BASIRA

CANCER METABOLISM
RESEARCH & TRAINING

بصيرة

TRAINING CURRICULUM

CHAPTER ONE

Thinking

How science reasons — and how not to fool yourself

Before any biology, any data, and any code: the habits of mind the rest of the curriculum depends on. This chapter is about how a claim earns belief, how you read a result honestly, and how one result becomes the next question.

Founder & Programme Lead	Aroob Gomosani
Prepared & delivered by	Haneen Marghalani
Version	2.0 · second cohort
Date	2026

Gomosani, A., & Marghalani, H. (2026). Basira Curriculum. Zenodo.
https://doi.org/10.5281/zenodo.21791751

1·0

Before you begin

This is the first chapter for a reason. Everything that follows — the biology, the datasets, the statistics, the code — is machinery. Machinery is only as good as the reasoning that decides what to point it at and what to believe once it has run. A researcher who can run a pipeline but cannot tell a lead from a finding is not yet a researcher.

You do not need any biology, statistics or programming to read this chapter. You need only a willingness to be wrong in public, which is the actual entry requirement for science.

WHERE THIS SITS

	Chapter	What it answers
▶️	1 · Thinking	How does science reason, and how do I avoid fooling myself?
	2 · Practice	How do we behave — integrity, collaboration, communication.
	3 · The Science	What is cancer, and how does a cell power itself?
	4 · Data & Tools	What data exists, and how do we handle it without breaking it?
	5 · The Toolkit	The statistics and code, as a reference you return to.
	6 · Build Your Own	Now go find a question of your own and run it.

WHAT YOU SHOULD BE ABLE TO DO BY THE END

- Place any claim on the ladder from guess to hypothesis to finding — including your own.
- List the rival explanations for a result, and say what test would eliminate each.
- Explain why a false positive costs a discovery science more than a false negative.
- Distinguish correlation from causation, and reproducibility from replication, in your own words.
- Read a result together with its caveats, and name the traps that ambush interpretation.
- Turn a result into a hypothesis that is testable, specific, falsifiable and consequential.
- Choose language that matches the strength of your evidence, neither over- nor under-claiming.

HOW TO READ THE COLOURED BOXES

These six appear throughout every chapter and always mean the same thing. Each carries a word and a symbol as well as a colour, so they still work in black and white.

Box	Use it to
◆ KEY IDEA	The thing to remember if you remember nothing else from the section.
■ DEFINITION	A term given a precise meaning, used precisely from then on.
▲ CAUTION	A trap that catches careful people. Read these twice.
● WORKED EXAMPLE	The abstract idea made concrete on a real case.
? CHECKPOINT	Stop and answer before continuing. Mid-chapter, not just at the end.
▸ GO DEEPER	Videos, articles and papers if you want more than the chapter gives.

1.1
What research actually is

Research is disciplined truth-seeking under uncertainty. Each word carries weight.

Truth-seeking, because the aim is to find out how things really are, not to win an argument or confirm what we hoped. **Under uncertainty**, because if the answer were already known we would not be doing research. And **disciplined**, because the human mind is extraordinarily good at seeing what it wants to see, and the whole apparatus of method exists to protect us from that tendency.

It helps to be precise about three things that ordinary speech blurs together. A guess is a claim with no evidence behind it. A hypothesis is a claim stated precisely enough that evidence could count for or against it. A finding is a claim that evidence has actually tested and, for now, supported. Moving a statement up that ladder is the daily work of research — and pretending a statement sits higher than it does is the most common failure of honesty in science.

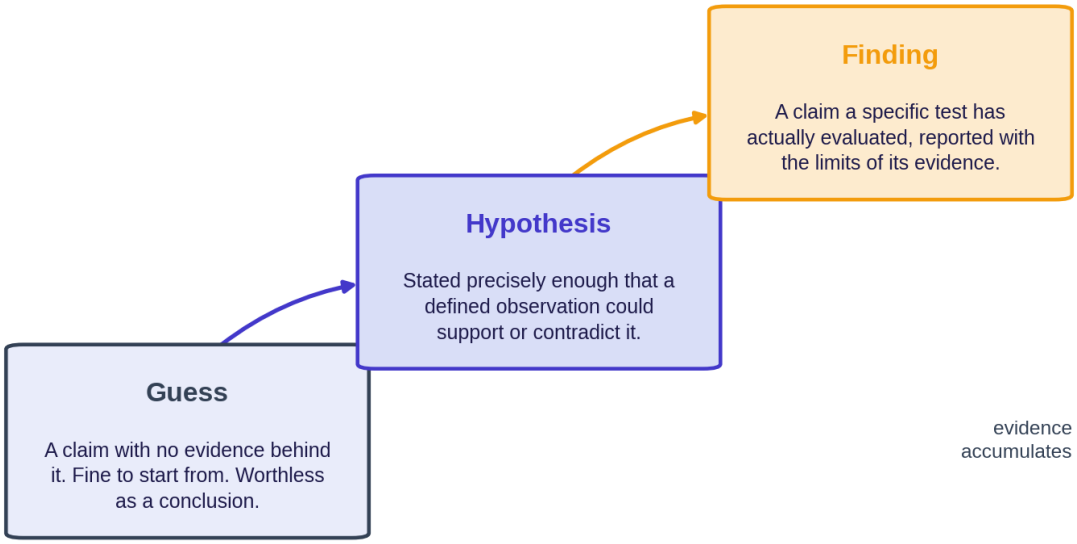


Figure 1.1 The ladder every claim climbs. Most of what goes wrong in science is a claim reported one rung higher than it earned.

◆ **KEY IDEA — The asymmetry at the heart of research**

It is easy to find support for a belief you already hold. If you go looking for confirmation, you will always find some. The discipline of science is the opposite habit: actively trying to break your own idea.

A hypothesis earns credibility not by accumulating confirmations but by surviving serious attempts to refute it. Your first loyalty as a researcher is to the possibility that you are wrong.

1.2
The method, as it is actually practised

The version of the scientific method taught in school — observe, hypothesise, test, conclude, in a tidy line — is a cartoon. Real research is iterative and messy. Hypotheses get revised mid-stream, analyses fail and are redesigned, dead ends are common, and the clean narrative is assembled only in hindsight.

That messiness is normal, and it is not a licence for sloppiness. One thing survives from the cartoon and is non-negotiable: a claim must be exposed to the genuine possibility of being wrong, and the competing explanations for a result must be actively ruled out.

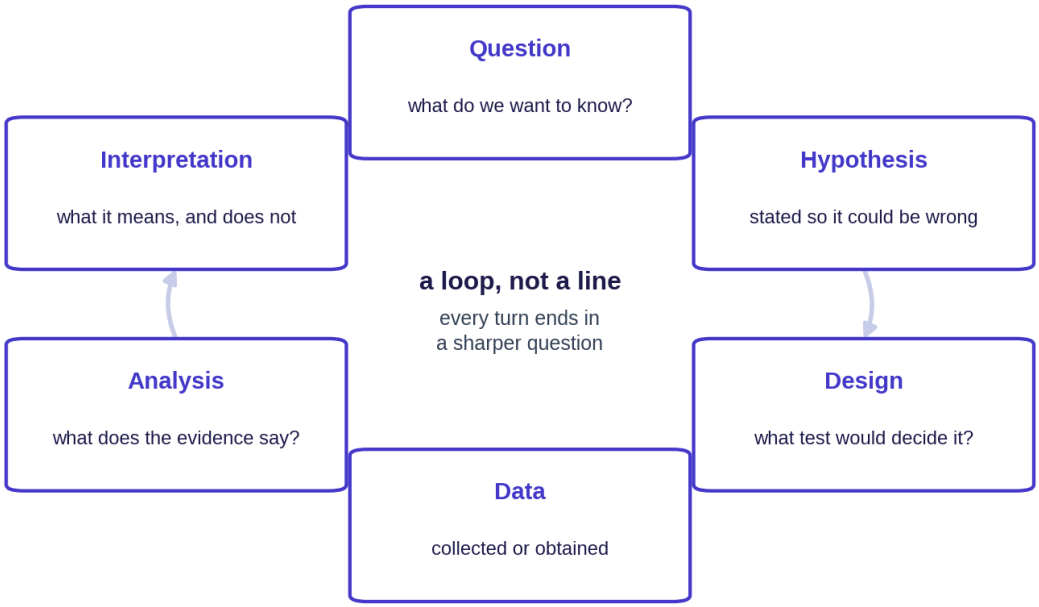


Figure 1.2 Research is a loop, not a line. The end of one turn is the beginning of the next — which is the subject of §1.9.

Research also never starts from nothing. Others have worked on your question, or on its neighbours. Knowing that prior work is a matter of efficiency — you should not spend a year rediscovering what is already known — and a matter of integrity, because building on others' work requires acknowledging it. Reading the literature critically rather than credulously is a core research skill, and one you will practise from your first week.

1.3
Strong inference

Ruling out alternatives deserves its own name, because it is the engine of good work. The practice is called strong inference: for any result you are excited about, deliberately enumerate the other explanations that could produce the same observation, then design the analysis that distinguishes among them.

A result is only as strong as the alternatives you have eliminated. A finding with three plausible rival explanations still standing is not a finding — it is a hypothesis wearing a finding's clothes.

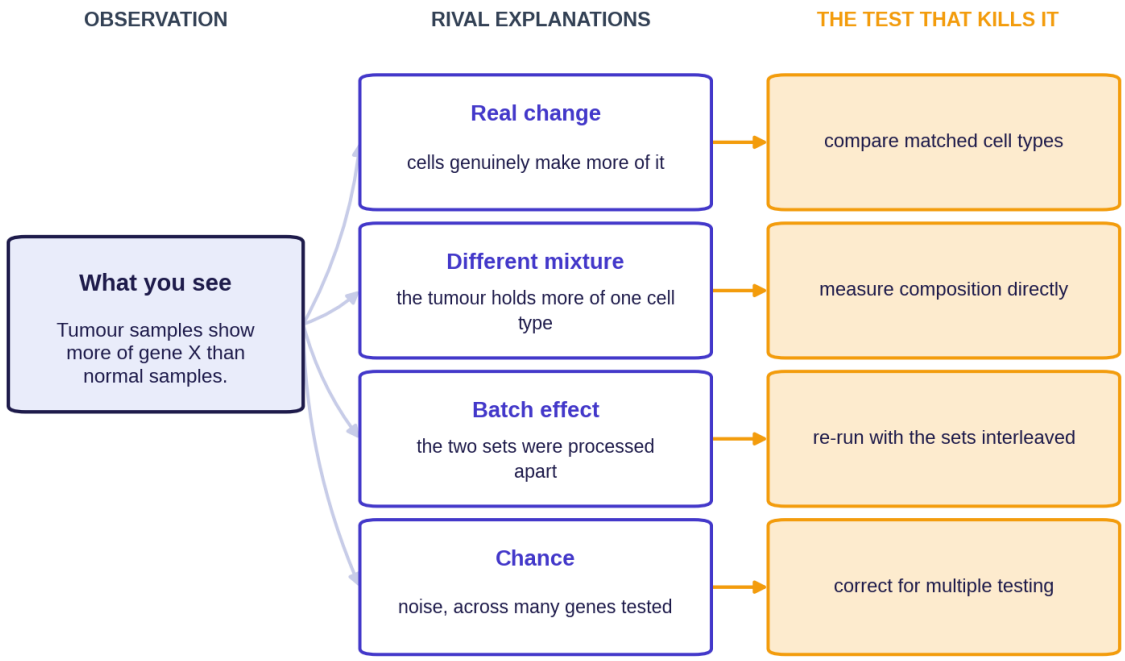


Figure 1.3 One observation, four rivals, four tests. Notice that the observation alone does not favour any of them.

◆ KEY IDEA — **Strong inference, as a habit you can actually perform**

For every result: write down the list of rival explanations — confounders, artefacts, coincidences — and ask what test would separate the real explanation from each rival. Then run that test.

A result with one obvious explanation and no proposed test is a dead end. A result with three competing explanations and a plan to tell them apart is a research programme.

? CHECKPOINT — **Checkpoint 1**

Answer these before continuing. If one gives you trouble, reread the section in brackets.

- State the difference between a guess, a hypothesis and a finding, and give an example of each from anything you know about. (§1.1)
- Why is trying to break your own idea more informative than trying to support it? (§1.1)
- A friend says “the scientific method is: observe, hypothesise, test, conclude.” What is right and what is misleading about that? (§1.2)
- You observe that students who attend more lectures get better grades. List three rival explanations, and one test for each. (§1.3)

1.4
Signal, noise, and two ways to be wrong

Every measurement contains both signal — the thing you want to know — and noise, which is everything else. The central difficulty of research is that you never see them separately. You see only their sum, and you must reason about which is which. Statistics, at bottom, is the discipline

of not fooling yourself about that: of quantifying how much of an apparent pattern could be noise, so you do not mistake a coincidence for a discovery.

Because you are reasoning under uncertainty, you can be wrong in two distinct ways, and they trade off against each other.

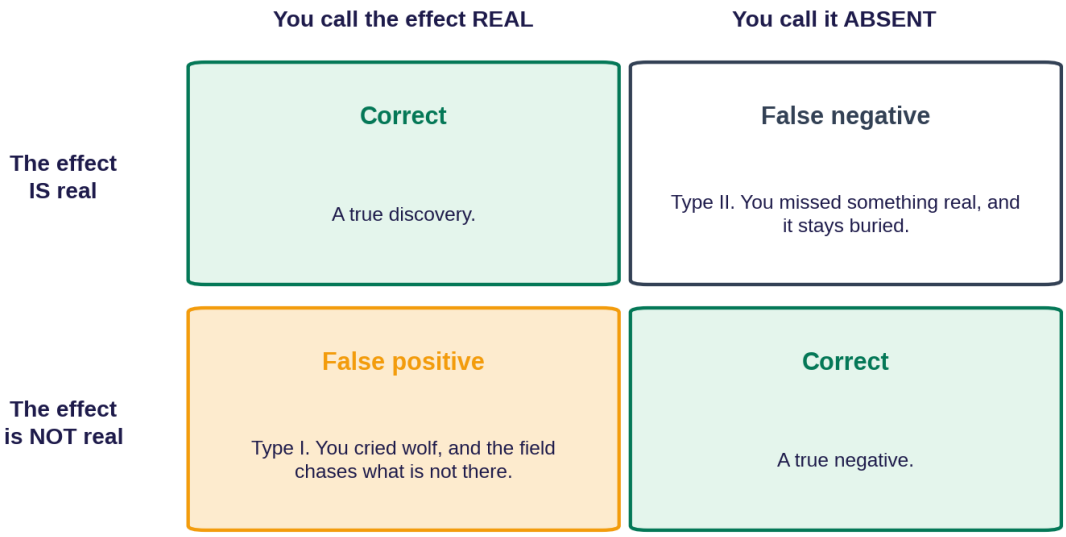


Figure 1.4 The two errors. Tightening against one loosens you against the other — there is no setting that avoids both.

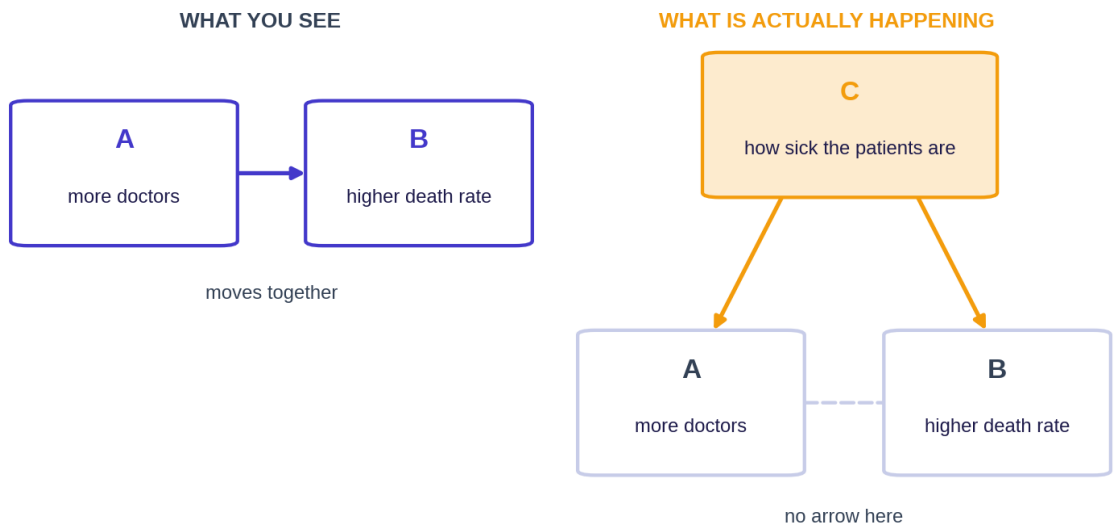
A false positive sends a whole field chasing something that is not there, wasting years and money that were not yours to waste. A false negative quietly buries something real. Both are costly, but in a discovery science the false positive is the more damaging, because it propagates: others build on it, cite it, and design new work around it. Much of the machinery you will meet later — multiple-testing correction, effect-size thresholds, positive and negative controls — exists to hold the false-positive rate down without blinding us entirely.

> ▲ CAUTION — **Absence of evidence is not evidence of absence**
>
> "We did not detect an effect" can mean two completely different things: the effect is not there, or our test could not have seen it even if it were. These must never be reported as the same thing.
>
> Basira keeps them separate in words. "Tested and null" is a result. "Not testable" is not a result — it is a gap, and it gets reported as one.

1.5 Two confusions that ruin conclusions

1.5.1 Correlation is not causation

Two things moving together does not mean one causes the other. A third variable — a confounder — may be driving both. This is the single most common error in reasoning from data, and it is common precisely because the causal story is so much more satisfying than the boring one.



Large hospitals receive the sickest patients. Severity drives both the staffing and the deaths.

Figure 1.5 A confounder produces a real, reproducible association with no causal link between the two things you measured.

● **WORKED EXAMPLE — Why the association is real and the conclusion is wrong**

Hospitals with more doctors per patient record higher death rates. Taken at face value, that suggests doctors are dangerous.

The association is genuine and will replicate in fresh data. But large, well-staffed hospitals are precisely the ones that receive the sickest and most complex patients, while stable patients are treated in smaller facilities. Severity drives both the staffing and the deaths.

The lesson is not that data lies. It is that an association constrains what can be true without telling you which of the surviving stories is true.

1.5.2 Reproducibility is not replication

These two words are used interchangeably in conversation and must never be used interchangeably in a claim.

Figure 1.6 Reproducibility is a floor, not a ceiling. It shows your analysis is honest, not that your conclusion is right.

? CHECKPOINT — Checkpoint 2

- In your own words: why is a false positive more damaging than a false negative in a discovery science? (§1.4)
- A study reports “no significant difference between groups.” What is the one question you must ask before believing the effect is absent? (§1.4)
- Give an example of a correlation from everyday life that you are confident is not causal, and name the confounder. (§1.5.1)
- A colleague says “I reran your code and got the same numbers, so the finding is solid.” What have they established, and what have they not? (§1.5.2)

1.6 Reading a result

The most important reframe in this chapter is also the simplest. When an analysis hands you a number, a ranked list, or a hit, it is tempting to feel the work is finished and the discovery made. It is not. A result is a lead: a hypothesis with some evidence behind it, pointing at where to look next.

This is not false modesty. A well-motivated lead is genuinely valuable — a specific reason to investigate a particular thing rather than the thousand alternatives. But its value is as a direction, and treating it as settled skips every step that would actually establish it.

Every result arrives with caveats, and interpreting it responsibly means holding those caveats in the same hand as the result. A result measured in one population may not hold in another. A borderline call may flip under a slightly different choice of method. An under-powered comparison needs replication before it is anything more than a lead. To interpret a result without its caveats is to interpret a different, more certain result than the one you actually have.

1.7 The traps

The moment a result confirms what you hoped, or promises to be important, is precisely the moment your judgement is least reliable — because now you want it to be true. Four traps do most of the damage.

- **Confirmation bias.** Weighing evidence that supports the exciting result more heavily than evidence against it. Not a decision you make; a drift you fail to notice.
- **Narrative seduction.** A clean, compelling story feels true. But the universe is under no obligation to be tidy, and a good story is not evidence. Be suspicious of results that are unusually satisfying.
- **HARKing.** Hypothesising After the Results are Known — building a hypothesis around a pattern you already saw, then presenting it as though you had predicted it. A prediction

that survives testing is strong evidence; a story fitted to data afterwards is not. HARKing disguises the second as the first.

- **Patterns in noise.** Test enough things and striking patterns arise by chance alone. Your eyes do not correct for this; they are, in fact, optimised to do the opposite.

> ▲ CAUTION — **The exciting result is the dangerous one**
>
> The defence is not to catch yourself in the moment — nobody reliably can. It is to lean on process: thresholds fixed before you look, controls run before claims, exploratory work labelled exploratory.
>
> A result you are thrilled by deserves more scrutiny than one you are indifferent to, not less. If you notice yourself arguing for a result rather than testing it, stop.

1.8
Saying exactly what the evidence supports

The cardinal virtue of scientific communication is calibration: claiming neither more nor less than the evidence warrants. Over-claiming is the familiar sin — it sends people acting on a certainty that is not there. But false modesty is also a failure, because burying a solid result in hedging wastes it. The goal is language that tracks your confidence so precisely that a careful reader can tell how much to trust each sentence from the words alone.

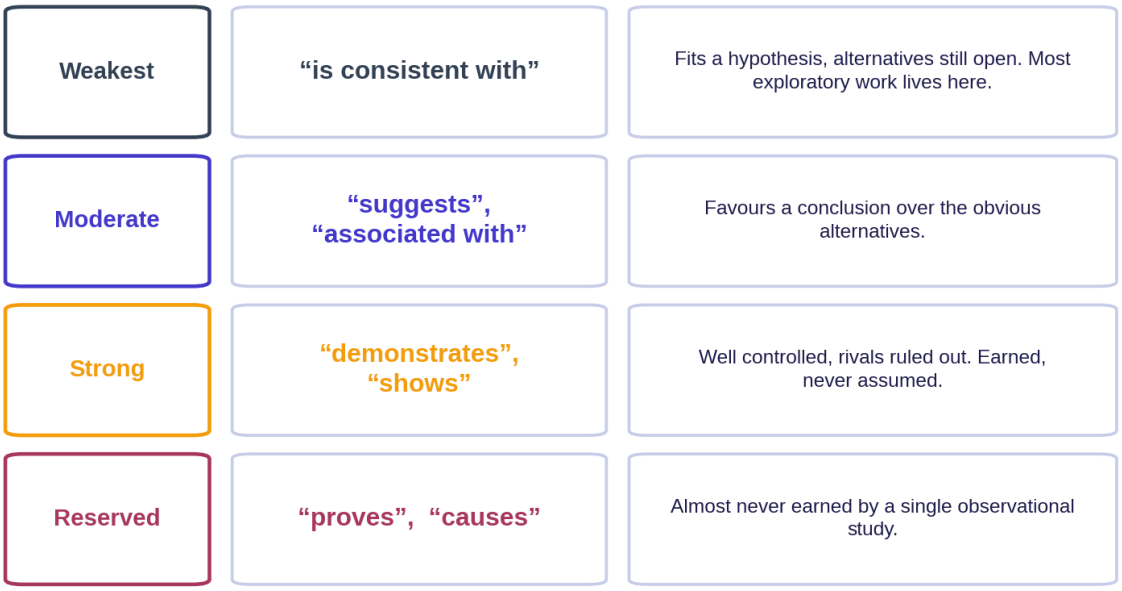


Figure 1.7 Climb this ladder deliberately. Most exploratory work belongs on the bottom rung, and saying so is a strength.

Two habits follow. Always distinguish exploratory from confirmatory in words: a result you went looking for after seeing the data is a hypothesis for the next study, not a confirmed finding, and it must be described that way. And state limitations plainly rather than hoping nobody notices — naming the weaknesses of your own work makes a reader trust everything else you say.

? CHECKPOINT — Checkpoint 3

- Why does a result you are excited about need more scrutiny than one you are indifferent to? (§1.7)
- Explain HARKing to someone who has never heard the term, and say why it inflates apparent evidence. (§1.7)
- Rewrite this sentence so it is honestly calibrated: “Our analysis proves that gene X causes resistance to treatment.” (§1.8)
- You find a striking pattern after trying twelve different comparisons. What must you say about it when you report it? (§1.7–1.8)

1.9
From result to hypothesis

A result becomes science when it generates the next question. This is the creative heart of research and the hardest part to teach, because it is genuinely generative — you are not following a procedure, you are proposing an explanation.

The move has a recognisable shape. A result presents a fact in need of explanation. A hypothesis is a proposed explanation specific enough to be tested. The reasoning that gets you from one to the other is called abduction, or inference to the best explanation. Where deduction moves from premises to a guaranteed conclusion, and induction generalises from many observations, abduction runs the other way: from a single striking observation to the hypothesis that would best account for it.

■ DEFINITION — Abduction

Reasoning from an observation to the explanation that would best account for it. Abduction supplies candidate explanations; it does not tell you which is right. That is what the next experiment is for.

● WORKED EXAMPLE — The generative move, performed slowly

Observation: patients taking a particular drug live longer than patients who are not.

Abduce — what could account for this? (1) The drug works. (2) Healthier patients are the ones well enough to be prescribed it. (3) Patients who tolerate it are those whose disease is less aggressive to begin with. (4) The two groups were followed for different lengths of time.

Distinguish — for each pair of rivals, what observation would tell them apart? Comparing patients matched on baseline health addresses (2). Checking who stopped treatment early addresses (3). Aligning follow-up windows addresses (4). What survives all three is a real candidate for (1).

Notice that this is strong inference from §1.3, run forward. The same engine that evaluates a result also generates the research plan.

1.10

What makes a hypothesis good

Not every question a result suggests is a good hypothesis. The generative move is only useful if what it produces can actually be tested and would actually matter.

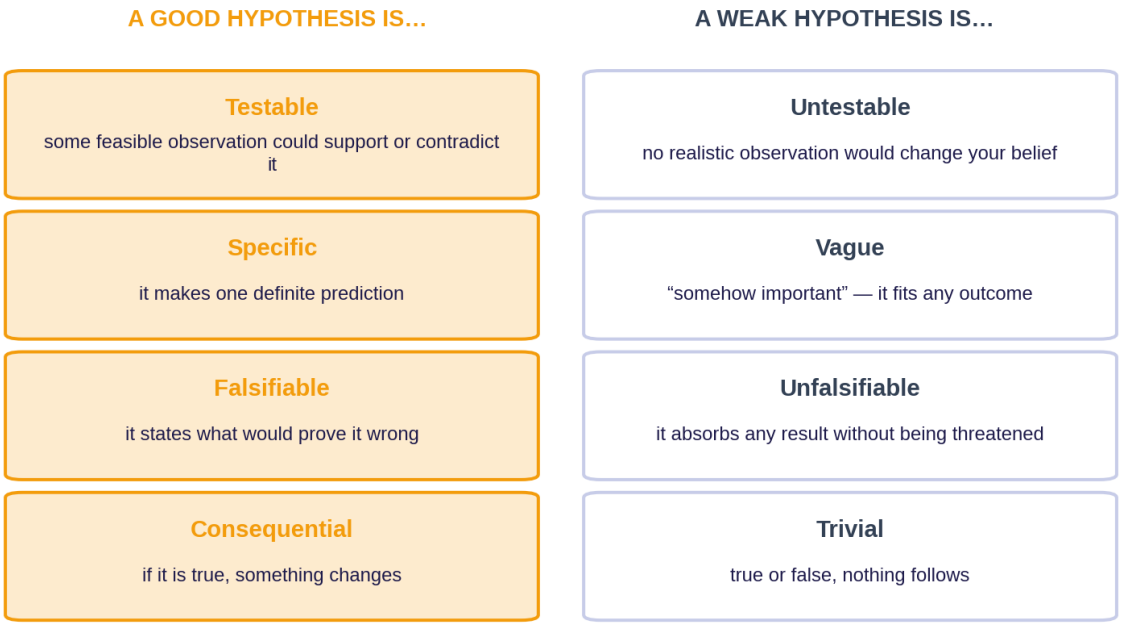


Figure 1.8 Four tests. Apply them to your own ideas before you invest months in one.

These are the same virtues demanded of any scientific claim, now applied at the moment the claim is born. A hypothesis that cannot fail is not a scientific hypothesis, however plausible it sounds. A hypothesis whose answer would change nothing is not worth the cost of testing. Hold your own new ideas to this standard, and hold them there before you are emotionally invested rather than after.

1.11
How a research project is built

A project has a shape, and knowing it helps you see where any given day's work fits. It runs from a question, to a design that could answer it, to data, to analysis, to interpretation, to communication. Weakness at any stage undermines everything after it: a brilliant analysis of the wrong data answers the wrong question, and a flawless result described dishonestly does harm.

Three practices hold the whole structure together, and you will meet all three again in Chapter 2 as matters of integrity, and in Chapter 5 as lines of code.

- **Pre-specification.** Decide the choices that matter — thresholds, which comparison, what counts as a hit — before you see the results. Choosing afterwards, tuning until the answer looks good, is one of the easiest ways to fool yourself.
- **Controls.** A positive control asks whether your method recovers something already known to be true; if it cannot, you do not trust it on the unknown. A negative control asks

whether it correctly stays silent where there is nothing. Both run before any novel claim is believed.

- **Provenance.** Every number in a final figure or table traces back to the data and code that produced it. Nothing is typed by hand. This makes the work auditable by a colleague, by a reviewer, and by you in six months when you have forgotten everything.

> ◆ KEY IDEA — **The Basira standard, stated plainly**
>
> Every number is computed; nothing is typed by hand. Thresholds are fixed before results are seen. Controls run before claims. What could not be tested is reported as untested, not as negative.
>
> Hold to these four and most of what follows in this curriculum takes care of itself.

1.12 End-of-chapter questions

The checkpoints tested recall. These ask you to reason. There is no single right answer to several of them — write a paragraph, and bring it to the session.

- Take a claim you have heard recently — in a lecture, in the news, from a friend. Place it on the ladder in Figure 1.1 and justify the placement.
- Choose any published finding you find interesting. List every rival explanation you can think of, then say which the authors addressed and which they did not.
- Describe a time you believed something because you wanted it to be true. What would have caught it?
- “A reproducible result can still be wrong.” Explain this to someone who thinks the two words mean the same thing, using an example.
- Write one good hypothesis and one weak hypothesis about the same topic. Explain what makes the difference, using all four criteria in Figure 1.8.
- You run twelve comparisons and one is significant. Write the sentence you would use to report it honestly, and the sentence a careless researcher would write.
- Pick a question you would like to answer one day. Run it through the funnel: is it testable, specific, falsifiable, consequential? Sharpen it until it passes all four.

1.13 Go deeper

Every link below is live and goes straight to the resource. [Watch] is video, [Read] is written, [Source] is the original paper or essay.

CRASH COURSE · SCIENTIFIC THINKING

A seven-episode series built for exactly this chapter. If you watch nothing else, watch this.

[Watch] **The full series — all 7 episodes** — Start here. Roughly ninety minutes end to end.

[Watch] **#2 Statistical Thinking in Science** — Signal, noise, and what a statistic can and cannot tell you. Pairs with §1.4.

[Watch] **#4 Peer Review and the Quest for Truth** — How claims get checked before anyone believes them. Sets up Chapter 2.

[Watch] **#5 How Scientists Build Consensus** — Why one result is never the answer. Pairs with §1.6.

[Watch] **#6 Evaluating Sources & Fact Checking** — Directly on the traps in §1.7.

[Watch] **#7 When Science Clashes with the Public** — Uses the smoking and cancer story. Pairs with §1.5 and §1.8.

[Read] **Crash Course · series homepage** — Episode list and teaching notes.

WORKING THROUGH IT YOURSELF

[Read] **Khan Academy — Statistics & Probability** — Free, worked, self-paced. The study-design and confounding sections pair with §1.5.

[Read] **Calling Bullshit — Bergstrom & West, University of Washington** — A full university course on spotting bad data reasoning, free online. The causality lectures are the ones for §1.5.

THE ORIGINALS

[Source] **Platt (1964) — Strong Inference** — Where §1.3 comes from. Short enough to read in one sitting.

[Source] **Ioannidis (2005) — Why Most Published Research Findings Are False** — The argument is more careful than the title suggests.

[Source] **Simmons, Nelson & Simonsohn (2011) — False-Positive Psychology** — With enough analytic freedom, almost any hypothesis can be made to look significant.

[Source] **Feynman (1974) — Cargo Cult Science** — Caltech commencement address, widely reprinted. On leaning over backwards to report what might prove you wrong.

1.14
Glossary

Abduction Reasoning from an observation to the explanation that would best account for it. Generates candidate hypotheses; does not decide between them.

Calibration Matching the strength of your language, and your belief, to the strength of your evidence.

Confirmation bias The tendency to weigh evidence supporting a belief you already hold more heavily than evidence against it.

Confounder A third variable that influences both of two things that appear associated, producing a real correlation with no causal link between them.

Controls (positive / negative) A positive control checks that a method recovers something known to be true. A negative control checks that it stays silent where there is nothing.

Falsifiability The property of a claim that some possible observation would show it to be false. A claim nothing could refute is not a scientific claim.

False negative (Type II) Concluding an effect is absent when it is real. The effect stays buried.

False positive (Type I) Concluding an effect is real when it is not. The costlier error in a discovery science, because others build on it.

Finding A claim that a specific test has evaluated, reported together with the limits of the evidence behind it.

Guess A claim with no evidence behind it. A legitimate starting point and an illegitimate conclusion.

HARKing Hypothesising After the Results are Known: presenting a pattern found after the fact as though it had been predicted in advance.

Hypothesis A claim stated precisely enough that a defined observation could support or contradict it.

Noise Everything in a measurement that is not the thing you are trying to measure.

Pre-specification Fixing the decisions that matter — thresholds, comparisons, what counts as a hit — before results are seen.

Provenance The traceability of every reported number back to the data and code that produced it.

Replication Reaching the same conclusion in new, independent data. Far stronger evidence than reproducibility.

Reproducibility Obtaining the same result from the same data and the same analysis. A minimum standard; a reproducible result can still be wrong.

Signal The part of a measurement that carries the information you are after.

Strong inference Enumerating the rival explanations for a result and designing the test that eliminates each.