the cost tells us. Expanding now scales a selection effect.
Assumptions	An assigned cohort is comparable, and two quarters is long enough.
Limitations	Ratings miss what the program may add that a score never measures.

The Same Thing, Said In Two Minutes

“The question in front of you is whether we expand the leadership program next year, hold it, or redirect the budget. It has run for two years and the decision lands this quarter.

Graduates do lead better-rated teams, 4.1 against 3.7. But when we checked, graduates rated 3.9 before they ever attended. Most of the gap was already there. My confidence is low to moderate.

So I recommend we hold, and run one cohort with managers assigned rather than volunteering. Two quarters and a fraction of the cost tells us whether to expand. Expanding now scales a selection effect at full budget.

We are assuming an assigned cohort is comparable. And ratings may not capture everything the program adds.”

Same content, same order. The canvas is how you build it. Reading it aloud is how you deliver it.

What We Have Earned, and What Is Left

Module 9 makes a recommendation defensible. Module 10 makes it land.

MODULE 9

MODULE 10

The question

What should we do?

How do we make it land?

The work

Evidence plus judgment

Meaning plus relevance

The tool

The Recommendation Canvas

The Data Story Arc

The standard

Defensible

Memorable

The failure to avoid

Outrunning the evidence

Distorting the evidence

You can now recommend an action proportioned to the evidence and the constraints.

Team Activity - The Two-Minute Brief (~20 minutes)

As a team

- Complete the Recommendation Canvas for your finding
- Deliver it in two minutes using Context, Tension, Evidence, Recommendation, Consequence
- State the confidence honestly, and name one limitation

The room checks

Is it a recommendation, not just a finding?

Is the confidence honest?

Is there a named limitation?

If yes to all three, it is ready for a real leader.

When Is Enough Evidence Enough?

There is no percentage. There is a rule, and it is not about how sure you feel.

Two Ways to Get This Wrong

ACTING TOO SOON

Committing before the evidence supports it

One metric, confidently read

Nobody named what would prove it wrong

The move is large and hard to reverse

Visible. Someone owns it. It gets reviewed.

NEVER DECIDING

Gathering because gathering feels safe

One more cut, then we will know

Waiting on evidence that will not arrive in time

The decision gets made by default, by the calendar

Invisible. Nobody owns it. It is never reviewed.

Research on decision avoidance: postponing, delegating or bypassing a choice reduces felt responsibility without reducing the consequences.

The Feeling of Not Knowing Enough Is Not a Signal

People gather more information when they feel less confident, even when their accuracy is identical

What that means

The urge to keep looking tracks how you feel, not how much you actually know. It is a mood, not a measurement.

And confidence is a poor guide

Meta-analytic work on clinical judgment finds only a small relationship between how confident a judge is and how accurate they are.

More information can help

Additional evidence that is independent, diagnostic and decision-relevant does improve accuracy.

But most added information is not

Redundant data, selectively searched data, or data read after a preferred answer has formed raises confidence without raising accuracy.

So the test cannot be “do I feel ready?” It has to be something outside your own head

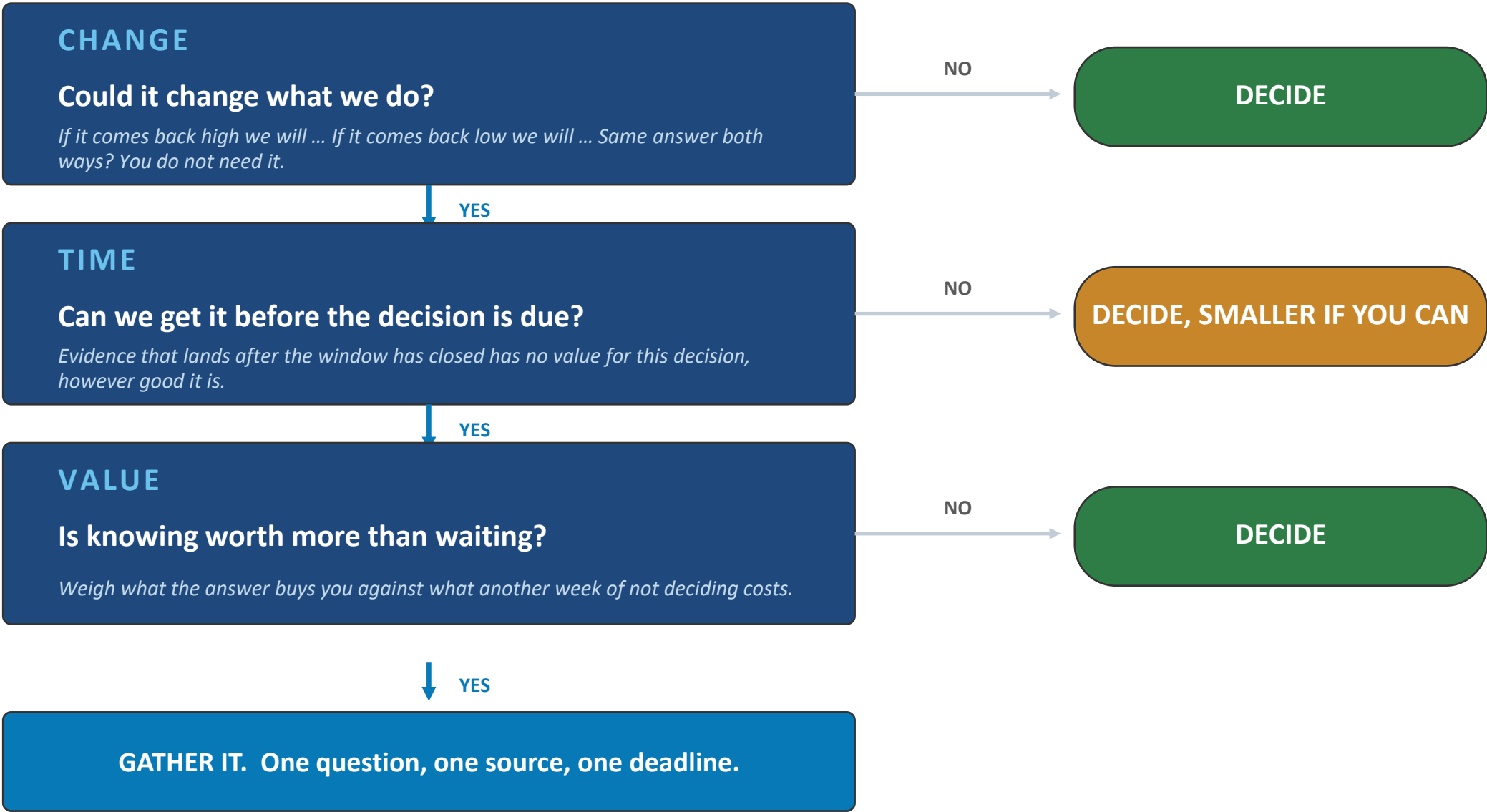
The Deadline Is Not Arbitrary, Work Backward

The evidence window is what is left after implementation takes what it needs

Latest date the action still delivers the benefit	The start of the hiring season, the budget lock, the deployment, the point after which it no longer helps
Minus implementation time	Approvals, procurement, configuration, training, communications, mobilisation
Minus contingency	Because implementation runs long, and it always has
= your evidence window	Everything left. Not a week you chose because it felt about right
ARMY PLANNING DOCTRINE, ONE THIRD AND TWO THIRDS Commanders use no more than a third of available time for their own planning, leaving two thirds for the people who have to prepare and execute. Analysis that overruns is not thoroughness. It is taking time from the people below you.	

Before You Ask for More Data

Three questions. Any no, and more data should not delay this decision.



Uncomfortable at any of the three exits? Reduce the commitment, not the analysis.

You Cannot Answer VALUE Without This

What does one more week of not deciding actually cost?

Capacity Vacancies open, coverage gaps carried another week	Premium Overtime, contractors, agency cover, manager time	People The at-risk group, times the weekly chance of exit	Benefit delayed Every week it is not live is a week it is not helping	Exposure Compliance, safety or legacy risk still being carried
---	---	---	---	--

Give a low, a base and a high figure. Never one number, and never zero.

You do not need precision. You need enough to answer one question:

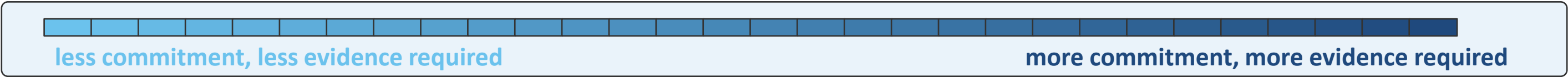
Is a week of waiting worth more, or less, than what the answer would buy us?

Delay always has a price. It is just never anybody's line item.

The Question Is Not Yes or No

Binary framing is what creates paralysis. If the only alternative to deciding is committing fully, waiting looks rational.

THE COMMITMENT	WHEN IT IS RIGHT	EVIDENCE IT NEEDS
Decide and monitor	Low exposure, reversible, fast feedback	<i>Adequate rationale and a rollback plan</i>
Test or pilot	A bounded action designed to learn	<i>Enough to justify the test, not proof of impact</i>
Safe-to-fail probes	Complex, where cause only becomes clear afterwards	<i>Only safety and plausibility. No predictive proof</i>
Stage the commitment	Proceed in steps, learning as you go	<i>Enough that withholding entirely would be wrong</i>
Full commitment	High exposure, or hard to reverse	<i>Strong evidence, alternatives considered, sign-off</i>



The real question: what is the smallest commitment that would safely answer the uncertainty that matters?

So, When Is Enough Enough?

You have enough evidence when more information is unlikely to change what you do, or when waiting for it costs more than acting with the uncertainty you have.

And if uncertainty remains, reduce the commitment rather than extending the analysis.

CHANGE

Could it change what we do?

If no, decide

TIME

Can we get it before the decision is due?

If no, decide smaller

VALUE

Is knowing worth more than waiting?

If no, decide

Enough evidence never meant enough to be certain. It means enough to make the next appropriate commitment.

HOW GOOD ANALYSIS GOES WRONG

How Good Analysis Goes Wrong

And the safeguards that make it harder.

Three Failures. Which One Was Caused by Bias?

Take a moment before you answer.

A-level grading, UK, 2020

A school's historical results were used to standardize individual students' grades. Around 36% of grades came out below teacher assessment, hitting able students in historically lower-attaining schools hardest.

Dutch childcare benefits

Risk profiling for fraud, with nationality reportedly used as a risk factor. At least 35,000 parents were wrongly accused.

Amazon recruiting tool

A model trained on a decade of hiring outcomes in male-dominated technical roles. It downgraded signals associated with women.

Most rooms answer: all three. That is true, and it is almost useless.

“Bias” names the outcome, not the mechanism. And you cannot fix a mechanism you have not identified.

Six Ways Analysis Goes Wrong And Only One Is Cognitive Bias

Different mechanisms need different controls. That is why the diagnosis matters.

THE SOURCE	WHAT IT SOUNDS LIKE	WHAT ACTUALLY CONTROLS IT
Cognitive Bias / Judgment	<i>“I favor evidence that supports what I already believe”</i>	Disconfirmation prompts, competing explanations
Assumption	<i>“I assume manager ratings mean the same thing across divisions”</i>	Surface it, then test only what could change the decision
Measurement	<i>“The metric systematically misses part of what we care about”</i>	Name the proxy and the gap. Triangulate with another method
Selection	<i>“The dataset contains only the people who survived the process”</i>	Ask who is missing before you analyze who is present
Incentives	<i>“The organization rewards a particular answer”</i>	Separate the analyst from the advocate. Reward calibration
Process	<i>“Nobody independently challenges the conclusion”</i>	Pre-commitment, replication, a real challenge function
Sometimes nobody is biased. The measurement system is. Telling those apart is what makes the fix findable.		

So Why Doesn't Bias Training Work?

Because five of those six sources are not in anyone's head.

The bias blind spot: people reliably see bias in others more readily than in themselves.

So what happens in the room

People leave able to name the bias their colleague has. Recognition of one's own is the part that does not arrive.

Does teaching about bias reduce it?

A systematic review found abstract knowledge insufficient on its own. Evidence of transfer into real decisions is scarce: only a handful of studies tested retention, and essentially one tested transfer beyond the training setting.

What does have some support

Practice on realistic cases with feedback beats a lecture. Even there, durable transfer into daily work is uncertain.

The conclusion worth drawing

Do not rely on anyone noticing their own bias at the moment it matters. Build steps that do not require it.

This is not an argument that bias does not matter. It is an argument about where to spend the effort.

How to Find the Assumptions You Did Not Know You Made

Four places to look. Run the claim through all four.

The claim: “Turnover is higher because of poor supervision.”

MEASURE

What are we assuming this measure represents?

Supervisor tenure stands in for supervisory quality. Those are not the same thing.

PEOPLE

What are we assuming about who is represented?

Which leavers are in this? Who never answered, and who left before the window opened?

COMPARISON

What are we assuming is comparable?

Are newer supervisors concentrated in the hardest units? Then we are comparing units, not supervisors.

CAUSE

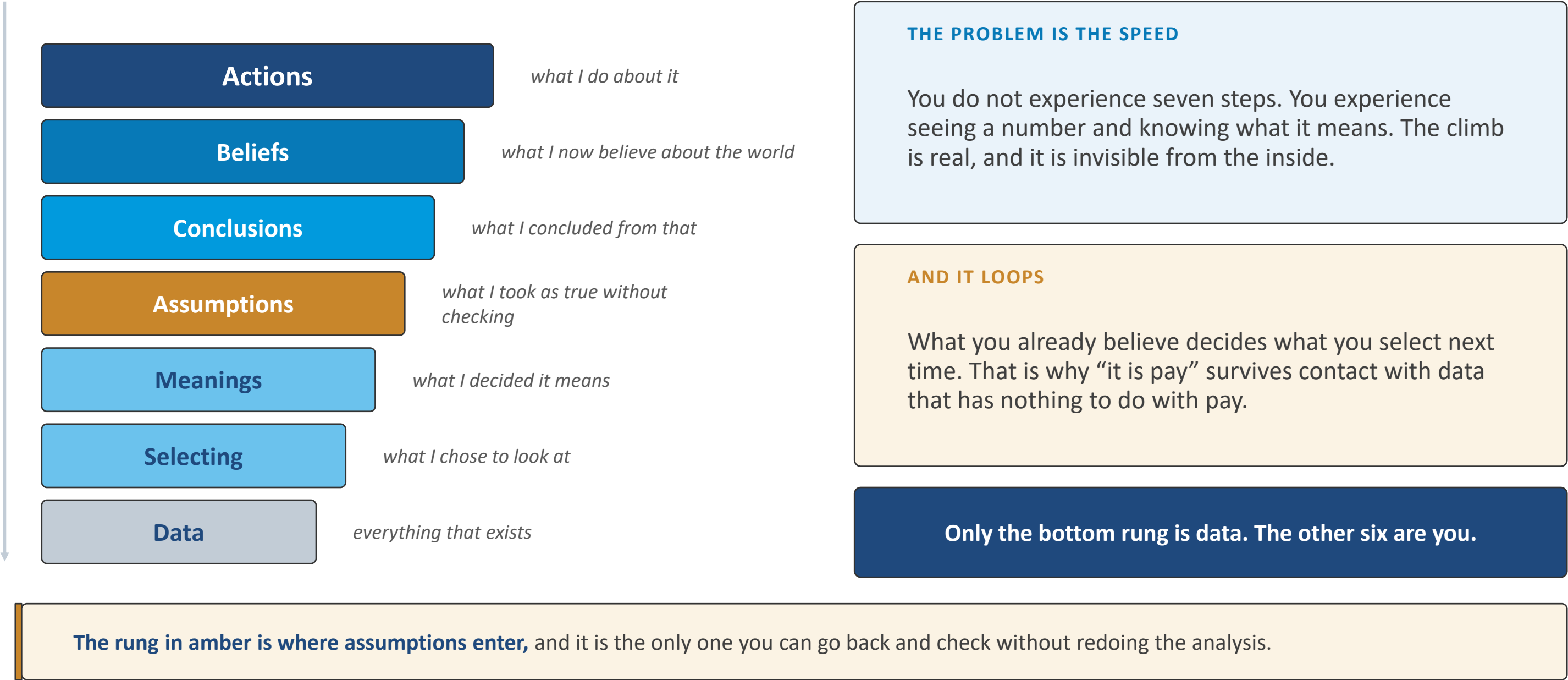
What are we assuming explains the pattern?

Turnover travels with supervisor tenure. Travelling with is not causing.

If you haven’t asked all four, you probably have assumptions you have not surfaced yet.

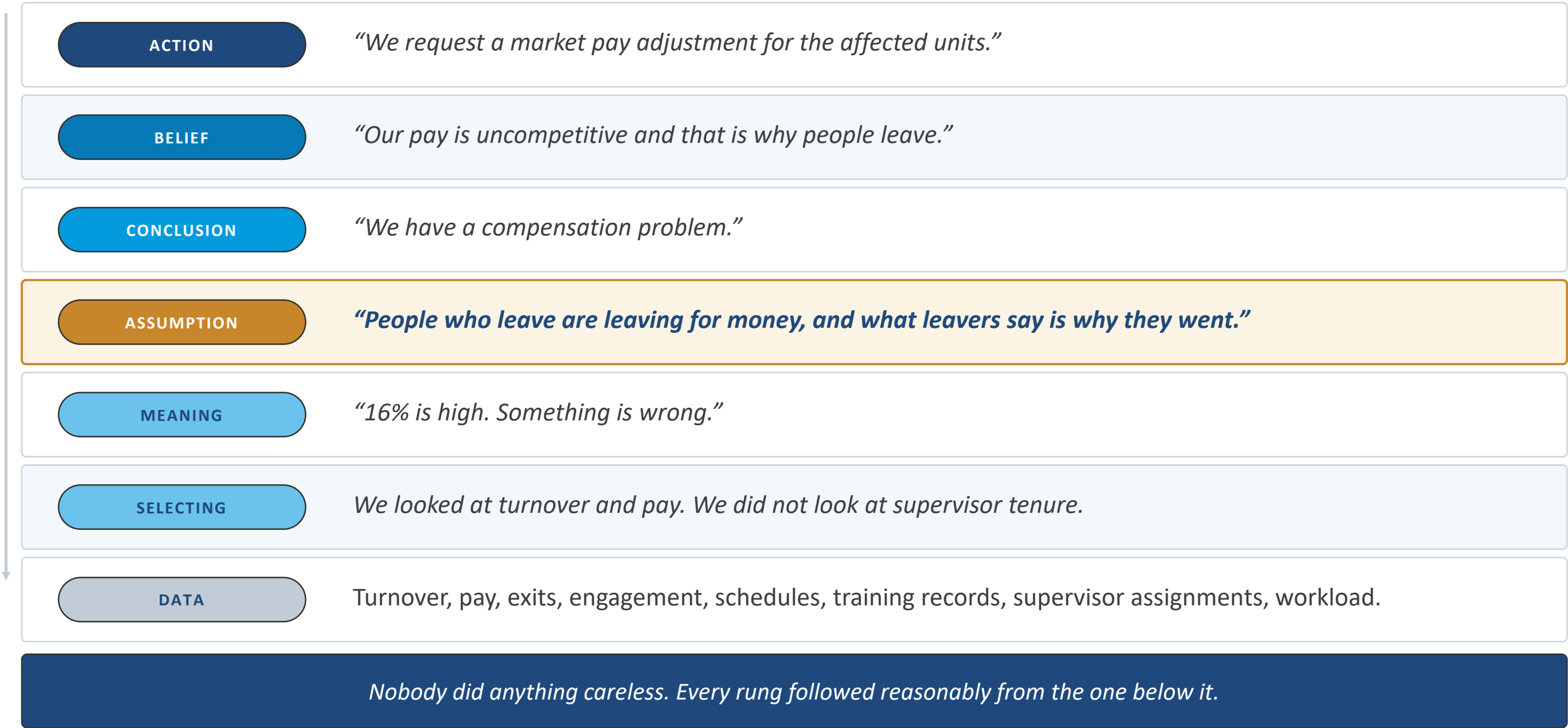
The Ladder of Inference

You climb seven rungs between a number and an action, and you do it in about a second.



The Climb, on Our Own Case

This is what happened in this room on the first morning. Read it from the bottom up.



Walking Back Down

One question per rung. Start at what you are about to do, and descend.

ASK YOURSELF	OUR ANSWER	WHY IT WORKS
1 What am I about to do?	<i>Request a pay adjustment</i>	Start at the action, because that is the only rung you can see clearly.
2 What do I believe that makes this obvious?	<i>Our pay is uncompetitive</i>	Say the belief out loud. It usually sounds less certain spoken than it felt.
3 What did I conclude, and from what?	<i>A compensation problem, from a turnover number</i>	Name the evidence you actually used, not the evidence that exists.
4 What did I take as true without checking?	<i>That leavers leave for money</i>	This is the rung you came for. Everything above it rests here.
5 What meaning did I add?	<i>That 16% is high</i>	Compared to what? A number is not high on its own.
6 What did I look at, and what did I skip?	<i>Turnover and pay. Not supervisor tenure</i>	The skipped data is usually still sitting there.

Six questions, about ninety seconds, and no new data required. Everything you need is already in your own reasoning.

Not Every Assumption Deserves an Investigation

Otherwise assumption-hunting becomes its own kind of paralysis.

NECESSARY?	Would our conclusion change if this turned out to be false?	<i>If no, note it and move on. It is background, not risk.</i>
SUPPORTED?	What evidence do we actually have that it is true?	<i>Untested and load-bearing is the combination that hurts.</i>
TESTABLE?	What could we check that would move our confidence either way?	<i>If nothing could, say so. That is a limitation, not a task.</i>
Could change the decision, and support is weak		Cannot change the decision
Test it. This is where the effort belongs.		Write it down and move on. Do not open a rabbit hole.
The question is never how many assumptions we have. It is which ones this decision actually depends on.		

You Have Been Surfacing Assumptions All Along

Every module asked a version of the same question.

Module 2	Frame better questions	<i>What are we assuming in the question itself?</i>
Module 3	Data as evidence	<i>What are we assuming this evidence represents?</i>
Module 4	Reading data correctly	<i>What are we assuming about the measure?</i>
Module 5	Exploring the data	<i>What are we assuming about the population and the comparison?</i>
Module 6	From pattern to cause	<i>What are we assuming explains the pattern?</i>
Module 9	Driving decisions	<i>What are we assuming will happen if we act?</i>

This module is not new material. It is the name for what the whole course has been doing.

What Actually Works, and How Well

One relies on people catching themselves. The other changes what the process requires.

ASKING PEOPLE TO THINK BETTER

Awareness training
Helps people recognize and name biases

Warnings and prompts
Reminds people to slow down or reconsider

Consider the opposite
Forces attention toward contrary evidence

Practice with feedback
Builds skill through repeated application

Trying harder
Relies on people catching themselves in the moment

CHANGING WHAT THE PROCESS REQUIRES

Pre-commit the analysis plan
Choices are visible before results are known

Blind what should not matter
Irrelevant information cannot influence judgment

Independent replication
Someone else has to reach the same result

Competing explanations
The preferred story has to beat alternatives

Outside view and comparable cases
Compare against what usually happens

The right column works because it does not require anyone to notice their own bias at the moment it matters

What an Organization Can Do That a Person Cannot

Three of the six sources are structural. No individual can fix those alone.

Separate the analyst from the advocate	Analysts report method, uncertainty and alternatives. Someone else owns the recommendation.	<i>Strong governance logic. Needs leaders who tolerate inconvenient findings.</i>
Standing challenge function	A red team with access, standing and a written response required from the decision owner.	<i>Works only with authority and protected dissent. Otherwise it is theater.</i>
Decision journal and review	Record forecasts, assumptions, confidence and alternatives. Review after outcomes.	<i>Good for calibration. Value depends on reviewing process, not punishing bad luck.</i>
Reward accuracy, not confidence	Recognize calibration, honest uncertainty and correction. Not speed or the desired finding.	<i>High face validity. Direct trials in bureaucracies are limited.</i>
Independent audit by group	Outcomes, data quality, overrides and drift, monitored over time.	<i>Particularly important for public-sector systems.</i>
No-blame learning review	Judge whether the process was sound given what was known at the time.	<i>A blame culture predictably produces selective evidence and buried uncertainty.</i>
The precondition for all six: if bad news is punished, every one of these becomes paperwork and the real analysis happens in the corridor.		

Dashboards and Models Do Not Remove Any of This

They move it somewhere harder to see, and lend it authority.

Dashboards decide what is visible

A default KPI set makes some outcomes visible and others invisible. Time-to-hire rewards speed and conceals pass-through rates by group, adjustment delays, and quality of appointment.

Defaults do the reasoning

Filters, aggregation, chosen denominators and visual salience direct attention before anyone consciously reasons about anything.

Models inherit the past

A system trained on prior hiring or performance outcomes learns historic preference, not job-relevant merit. That is a selection problem, not a psychological one.

Automation bias

People use system output as a substitute for looking. Erroneous recommendations get followed, and omissions go unnoticed, especially when the tool is presented as objective.

THE MINIMUM CONTROLS FOR ANY HR TOOL

State the purpose, population and fairness definition before you buy or build it · Test error rates by group, not just overall · Audit inputs, proxies, missing data and overrides, not only outputs · Give reviewers real authority and enough time · Monitor drift after deployment

A fairness score is not governance. The whole decision process has to be explainable, challengeable and auditable.

Do Not Try to Be Less Biased

Build a process where it is harder to go wrong quietly.

Make the assumptions visible. Make the preferred explanation face alternatives. Commit the choices before you look. **And let someone else check the ones that matter.**



Before you look

Write the question, the population, the outcome and what would change your mind.



While you work

Log the choices. Label exploration as exploration.



Before you conclude

Run MEASURE, PEOPLE, COMPARISON, CAUSE. Put the strongest alternative next to your answer.



Before you act

Name the assumptions the decision depends on, and who checked them.



After the outcome

Review whether the process was sound, given what was known at the time.

Data literacy is not knowing how to eliminate bias, uncertainty or judgment. **It is building a decision process that stays defensible despite them.**

AI as a Thought Partner to Mitigate Bias

AI for Pattern Checking

Highlight patterns humans may miss due to availability or confirmation bias.

AI for Diverse Perspectives

Compare how AI frames insights differently than your team might.

AI for Counterfactuals

Ask “what if” scenarios to test assumptions and reveal blind spots.

AI as a Consistency Checker

Validate calculations, logic, or statistical methods to avoid simple cognitive traps.

AI as a “Second Pair of Eyes”

Use AI to surface alternative interpretations of the same data.

AI for Red Teaming

Generate critiques of your conclusion to simulate dissenting views when groupthink pressures are high.



AI As a Thought Partner

Activity: Walk Down Your Own Ladder (~15 minutes)

Not the readiness case. Something you are actually carrying at work.

ON YOUR OWN · 6 MINUTES

- 1 Write down one thing you currently believe about your workforce, your team, or a program. Something you would act on.
- 2 Descend the six questions. Write the answers, even the ones that feel obvious.
- 3 Stop at the assumption rung and write that one out in full.

IN PAIRS · 8 MINUTES

- 1 Four minutes each. One person descends out loud, the other asks the six questions.
- 2 Your only job as the partner is to keep asking. Do not solve it, do not reassure, do not agree.
- 3 When they reach an assumption, ask once more: what would have to be true for that to hold?

THE SIX QUESTIONS, IN ORDER

1. What am I about to do?
2. What do I believe that makes this obvious?
3. What did I conclude, and from what?
4. **What did I take as true without checking?**
5. What meaning did I add?
6. What did I look at, and what did I skip?

You need no data for this. Everything required is already in your own reasoning.

Debrief

One question. Everything else stays in the pairs.

What assumption did you find, and had you ever written it down before?

Almost nobody has written it down

That is the finding. The assumption was load-bearing, it was doing work every day, and it existed nowhere except in somebody's head.

Notice which rung was hardest

Most people stall between belief and conclusion. Saying the belief out loud is uncomfortable because it sounds thinner spoken than it felt.

Notice what you skipped

Question six almost always turns up data that exists and was never selected. It is usually still sitting there.

And notice you did not need permission

No new analysis, no request to anybody, no tool. Six questions and ninety seconds.

Take that assumption back with you. Write it into the next brief you produce, under Assumptions, where somebody else can challenge it.

MODULE 10

Telling the Story

Make the recommendation land, stick, and spread

MODULE 10

Communicating with Data

Make the recommendation land without changing the truth.



<https://dataintowisdom.com/EMBL/>

What Just Happened

Bounded rationality & sensemaking

Your readiness to decide moved between Round 1 and Round 2.

Simon, 1956; Weick, 1995

Cognitive load theory

Working memory holds 3-4 chunks. Round 1 exceeded that capacity; the second worked within it.

Sweller, 1988

Choice architecture

When the speaker designs for the decision and places the ask deliberately, the audience can detect it.

Thaler & Sunstein, 2008

Narrative encoding

Round 1 read as a list of facts. Round 2 read as a coherent argument. Two modes of cognition.

Bruner, 1990; Hasson et al., 2010

The signal that mattered most was your readiness to decide

Why Is This So Hard?

Three Disciplines Must Converge (in one person)

The disciplines are taught separately when they are taught at all. Most professionals receive formal training in only one

Analytical thinking

The work of producing trustworthy findings

**Thoroughness, Hedging,
Disclosure**

Audience thinking

The work of reading who you're talking to

Empathy, Flexibility, Adaptation

Communication design

The work of making the finding land

**Selection, Compression,
Confidence**

The communicator must hold conflicting epistemologies in one act

Rigor wants thoroughness; design wants selection

Audience-reading wants flexibility; analysis wants consistency

Method wants disclosure; communication wants confidence

Why Does It Persist?

Communication Fails In Four Places At Once

Four patterns that all reinforce each other

The Communicator

The Analyst’s Curse

Days deep in the data
Rich mental model built from that immersion
Forgot what they had to figure out
Audience starts where they started days ago
The gap is invisible to them

The Translator’s Distortion

Communicating findings they didn’t produce
Headline known. What’s behind it, not
Hedges smoothed into performative confidence
Their own frame imported into the retelling
Data sometimes used as alibi for a prior choice

The Audience

The Receiver’s Filter

Arriving with depleted attention
Uncomfortable with numerical claims
Confirmation pull on prior beliefs

Defensive "what about" reflexes
Trust attribution by source, not content

The System

The System Around Them

Tools push toward maximalism
Incentives reward analysis, not communication
Education didn’t cover it

No feedback loop on communication quality
Career risk is asymmetric. Over-disclose feels safe

How Do We Improve?

The fix is not motivational

It is architectural

We don't ask people to become better integrators of three converging disciplines

We move the integration out of the individual mind and into a visible structure

Workforce Services — Success Metrics for 1H

Somebody sent this to leadership last week. What is wrong with it?

Hiring actions processed

1,240 requisitions completed across all directorates

Recruiting reach

18 job fairs attended, 3,250 candidate contacts logged

Training delivery

24,000 training hours delivered, up 11% on last year

Onboarding

410 new employees in-processed

Records management

9 systems maintained, 96% data completeness

Case handling

2,180 employee relations cases opened and closed

Lacks context

Data without interpretation

No emotional or personal connection

Does not engage the audience

Focuses on outputs, not outcomes

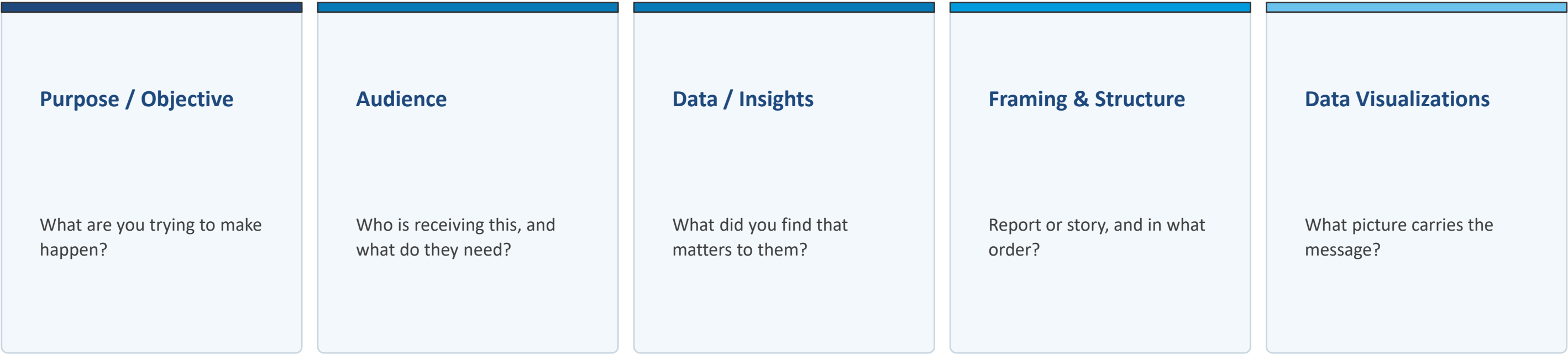
No clear “so what” factor

Every number is true. Not one of them tells anybody what to do.

This is activity reported as achievement. It answers “what did we do?” when the reader is asking “is anything better?”

Communicating with Data — Key Components

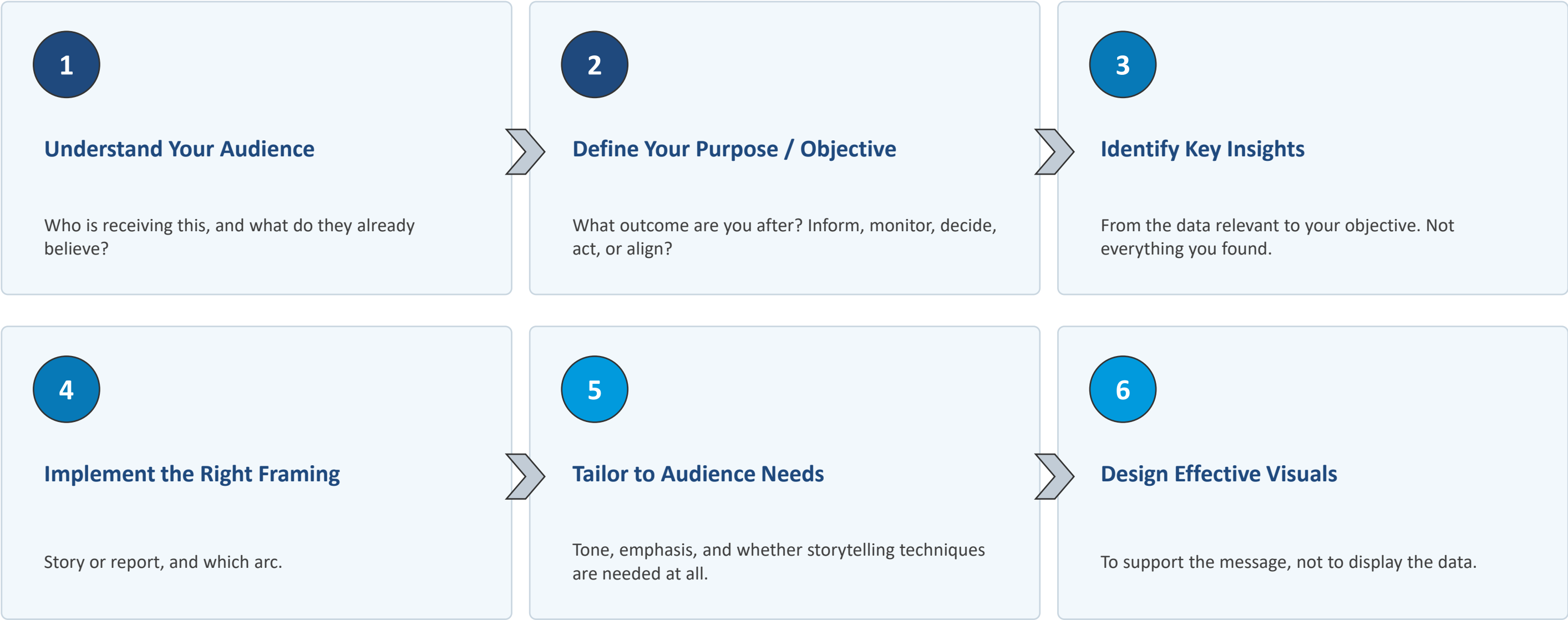
Five things every data communication has, whether you chose them deliberately or not.



Most people start at the visualization. It is the last of the five, and it is the one that matters least if the first two are wrong.

The Communicating with Data Process

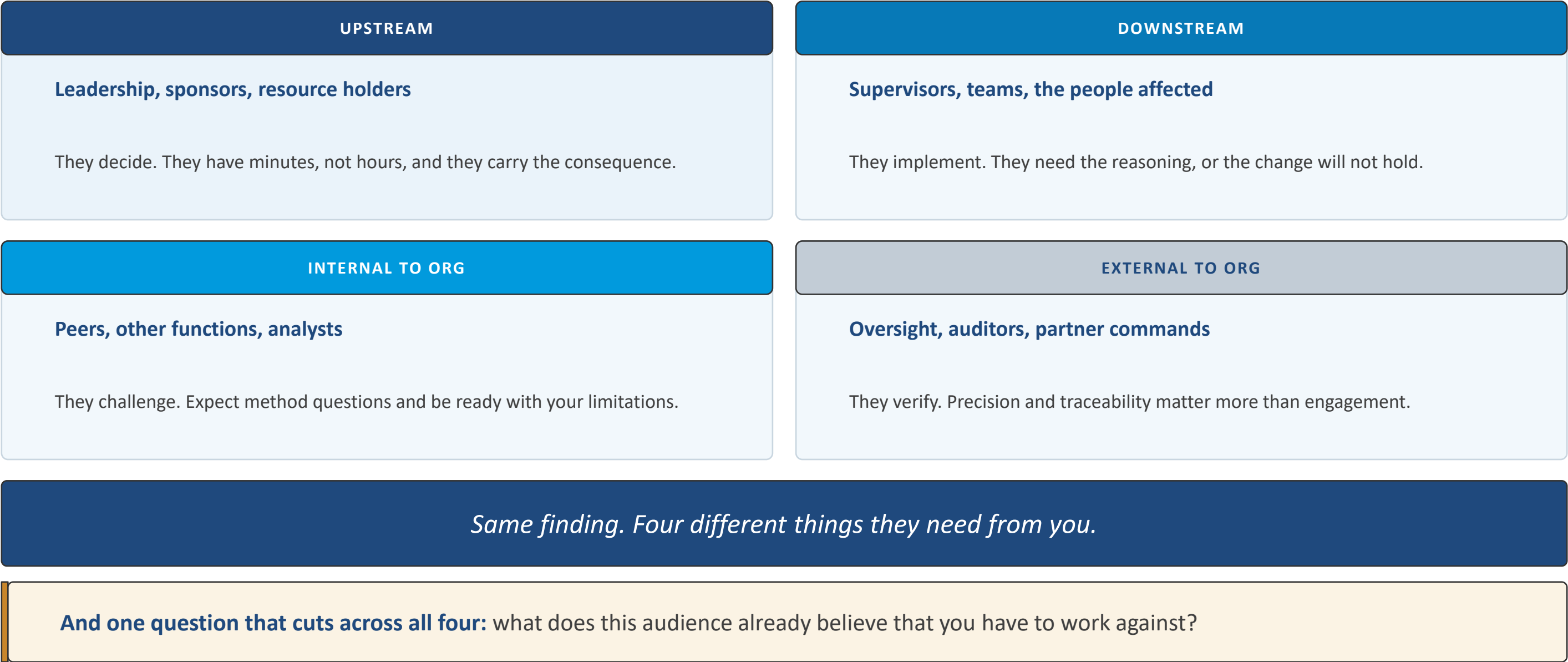
Six steps. Everything in this module hangs on this sequence.



Steps 1 and 2 decide everything after them. If you cannot answer them, do not start building slides.

Step 1: Understand Your Audience

Four directions, four different conversations.



Questions for Each Stakeholder Group

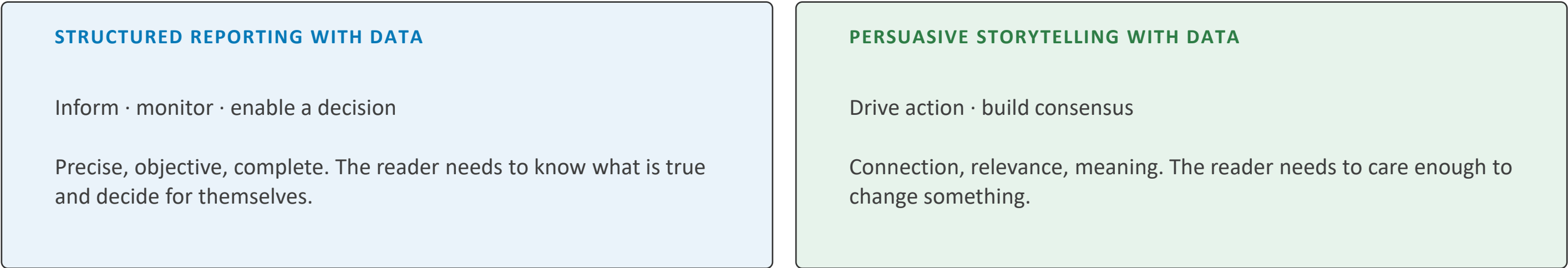
Four questions to ask before you write anything.

	QUESTIONS	IMPLICATIONS
Our purpose / goal	What is our purpose for the communication? What is the key decision and insight you want to share?	<i>Decide the outcome before the format. Everything else follows from it.</i>
Stakeholder's role	What decisions do they make? What questions do they need answered? What matters to them?	<i>Structure the communication to make it easy and direct to answer their high-priority questions.</i>
Their data comfort and skills	How knowledgeable are they with using data? Do they enjoy digging into the numbers?	<i>Match the data and the visuals to their capabilities.</i>
Their industry and data expertise	How familiar are they with the specific metrics and terminology?	<i>Embed explanations and definitions as needed to match their knowledge.</i>

Row one is about you. The other three are about them. Most communications get built having answered only row one.

Step 2: Define Your Purpose / Objective

What are the specific outcomes or goals?



When defining your objective, ask: do I need to inform, or do I need to persuade?

Step 3: Identify the Right Insights

From data overload to a focused narrative

THE QUESTION		THE PURPOSE	
F	Focus	What decision or action am I trying to drive?	Define the narrative purpose upfront
O	Outcome-Relevance	Which insights directly impact this goal?	Filter for insights that are actionable and outcome-relevant
C	Critical Data	What minimal data is required to support the insight?	Limit to must-know data that makes the insight credible
U	Usefulness Check	Does this data or insight help the audience act or decide?	Exclude data that does not change the audience's understanding or decision
S	Simplicity	Can I explain this clearly in one to two minutes?	Prioritize clarity and brevity to avoid cognitive overload

Step 4: Implement the Right Framing

Apply the structure and techniques that match your chosen approach.

STRUCTURED REPORTING WITH DATA

Focus on precision, clarity, and objectivity

Present facts, insights, and recommendations directly

Use logical, linear arcs: Situation → Data → Conclusion

Emphasize accuracy and completeness

PERSUASIVE STORYTELLING WITH DATA

Focus on connection, engagement, and persuasion

Frame insights within a narrative arc to create meaning and drive action

Use human-centered arcs: Challenge → Change → Impact

Emphasize relevance, emotion, and simplicity

Communicating with Data, and Storytelling

Two different jobs. Each covers the other's weakness.

COMMUNICATING WITH DATA

Precision

Focuses on accuracy and clarity

Logical

Engages the analytical part of the brain

Objective

Helps make informed, fact-based decisions

STORYTELLING

Engagement

Taps into emotion for lasting impact

Context

Provides meaning behind the numbers

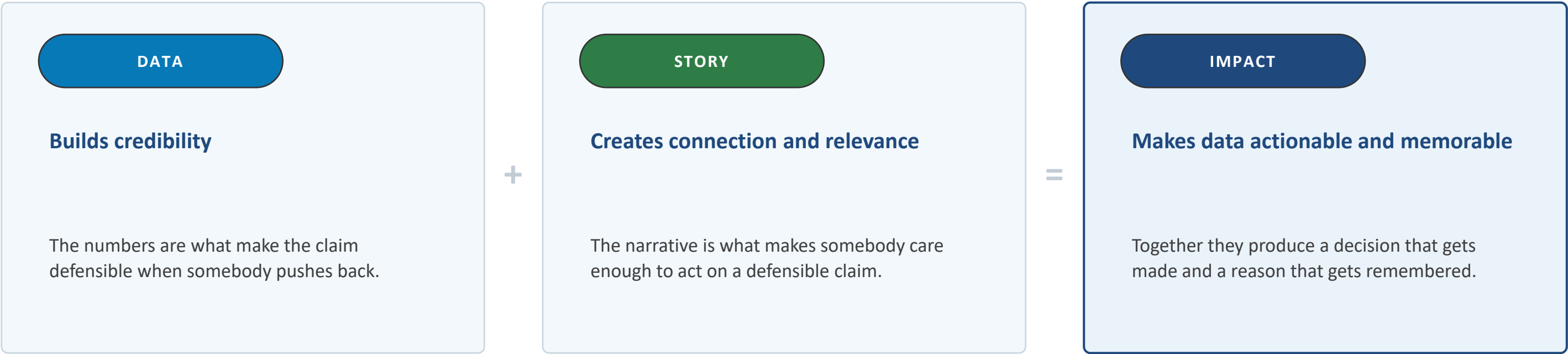
Memorable

People recall the story long after the numbers

Data without story does not move anybody. Story without data is just an opinion delivered well.

The Power of Combining Both

When data meets story, insights become actionable and decisions become transformative.



Credibility gets you heard. Connection gets you acted on. You need both.

Structure Your Message With a Narrative Arc

Seven stages. You will not use all of them every time.

Hook	Grabs attention by highlighting a compelling insight from the data analysis
Context	Provides necessary background to frame the issue and establish its significance
Process	For sophisticated analyses, walks through the methodological detail to give transparency into the rigor behind the insights
Insight	Distills and reveals the key data-driven findings, conclusions and implications from the work
Impact	Discusses the broader relevance of the insights with respect to the audience's interests, goals and concerns
Resolution	Outlines proposed solutions, recommendations and next steps implied by the insights uncovered
Future	Concludes by projecting the possibilities ahead and inspiring further progress fueled by the insights

Challenge, Change, Impact is the arc. These seven stages are how you build those three beats.

Determine Arc Format

Same arc. Different entry point, depending on who is listening and how long you have.

STANDARD

Introduction → Development → Key Insights → Implications and Call to Action

WHEN TO USE IT

Use when you need a comprehensive narrative that gradually builds understanding and concludes with a strong point or action.

INVERTED

Start with the Recommendation → Key Insights → Supporting Details → Conclusion

WHEN TO USE IT

Use when the audience is expert or time-constrained and benefits from immediate clarity on the outcome or action needed.

TRUNCATED

Key Insights → Supporting Details → Quick wrap up

WHEN TO USE IT

Use when time is limited, or the message needs to be delivered succinctly without overloading the audience.

The recommendation does not change. Only where you enter it, and how long you stay.

Example: Presenting a Workforce Readiness Review

During the last quarter, the Human Capital directorate reviewed workforce readiness across seven units, covering 1,350 mission-critical positions. Data was drawn from the HRIS, the learning management system, exit interview records, the annual engagement survey, and recruiting logs.

The workforce readiness index fell from 78 to 71 over four quarters. Average time to fill a vacancy rose from 46 to 60 days, an increase of 30%. Training completion held steady at 82%. Voluntary turnover rose from 12% to 16%. The engagement score fell from 3.7 to 3.4 out of 5.

Turnover was segmented three ways. By unit: Unit A 9%, B 11%, C 27%, D 10%, E 8%, F 25%, G 12%. By supervisor time in role: under one year 24%, one to three years 14%, three years and above 8%. By pay position against market: more than 10% below market 13%, within market 12%, above market 11%.

Exit interviews were coded for 40 leavers. Manager or supervision was cited by 45%, workload by 30%, career growth by 20%, pay by 15%. Engagement by supervisor tenure showed 2.9 for supervisors under one year, 3.5 at one to three years, 3.8 at three years and above. Certification lapses rose in two units, beginning after the departures rather than before.

Example — Told With the Standard Arc

Presenting a workforce readiness review. Comprehensive, building understanding.

Hook	Workforce readiness has fallen for four straight quarters, and the reason is not the one we were asked to check.
Context	Leadership asked us to look at compensation. Readiness is down seven points, vacancies are lasting a third longer, and voluntary turnover has risen from 12% to 16%.
Process	We segmented turnover three ways: by unit, by supervisor time in role, and by pay position against market. We then checked exit themes and engagement against the same cuts.
Insight	Turnover is not general. It sits in two units at 25 and 27 percent against 8 to 12 elsewhere. It falls steeply as supervisor tenure rises, from 24% under a first-year supervisor to 8% under one with three years. It does not track pay position, which is flat at 13, 12 and 11 percent.
Impact	If we fund a broad pay adjustment we spend the budget where the problem is not, and the two units carrying most of the loss stay where they are. Certification lapses in those units are already rising as a consequence.
Resolution	Fund targeted supervisor support in the two affected units this quarter, at moderate confidence, rather than a small across-the-board step.
Future	Two quarters will tell us whether it moves. We will also check whether the newest supervisors were assigned to the hardest units, which is the one thing that would change this answer.

Example — Told With the Inverted Arc

Same review. For a time-constrained decision-maker.

Recommendation	Fund targeted supervisor support in the two outlier units this quarter, rather than an across-the-board pay step.
Key insight	Turnover is concentrated in two units and tracks supervisor tenure, not pay. 24% under a first-year supervisor against 8% under a three-year one.
Supporting detail	Pay position is flat across the range at 13, 12 and 11 percent. Exit themes and engagement data point the same way. Three independent sources agree.
Impact	A broad adjustment spends the budget where the problem is not. A targeted one costs less and tells us within two quarters whether we are right.
Confidence and limitation	Moderate confidence. We cannot say supervision causes this, and we have not checked whether the newest supervisors were assigned to the hardest units.
Conclusion	Two units, this quarter, and a check on assignment patterns before anything wider.

Example — Told With the Truncated Arc

Same review. Ninety seconds at a steering update.

Key insight

Turnover is concentrated in two units and moves with supervisor tenure, not pay. 24% under a first-year supervisor, 8% under a three-year one, and flat across every pay grade.

Supporting detail

Three independent sources agree, at moderate confidence. What we cannot yet say is whether the newest supervisors were assigned to the hardest units.

Wrap up

We recommend funding supervisor support in those two units this quarter, and checking assignment patterns before anything wider.

Example: Presenting the Results of Hiring Operations

A second case. Same problem, different subject.

Over the past quarter, the staffing office processed 138 hiring actions across four directorates, covering GS-11 through GS-14 mission-critical roles. Applicant tracking data was drawn from 17 requisition records and six referral sources, producing more than 15,000 individual application events.

Across all actions, 14% experienced certification delays lasting from three to fifteen working days, primarily during periods of concurrent hiring surges. Average time to fill rose from 46 to 60 days. Offer acceptance held at the usual rate of 84%. Panel scheduling accounted for a mean of 9 days per action, reference checks 4 days, and security processing 11 days.

Analysis of the delayed actions found that the delays did not correlate with requisition complexity, grade level, or directorate. They correlated with weeks in which more than twelve actions were open simultaneously in the same classification specialist's queue.

Example — Told With the Standard Arc

Presenting hiring operations. Comprehensive.

Hook	One in seven hiring actions this quarter was delayed, and the cause is not the one everybody assumed.
Context	The staffing office ran 138 actions across four directorates. Time to fill rose from 46 to 60 days, and 14% of actions hit certification delays of three to fifteen days.
Process	We tested the delays against requisition complexity, grade level, and directorate. None of them explained the pattern. We then tested against specialist workload.
Insight	The delays track queue load, not difficulty. They cluster in weeks where a single classification specialist had more than twelve actions open at once.
Impact	Hiring managers experience this as an unpredictable process, which drives workarounds and repeat requests. Candidates in a 60-day pipeline are the ones we lose to faster employers.
Resolution	Cap concurrent actions per specialist at twelve and redistribute above that threshold, rather than adding headcount.
Future	We will know within two months whether time to fill moves. If it does not, the constraint is somewhere we have not looked yet.

Example — Told With the Inverted Arc

Same hiring review. For a director with five minutes.

Recommendation	Cap concurrent hiring actions per classification specialist at twelve and redistribute above that, rather than adding headcount.
Key insight	Delays track queue load, not requisition difficulty. They cluster in weeks where one specialist had more than twelve actions open.
Supporting detail	We tested complexity, grade, and directorate first. None explained the pattern. 14% of 138 actions were delayed three to fifteen days.
Impact	Time to fill has risen from 46 to 60 days. A cap costs nothing and can be tested this quarter. Adding headcount costs a position and takes six months.
Conclusion	Cap it, watch time to fill for two months, and revisit if it does not move.

Example — Told With the Truncated Arc

Same hiring review. A written update, three lines.

Key insight

Hiring delays track specialist queue load, not requisition difficulty. They cluster above twelve concurrent actions.

Supporting detail

Complexity, grade and directorate were tested first and explained nothing. 14% of 138 actions were delayed.

Wrap up

We propose capping concurrent actions at twelve and redistributing above that. Two months will tell us whether it works.

Same evidence, same recommendation, three different lengths. Nothing true changed between them.

Step 5: Tailor to Audience Needs

What are the specific outcomes or goals?

Inform or Educate	Monitor / Track Performance	Enable Decision-Making	Drive Action or Change	Build Consensus and Buy-In
Focus on clarity and completeness Provide background and unbiased facts Use neutral, objective language to explain findings clearly	Prioritize accuracy and transparency Present consistent metrics or KPIs over time Use simple, repeatable visuals for tracking trends	Highlight key findings quickly Present clear options and recommendations Use visuals to reduce complexity and support decision confidence	Frame the problem clearly and emotionally Use human-centered stories, real-world examples, actionable insights Focus on what needs to happen next and why it matters	Acknowledge different perspectives Use relatable stories, shared goals, data-backed common ground Aim to create shared understanding that drives agreement

Add Storytelling Components (If Needed)

Make it engaging, memorable, and persuasive.

Curiosity

heightens attention

Open a loop before you close it.
Make them want the answer before
you give it.

Visual Imagination

creates ownership

Give them something to picture. A
room, a person, a shift. Abstractions
do not stick.

Emotion

is the glue

Not manipulation. The recognition
that these are people leaving, not
units of turnover.

Self-Reflection

motivates action

Let them arrive at the conclusion. A
question they answer themselves is
theirs.

These techniques are especially useful in persuasive storytelling, but may not be needed in structured reporting.

Examples of Storytelling Techniques (If Needed)

Choose techniques that fit your objective and audience. Not all are required every time.

Curiosity	Visual Imagination	Emotion	Self-Reflection
Pose a provocative question	Metaphors and analogies	Anecdotes	Anecdotes the audience sees themselves in
Rhetorical techniques	Vivid imagery and descriptive language	Emotional appeal	Real-world impact illustration
Open a loop you close later	Dialogues and imagined scenarios	Contrasting sentiments	Case studies
	Action verbs and strong adjectives	Real-world impact illustration	

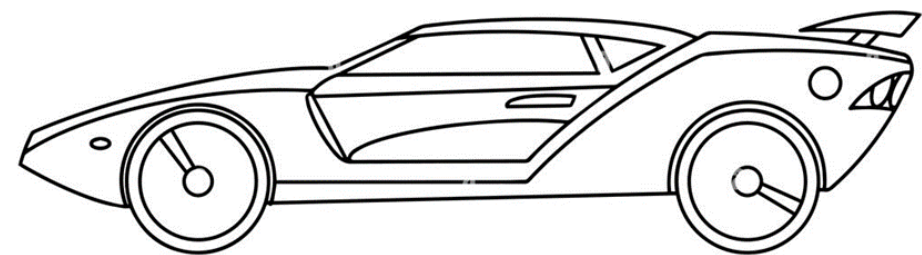
Use them to enhance persuasion and engagement. Never to close a gap in the evidence.

Example: Analogies to Avoid the “So What” Response

Today we have the industry’s first-shipping 32-nanometer process technology

Processor Comparison: 4004 vs 32-nanometer

METRIC	4004 PROCESSOR	32-NANOMETER PROCESSOR	IMPROVEMENT
Speed	1x (baseline)	5,000x	5,000 times faster
Transistor Cost	100,000x (baseline)	1x	100,000 times cheaper
Year Introduced	1971	2009	38 years of development
Transistor Size	10,000 nm	32 nm	~312 times smaller



470,000 mph
100,000 mpg
Cost: 3 cents

Example: Analogies to Avoid the “So What” Response

THE NUMBERS AS REPORTED

Metric	Unit C	All other units
Voluntary turnover	27%	8 to 12%
Headcount	150	1,200 combined
Leavers this year	40	139
Time to backfill	60 days	46 days
Certifications lapsed	Rising	Stable

“So what? Every organization has some turnover.”

THE SAME NUMBERS, WITH AN ANALOGY

“Unit C is a hundred and fifty people. At 27 percent, we are losing forty of them a year.

That is a full watch section, every year, from the one place we can least afford it. And it takes us sixty days to replace each one, so the section is never actually full.”

“A full watch section. Every year.”

Not one number changed. The analogy converted a rate into something the room can picture.

Step 6: Design Effective Visualizations

Two different jobs, and most people use the wrong one.

EXPLORE

Explore and gain insights · describe and diagnose

For you. Many charts, many cuts, no styling. You are looking for something and you do not yet know what.

Nobody else should ever see these.

COMMUNICATE

Communicate insights and decisions

For them. One chart, one message, deliberately designed. Leverage pre-attentive attributes and Gestalt principles so the eye lands where the meaning is.

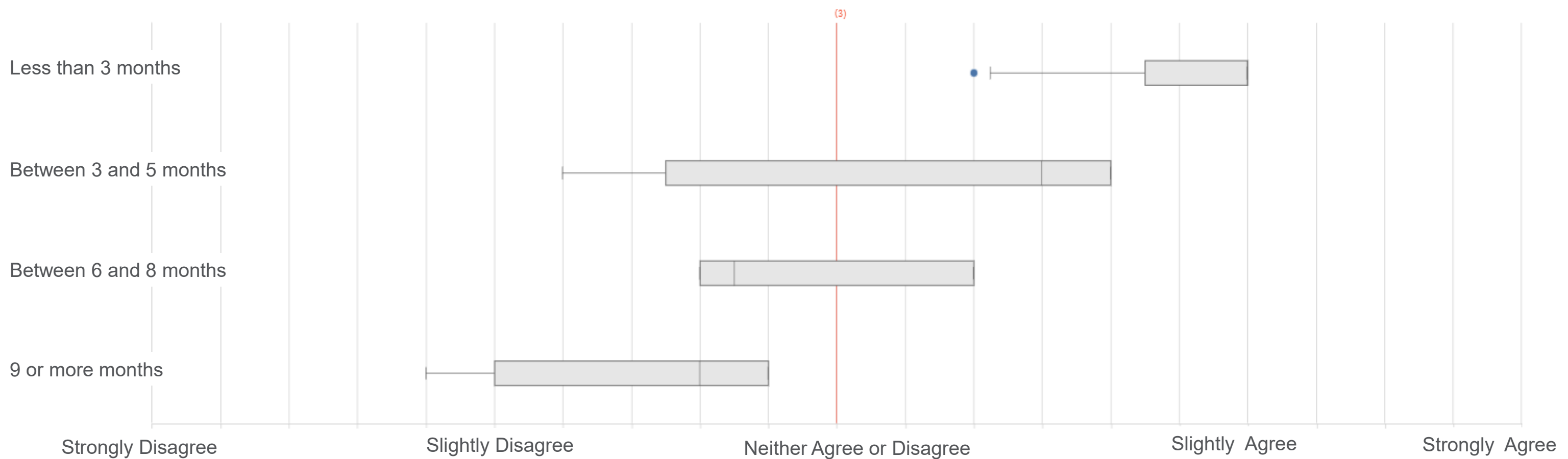
This is the only kind that should ever be in a briefing.

The most common visualization failure is showing an explore chart to an audience.

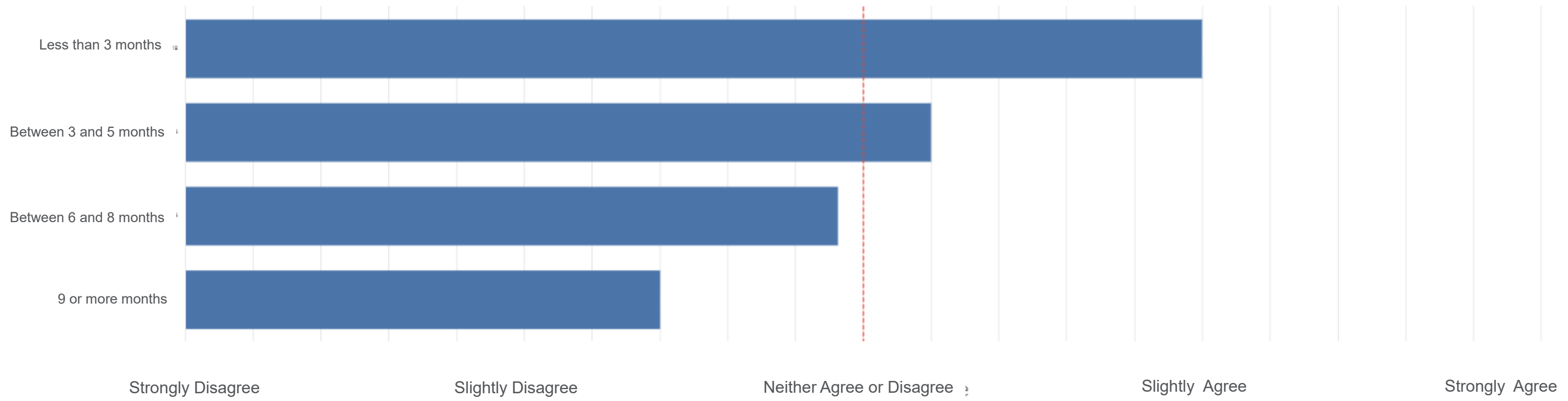
An explore chart says “here is everything, you work it out.” A communicate chart says “here is the one thing that matters, and here is why.”

Example – Explore Visualization

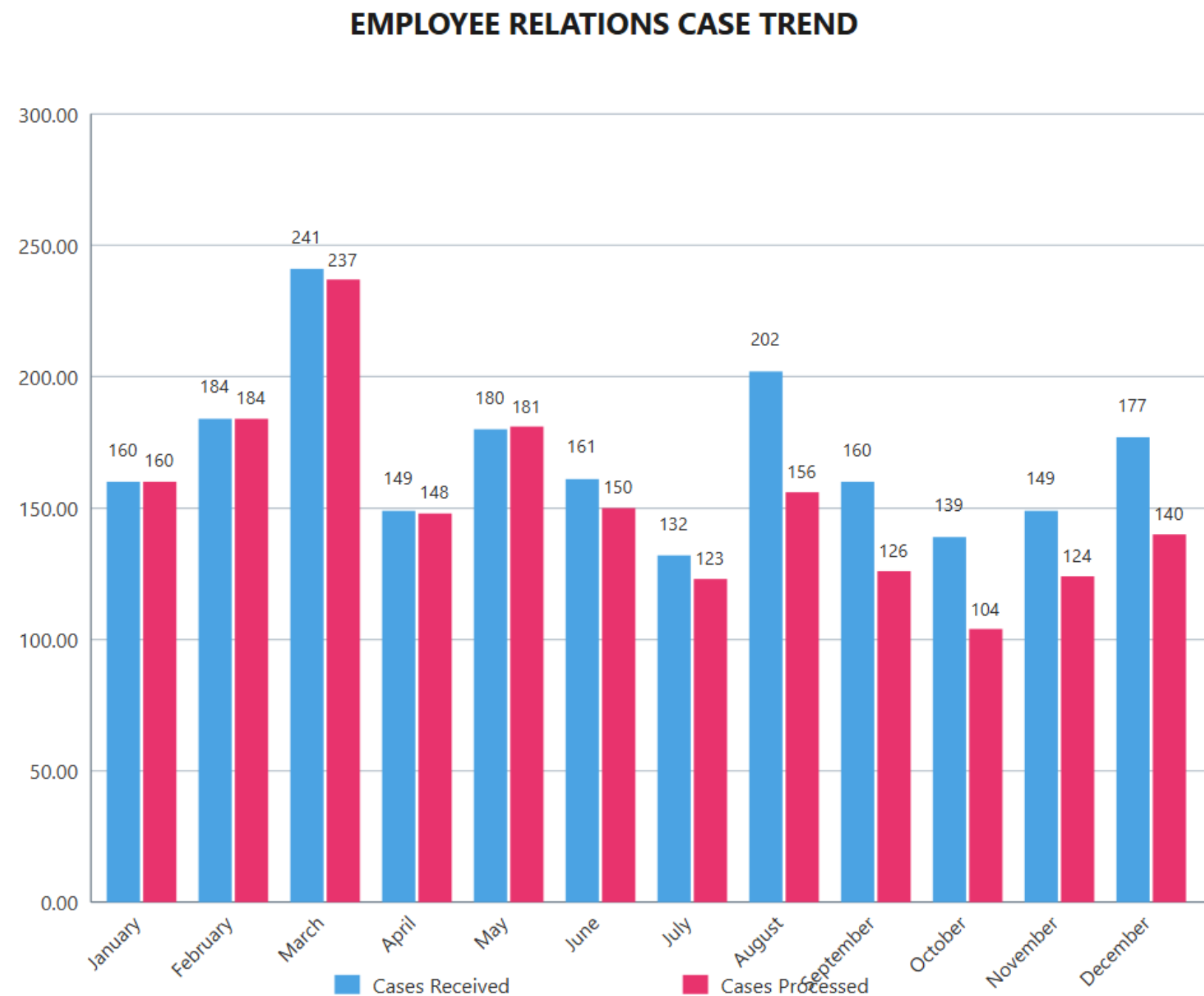
How did feedback vary based off employees' tenure within the first year?



Example – Communicate Visualization



Focus on Insights Over Raw Data

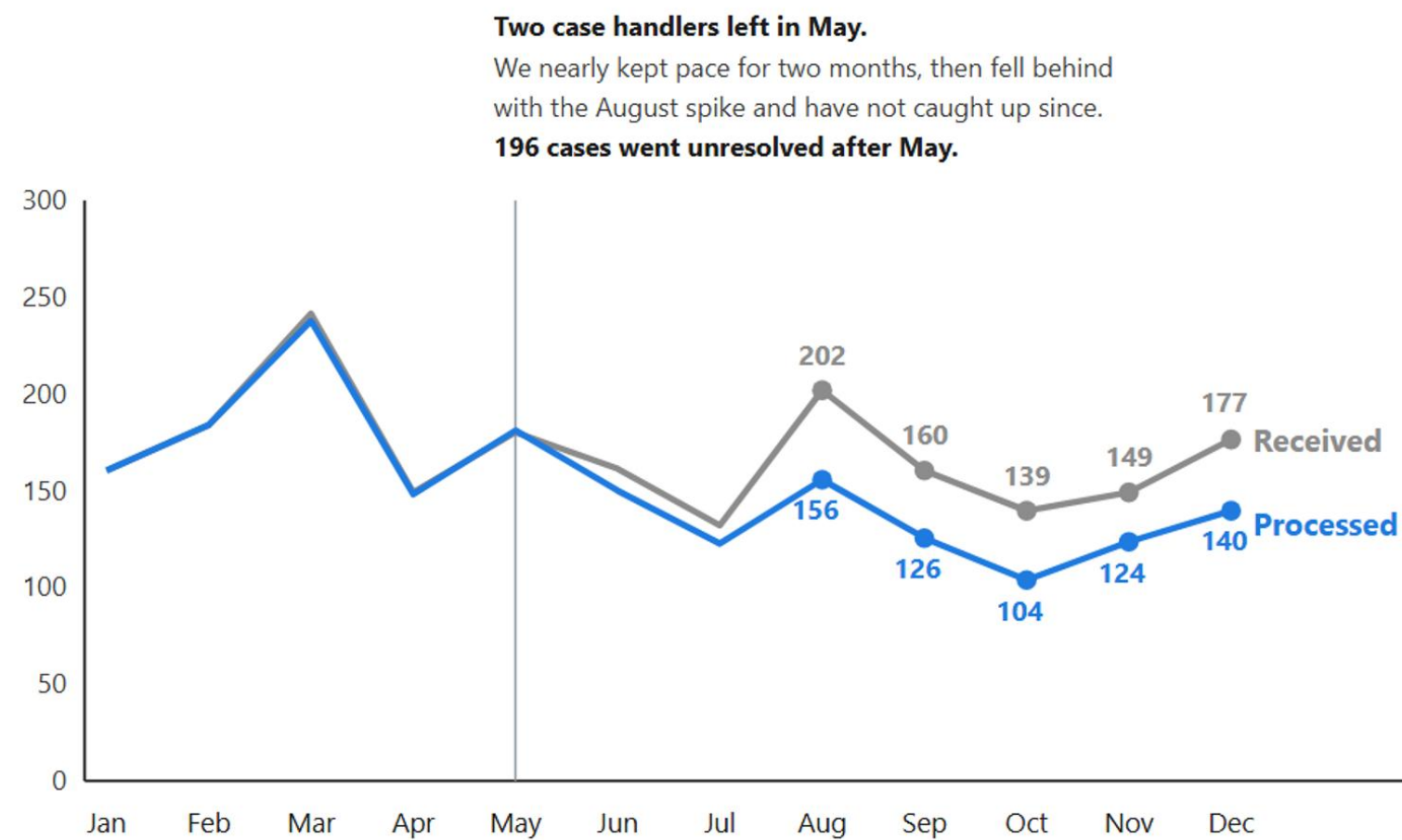


Focus on Insights Over Raw Data

Please approve two case handlers

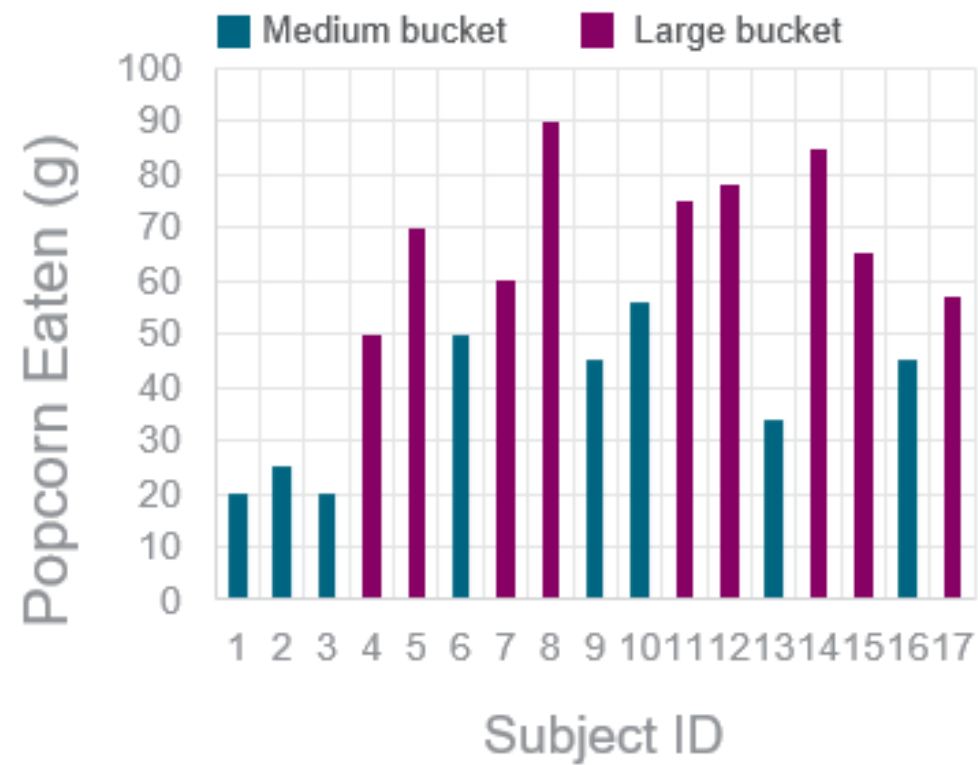
to backfill the two who left in May

ER case volume over time



Example – Focus on Insights Over Raw Data

Do people with larger buckets eat more stale popcorn?



People with larger buckets ate 53 percent more popcorn than people with the medium size

Design Has Two Jobs

Everything that follows is one of these two

Clarity

Help the eye find the point fast. Remove what does not carry meaning, and guide attention straight to the finding.

Honesty

Never let the chart argue what the data does not. The baseline, the denominator, the distribution, the color.

Clear but dishonest persuades of the wrong thing. Honest but cluttered persuades of nothing. You need both.

The Two Jobs, in Full

Every design decision serves one of these

CLARITY · HELP THE EYE FIND THE POINT

Remove what is not carrying meaning

- Cut gridlines, borders, and backgrounds
- Kill the legend when direct labels will do
- If it does not encode data, it competes with it

The eye sees before it reads

- One emphasis color, everything else grey
- Direct-label the finding
- If everything is highlighted, nothing is

Let the title say the finding

- Not “Turnover rate by unit, 2025”
- But “Turnover is concentrated in three units”
- Annotate the finding where it happens

THE CLARITY TEST

Can somebody find the point in five seconds without you talking?

HONESTY · ARGUE ONLY WHAT THE DATA SUPPORTS

The baseline makes the first claim

- Bars need a true zero, or length lies
- Keep the aspect ratio neutral
- One scale. Never a dual axis

Show the denominator and the spread

- Rate, not raw count. Print the base, n =
- An average alone hides its distribution
- A percent with no denominator misleads

Color judges, and time misleads

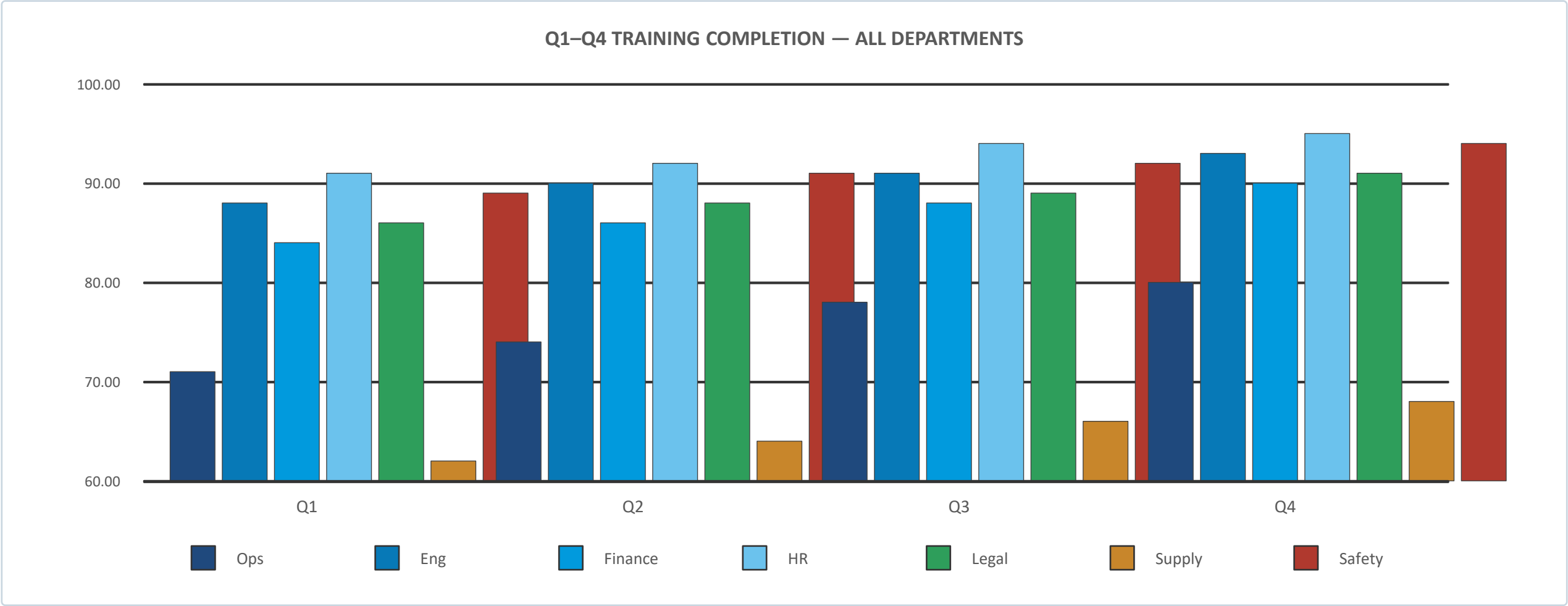
- Red reads as bad before any label
- Sequential for magnitude, diverging only with a true midpoint
- Vary lightness, never encode by color alone
- Dashed beyond today. Never let hope look like measurement

THE HONESTY TEST

Would every choice survive being said out loud to a skeptic?

Worked Example - Before

Mandatory training completion. Every number is accurate.

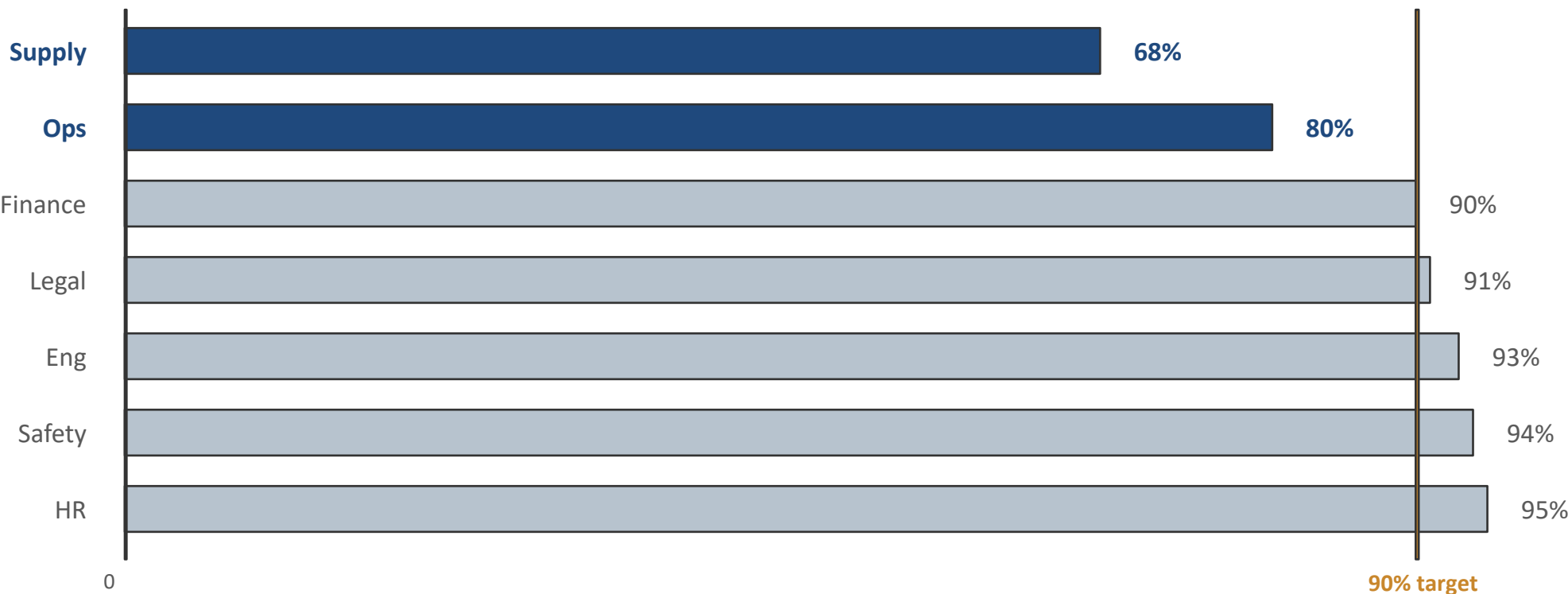


Worked Example - After

Same data but one question answered

Two departments are below the 90% training target

Q4 completion, all departments



Nothing was added. Three quarters, seven colors and a legend were removed, and the title now says the finding.

Worked Example - Which Moves Did That Use?

Clarity #1	Remove what is not carrying meaning	Dropped three quarters of history and the seven-color legend. The finding lives in Q4.
Clarity #2	The eye sees before it reads	Two departments in navy, the other five in grey. Attention goes where the problem is.
Clarity #3	Let the title say the finding	“Q1–Q4 Training Completion” became “Two departments are below the 90% target.”
Honesty #1	The baseline makes the first claim	The before chart started at 60, which exaggerated the gaps. The after starts at zero.
Honesty #2	Show the spread	All seven departments stayed in. Showing only the two would have hidden that five are fine.
Honesty #3	Color judges	Grey and navy, not red and green. The chart points; it does not scold.

Two questions, six moves. That is the whole module, and it is what the activity asks of you next.

Capstone - Present the Full Case

The Brief

Leadership asked what is actually happening, and what we should do. Answer them

What you have

The whole case. Your canvases, your cards, and every reveal from the two days.

20 min to prepare

What to produce

One telling: the recommendation, told as a story to an audience you name.

3 to 4 minutes

What to end on

What the analysis does not tell you. Say it plainly, and do not apologise for it.

One sentence

Use the Recommendation Canvas to build it and the Data Story Arc to arrange it. Both cards are on your desk.

On Monday morning this room was certain it was compensation. Show us what you have instead.

The Whole Journey

One continuous line, from a vague number to a story that moves people

QUESTION



EVIDENCE



READ



EXPLORE



EXPLAIN



VISUALIZE



RECOMMEND



TELL



You framed the question, weighed the evidence, read it, explored it, explained it, pictured it, recommended, and told it.

With judgment held honestly at every step. That is data literacy

The story changes but the evidence does not