decidenda

FRAME and the eighteen methods

Five things to write down before anyone knows how a decision will turn out. Then you can judge the decision itself, not its result.

F	R	A	M	E
Focus	Relevant evidence	Alternatives	Manage uncertainty	Execute and learn

decidenda by Business Case Weekly · decidenda.com
Edition of October 2026

Where are you stuck?

Find the step where you are stuck, then the line that sounds like your problem. Each line takes you to the chapter on the method that helps. The last two methods are for checking a finished memo before you send it.

Introduction: FRAME and the five checks

F Focus

What exactly is being decided?

The question sounds big, but nobody can say what is being decided	Clarity test
The goal is a metric, and nobody has asked what it leaves out	Map is not the territory

R Relevant evidence

Which facts would change the answer?

Several explanations fit the same facts	Competing hypotheses
The argument rests on a famous winner's example	Survivorship bias
The argument rests on experience: "we have done this before, and it worked"	Kind vs wicked learning environments

A Alternatives

What is rejected, and at what price?

What is rejected

Every option comes back as the same safe answer	Inversion
The rejected option is described so that it is easy to reject	Steel-man

At what price

You can say what the choice gains but not what it gives up	Opportunity cost & trade-offs
The first effect is obvious; the reaction to it is not	Second-order thinking

M Manage uncertainty

What is unknown, and what would change the decision?

What is unknown

The plan rests on a forecast or on “we are different”	Outside view
The plan feels safe and you cannot say how it would fail	Pre-mortem
Everything hangs on one customer, supplier or person	Fragility audit

What would change the decision

You cannot name the news that would make you drop the decision	Falsifiability

E Execute and learn

What happens first, and when is it checked?

What happens first

You are agonising over a small bet or rushing a big one	Two-way / one-way door
Others will carry out the plan, and it will change on the first day	Commander’s intent

When it is checked

You will not know the result for months	Decision journal

Before you send

Will the memo hold up?

The memo reads well, and you need its weakest point	Decision Quality: six links
Nobody who would argue against the memo has read it yet	Red team

Introduction: FRAME and the five checks

Five things to write down before anyone knows how a decision will turn out. Then you can judge the decision itself, not its result.

In the middle of 1985 Andy Grove, Intel's president, sat in his office in Santa Clara. With him was Gordon Moore, Intel's chairman and chief executive. Through the window Grove could see the Ferris wheel of the Great America amusement park turning in the distance. Intel had started out making memory chips, and now Japanese rivals were taking that market by selling cheaper chips. For almost a year Intel had drifted, with no plan for its memory chips. Grove turned from the window and asked Moore a question. If the board fired them both and brought in a new chief executive, what would the newcomer do? Moore answered without hesitating: the newcomer would get Intel out of memory chips. Grove stared at him, numb. Then he asked why the two of them should not walk out of the door, come back in and do it themselves.

That autumn Intel said it would stop making DRAM, the memory chip it had pioneered. The next year it lost $173 million, its first yearly loss since its shares went on sale in 1971. By 1993 Intel was the world's largest chipmaker. It made its money from processors, the main chips in personal computers.

People usually tell this story as a lesson in courage, and that is the mistake. The decision looks right because everyone knows what came next. Suppose IBM and the companies copying its computers had switched to another company's processor. Then that talk by the window would be remembered as the day Intel's leaders gave up the

business they had built. In 1985 nobody could know for sure which of the two stories would come true. Grove described the scene in a book eleven years later, when he already knew the ending.

Judging a decision by its ending is a natural habit with three faults. First, the ending comes late, often years after the choice. Second, the ending mixes judgement with luck. A careful decision can fail because a rival moved first, and a careless one can succeed because the whole market grew. Third, knowing the ending changes how we read the beginning. The psychologist Baruch Fischhoff tested this for his doctorate at the Hebrew University of Jerusalem and published it in 1975. He gave people short descriptions of events from history and medicine, each with four possible outcomes. He told some readers which outcome had happened. Often he named one that never happened. Whatever outcome they heard seemed more likely to them than to readers who heard nothing. They also believed they would have seen it coming.

So before the ending, you can judge only the reasoning. Reasoning is far easier to judge once it is written down: others can check it, and the writer cannot quietly change it later. Nobody wrote down Grove and Moore's talk. Many decisions are like that: settled in a room and explained afterwards. But some decisions are written up before they are approved. The person who approves such a page cannot know the future, but can still ask five questions. Does it name the real decision? Does it rest on the facts that matter, or on a story that sounds right? Does it say which option it turns down, and at what cost? Does it say what is unknown, and what would change the writer's mind? Does it commit to a first step, and a date to check the result? The five checks are Focus, Relevant evidence, Alternatives, Manage uncertainty, and Execute and learn. Their first letters spell FRAME.

Here is such a page, dated September 1985. Grove and Moore never wrote it. We did, using only facts Intel had at the time.

F · Focus: what exactly is being decided?

> By the end of September we decide whether Intel stays in the market for DRAM or leaves it. Either way, we will call the decision a success if Intel makes an operating profit again within two years. Our other memory chips, such as EPROMs, are not part of this decision.

Before Grove's question, Intel had spent a year arguing about how to win in memory chips. An argument like that can go on for ever. With one question and one answer, Grove and Moore turned it into a yes-or-no choice. We added what their talk left out: a date, a test of success and a boundary. We need the boundary because Moore spoke of all memory chips, but the memo decides only about DRAM. The first lines of any decision memo need a yes-or-no choice with those three things.

R · Relevant evidence: which facts would change the answer?

> Only one of our eight factories still makes DRAM. Our processors have been inside IBM's personal computers since 1981. Our 286 runs IBM's most powerful machine, the PC AT, and our 386 is about to come out. We assume that personal computers keep selling and keep being built around our processors.

The first fact tells the most: the decision was already half made. Robert Burgelman, a Stanford professor who later studied the episode, found out how. Factory space at Intel was scarce. Middle

managers gave it to whatever earned the most profit per silicon wafer, the thin disc that chips are cut from. Under that rule, production had been moving from memory to processors for years. So the managers had started the move long before Grove and Moore decided on it.

A · Alternatives: what is rejected, and at what price?

> Our managers have proposed three ways to stay. One is a giant factory just for memory, to beat the Japanese on cost. The other two are a new technology they cannot copy and special memory for narrow markets. The factory is the strongest, because it is the only one that goes straight at our costs. It works only if a new plant can make memory for less than the Japanese sell it for. None of our plants can do that today. The factory would also take money and engineers the processors need.

In his book Grove lists all three. He believed less and less in narrow markets, because memory was becoming a commodity: one product at one price all over the world. In a narrow market you can keep a high price only until rivals crowd in.

M · Manage uncertainty: what is unknown, and what would change the decision?

> The open question is whether Intel can live on processors alone. If it cannot, we will have given up our oldest business with nothing big enough to replace it. Leaving is hard to undo: once the DRAM teams break up and the plant is converted, coming back would

> take years. We will reopen the decision if PC makers start designing their next machines around another company's processor. We will check for that every quarter.

The last two sentences matter most. In them the writers name news that has not come yet, and decide now what to do if it comes. So they choose their reaction while they are calm.

E · Execute and learn: what happens first, and when is it checked?

> We leave DRAM and announce it on 10 October. Andy Grove is in charge of the exit. At the end of 1987 we will compare our results with this page and note what we got right and wrong.

Intel did announce its exit on 10 October 1985. But neither Grove's book nor Burgelman's study mentions a date for comparing the result with what anyone expected. Intel learned from the decision all the same. Grove wrote the lessons down only eleven years later, in a memoir, when he knew the ending.

What the page is worth

One warning about this memo. Intel knew every fact in it in September 1985. But we wrote it knowing the ending, and that helped us choose which facts to put in. This is what Fischhoff found, and it happened to us. That is why the page has to be written before the ending and dated. A page written that September would be worth the same whichever way the processor market went. Either way, we would judge it on what the writer saw, what they assumed and which news they promised to watch for.

The page also makes the ending useful. From one ending you learn little about anyone's judgement. From a run of endings, each set against a page written beforehand, you learn much more. In forecasting tournaments, people write down forecasts with a date and a probability, and each is scored when the outcome arrives. Over hundreds of forecasts, you can tell skill from luck. From 2011 to 2015 IARPA, the research agency of America's intelligence services, ran such a tournament. The Good Judgment Project won it. Its leaders were Philip Tetlock, Barbara Mellers and Don Moore. Moore teaches at the University of California, Berkeley, and is no relation to Gordon. A manager makes far fewer decisions than a forecaster makes forecasts, and rarely gives them a probability. So the scoring is rougher. But the habit is the same: a dated page for every decision that matters, compared with the outcome when it comes.

A common mistake is to fill all five parts with empty phrases: "the framing here is strategic", "there is significant uncertainty", "alternatives were considered". Every part is filled in, but nothing has been checked. Under the right heading, an empty sentence looks like a finished check. So take the headings away and test each sentence. Could someone else check it against something outside the page? That might be a number, an option not taken, or news that would make the writer drop the recommendation. If nobody else could check a sentence, it is only decoration.

Where the five checks come from

FRAME is our own list, but each of its checks has a source. Focus comes from decision analysis, the field Ronald Howard of Stanford named in the 1960s. In decision analysis, you start by defining the

problem. Relevant evidence draws on Richards Heuer, a CIA analyst for about 45 years. He taught analysts to look for the few facts that tell one explanation from another.

Alternatives comes from Paul Nutt of Ohio State University. He followed 356 decisions in organisations, and about half failed. Failure was more common where a manager picked a solution early and stopped looking for others.

Manage uncertainty joins two ideas. In Gary Klein's premortem, a team imagines its plan has already failed and explains why. Chip and Dan Heath's tripwire is an event you agree on in advance: if it happens, you look at the decision again. The dated check in Execute and learn comes from the forecasting tournaments.

Using FRAME on the cases

Every case on this site stops where someone had to decide and asks what you would do. Write your page there, before the story goes on: a few lines for each of the five checks. The ending will not tell you whether you reasoned well. But your page will show what you saw, what you assumed and what would have changed your mind. You can compare that with what the company knew at the time.

An AI reviewer scores your answer on these five checks without seeing how the story ends. Each check gets 0 to 20 points, and the five add up to a rating out of 100. The rating is only advice: you can rewrite the answer and ask for a new rating. Below are the methods for each check; they open next to the page. Find the check you are stuck on. The last line adds two methods for checking a finished memo:

- **F:** clarity test, map is not the territory

- **R:** competing hypotheses, survivorship bias, kind and wicked learning environments
- **A:** inversion, steel-man, opportunity cost, second-order thinking
- **M:** outside view, pre-mortem, fragility audit, falsifiability
- **E:** two-way and one-way doors, commander's intent, decision journal
- **Before you send:** Decision Quality: six links, red team

References

- Andrew S. Grove, *Only the Paranoid Survive*, Currency Doubleday, 1996.
- Robert A. Burgelman, "Fading Memories: A Process Theory of Strategic Business Exit in Dynamic Environments", *Administrative Science Quarterly* 39:1 (1994), 24–56.
- Intel, "Farewell to DRAM" and "Financial Year 1986", Intel timeline.
- Baruch Fischhoff, "Hindsight ≠ foresight: The effect of outcome knowledge on judgment under uncertainty", *Journal of Experimental Psychology: Human Perception and Performance* 1:3 (1975), 288–299.
- Ronald A. Howard, "Decision Analysis: Applied Decision Theory", *Proceedings of the Fourth International Conference on Operational Research*, 1966.
- Richards J. Heuer Jr., *Psychology of Intelligence Analysis*, CIA Center for the Study of Intelligence, 1999.
- Paul C. Nutt, "Surprising but true: Half the decisions in organizations fail", *Academy of Management Executive* 13:4 (1999), 75–90.
- Gary Klein, "Performing a Project Premortem", *Harvard Business Review*, September 2007.

- Chip Heath and Dan Heath, *Decisive: How to Make Better Choices in Life and Work*, Crown Business, 2013.
- Philip E. Tetlock and Dan Gardner, *Superforecasting: The Art and Science of Prediction*, Crown, 2015.

Clarity test

F Focus · 6 min

When you need it

The question sounds big, but nobody can say what is being decided

Write the decision so plainly that someone else could later check whether you were right.

How to use it

1. Say who decides and what they are choosing.
2. Add the deadline and what the decision does not cover.
3. Say which result will count as success.

Try it in your answer

Rewrite the case question: who must decide what, by when, within what limits, and to get what result?

Practise it on a case

Booking.com question 1 · **Braun** question 1 · **Y Combinator** question 1

What it is

Imagine someone who knows everything: every fact, every number, everything that will ever happen. You hand them your question. Could they answer it with a plain yes or no, or give you a number, without asking what you mean? If they could, your question is clear. If they would first have to ask "what do you mean by that?" or "by when?", it is not clear yet. That is the clarity test.

For a business decision, write the question as one sentence. It names four things: who decides, what they will do, by what date, and which result will count as success. A question like that can be answered, and later you can check whether the answer was right. A vague topic cannot. A team can research it for months and never finish.

Where it comes from

The test comes from Ronald A. Howard, a professor at Stanford University. In 1964 he gave the field of decision analysis its name. He described the clarity test in 1988, in a paper called "Decision Analysis: Practice and Promise" in the journal Management Science.

To explain the test, Howard imagined a clairvoyant. This person knows the future: every newspaper that will ever be printed, every measurement that will ever be taken. Later Howard added that the clairvoyant knows everything that can be measured, but has no opinions. So the clairvoyant can tell you what happened, but not what you meant.

Howard used the test on every part of a decision: the options, the unknowns, the possible results and how you score them. A part passes when the clairvoyant could answer without asking a single question back. In 2004, in "Speaking of Decisions: Precise Decision Language", he asked for words about decisions exact enough to be checked. This guide uses the test on the sentence that states the decision. That is common practice, but it is not in the 1988 paper.

What it corrects

Trouble starts with a question that sounds important but does not name a choice. Take "Should we enter Europe?" A good team gets to work at once: it measures the market, looks at competitors, studies

prices and checks the legal side. But nobody has said what exactly will be decided. So nobody can tell which research is beside the point, or when to stop. The meeting finishes with a summary instead of a decision.

Working harder makes this worse. The deeper the research, the more solid the topic feels. Nobody notices until late that the decision itself was never written down. There is a second problem. Two people can agree on the answer to a vague question while meaning different actions. Afterwards nobody can say whether the decision was right, so the team learns nothing from it.

How it works

1. Write the choice as one sentence: who will do what, and by when.
2. Say what the decision does not cover, and what stays the same while you decide.
3. Define success as a number that someone could read in a named report on a named date.
4. Imagine handing the sentence to the clairvoyant. Could someone who knows every fact, but has no opinions, say whether that result happened? Fix every word they would have to ask about.
5. Check that at least two really different actions still fit the sentence. If only one fits, you have written a plan, not a decision.

Worked example

"Are hedge funds worth their fees?" is the kind of argument that can go on forever. Hedge funds are investment funds for wealthy clients, and they charge high fees. On 19 December 2007 the investor Warren Buffett turned it into a bet that could be settled. He registered it on the website Long Bets. The S&P 500 is an index of 500 large American

companies. Buffett bet that over ten years a fund that simply tracks it would beat a group of funds that invest in hedge funds. All fees, costs and expenses would count.

Every word of the bet could be checked. The index was named. The other side, the firm Protégé Partners, chose five funds that invest in hedge funds. Between them they held stakes in more than 200 hedge funds. The bet ran from 1 January 2008 to 31 December 2017. And fees, which such arguments usually leave vague, were written into the bet.

Buffett gave the result in his 2017 letter to the shareholders of his company, Berkshire Hathaway. Over the ten years the index fund grew 125.8%, or 8.5% a year on average. The five funds grew 21.7%, 42.3%, 87.7%, 2.8% and 27.0% in total, and one of them was closed in 2017. The prize was $1 million for a charity named by the winner. In 2012 the money was moved from US Treasury bonds into Berkshire shares, and it grew further. Girls Inc. of Omaha, the charity Buffett had chosen, received $2,222,279.

The clear terms did not help Buffett win. But they meant nobody could argue with the result. The measure, the period and the way fees were counted were fixed in advance. So neither side could read the result in its own favour.

In a Business Case Weekly case

The first decision in the prediction markets case is set in Iowa City in 1988. That year three economists at the University of Iowa started the Iowa Political Stock Market. Its 192 traders predicted George H. W. Bush's share of the vote to within 0.2 percentage points. The case asks whether you would put $10,000 into prediction markets as a business.

"Prediction markets as a business" is a topic, not a decision. To turn it into a decision you could be held to, answer four questions. Which company do you fund? Which regulator oversees it? Which step by that regulator, such as a licence or a penalty, do you need to see, and by when? What return, over how many years, would make the money worth it? The case is a good place to practise. A prediction market faces the same test with every contract it lists. Before anyone can trade, the contract must say where the answer will come from and by what date.

In your answer

- "The decision is whether [who] will [do what] by [date]. It does not cover ..."
- "Success means [number] on [date], as shown in [named report]."
- "Someone who knew every fact could settle this by checking ..."
- "Two really different options fit this question: ... and ..."

Common misuse

The hardest mistake to spot is being exact about the wrong thing. "Will we reach 40,000 downloads by March?" passes the test: the clairvoyant could answer it. But the answer tells you nothing about whether the business works. A quick check: say what you would do if the answer were yes, and what you would do if it were no. If you would do the same thing either way, the sentence is exact but useless, and the real decision is still missing.

References

- Ronald A. Howard, "Decision Analysis: Practice and Promise", Management Science 34(6), 679–695, 1988. The paper that sets out the clarity test. Short enough to read in one sitting.

- "Clarity test", Wikipedia. A one-page summary, with Howard's description of the clairvoyant and its sources.
- Warren Buffett, "2017 shareholder letter", Berkshire Hathaway, pages 11–12. The full terms of the ten-year bet and the final results.
- Ronald A. Howard, "Speaking of Decisions: Precise Decision Language", Decision Analysis 1(2), 71–78, 2004. The record and summary of the paper; the full text is behind a paywall.

Map is not the territory

F Focus · R Relevant evidence · 5 min

When you need it

The goal is a metric, and nobody has asked what it leaves out

Name what your metric or model leaves out, then check it against the real thing.

How to use it

1. Name the map your decision rests on: a metric, model, label or story.
2. Say what it only stands in for, and list what it leaves out.
3. Before you trust it, check it against the real thing: a look for yourself, or a metric that could disagree.

Try it in your answer

Name the map your argument rests on, what it does not measure and one direct check that could prove it wrong.

Practise it on a case

Nokia question 2

What it is

A road map shows roads, not potholes. It helps you plan a route, but cannot tell you whether a road is open today.

In business, a map is a short version of something bigger: a metric, a model, a label, a forecast, a story about a case. Each keeps a few features of the real thing and drops the rest. With this method you

name what was dropped and ask whether it changes your decision. Every decision needs maps, so you keep them and check how far to trust them.

Where it comes from

The phrase comes from Alfred Korzybski (1879–1950), a Polish-American writer who studied how words fit facts. On 28 December 1931 he gave a talk to the American Mathematical Society. It was part of the New Orleans meeting of the American Association for the Advancement of Science. The talk was called "A Non-Aristotelian System and its Necessity for Rigour in Mathematics and Physics". In 1933 he reprinted it as Supplement III of his book *Science and Sanity*. Of the talk's four numbered points, the third is one short sentence: "A map is not the territory."

In his example, Paris, Dresden and Warsaw lie in that order, west to east. Now draw a map that puts them as Dresden, Paris, Warsaw. Travelling by that map, he wrote, "would be misguiding, wasteful of effort", and "in case of emergencies, it might be seriously harmful".

His first point is just as important: a map may or may not share the territory's structure, such as the order of the cities. A good map is useful because what holds on paper also holds on the ground.

What it corrects

The dashboard says weekly active users are up 9%. A careful manager checks the sums and concludes that people use the product more. But the number only counts who came back at least once that week. It says nothing about what they did.

An analyst builds a number for one purpose, and it works there. Then people use it to settle a different question. Nobody is being lazy. The trouble is how a number is passed on. The analyst knows its limits, but people at the next meeting do not.

Being more careful does not help, because the care goes into the map. Reviewers ask whether the database query was right, whether the model runs, whether the slide adds up. Nobody's job is to ask what the number leaves out. And the longer a map has worked well, the less anyone thinks to ask.

How it works

1. Name the map your argument rests on: a metric, model, label, forecast or story.
2. Say what it measures directly, and what it only stands in for.
3. List what it leaves out, and mark what could change this decision.
4. Check it against the real thing: a look for yourself, a metric that could disagree, or an unusual case.
5. Say how big the gap would have to be to reverse your call. Then decide.

Worked example

Google Flu Trends estimated US flu levels from Google searches. Its builders tested 50 million search terms. They kept those that best matched 1,152 weekly figures from the US health agency, the Centers for Disease Control and Prevention (CDC).

The map worked, then went wrong. In 2009 it missed the H1N1 pandemic completely, because that flu came outside the usual season. The model had learned to spot winter more than flu. In February 2013 *Nature* reported that it estimated the share of doctor

visits for flu-like illness at more than double the CDC's figure. Four researchers, Lazer, Kennedy, King and Vespignani, then checked it. From 21 August 2011 to 1 September 2013 it was too high in 100 of 108 weeks. It overshot the 2011–2012 season by more than 50%.

The fix is the useful part. The same researchers found that a forecast built on the CDC's own figures, though they were two or three weeks old, tracked current flu better than the live search model. The average miss, the mean absolute error, was 0.311 for that forecast and 0.486 for the searches. Combining the two beat both, at 0.232. The searches held a real signal, but it became reliable only when tied to a direct count of flu.

In a Business Case Weekly case

In the Tether case, the company told holders of its tokens that a dollar backs each one. At the fork, set in 2018–2021, you sit on the board. The reserves now include receivables (money owed to the company) and other non-cash assets.

"Backed one to one" is a map of whether the company can pay. To see the territory, ask five questions. How good are the assets? When will they turn into cash? Who holds them? Is too much in one place? How fast could holders cash out?

List those five before you rank the options. You can stop issuing tokens, disclose the gap or carry on with attestations, the accountants' reports on the reserves. Say which of the five would change your answer. The choice is still yours.

In your answer

- "The figure here is X. It stands in for Y, and it does not show Z."

- "The missing part that could reverse this call is ..., and it would have to be at least ... to matter."
- "Before treating this number as the business, I would check it against ..., which comes from the territory, not the model."
- "This map is accurate enough for the decision about ..., but not for the decision about ..."

Common misuse

The fake version is the statistician George Box's line "All models are wrong". It sounds modest. But it names no missing part, and the recommendation stays where it was. Ask one question: after you said it, did your evidence or your confidence change? If you cannot name the missing part, the decision it affects and the check you would run, you have only quoted Box.

The opposite mistake is throwing out a map that is good enough. A model that leaves out something irrelevant to this decision needs no fixing. Saying so is also an answer.

References

- Alfred Korzybski, *Science and Sanity: An Introduction to Non-Aristotelian Systems and General Semantics* (1933), first edition on the Internet Archive. Supplement III is the 1931 talk.
- David Lazer, Ryan Kennedy, Gary King and Alessandro Vespignani, "The Parable of Google Flu: Traps in Big Data Analysis", *Science* 343 (14 March 2014), 1203–1205, open copy in Harvard's DASH repository. Three pages, the source of every figure above.
- Institute of General Semantics, "Alfred Korzybski". His biography, with the date of the 1931 talk.

Competing hypotheses

R Relevant evidence · A Alternatives · M Manage uncertainty · 6 min

When you need it

Several explanations fit the same facts

Try to disprove each explanation, instead of collecting support for your favourite.

How to use it

1. Write down the explanation you favour, a rival one and a surprising one.
2. Check each fact against every explanation and look for contradictions.
3. Keep the explanation that survives your attempts to disprove it, and name the next fact to check.

Try it in your answer

Name two explanations for the main fact in the case, and one piece of evidence that would tell them apart.

Practise it on a case

Prediction Markets. Kalshi & Polymarket question 2 · **Singapore** question 3 · **n8n** question 3

What it is

Growth has slowed, and the right move depends on why. Maybe demand is weak. Maybe new customers find the product hard to start using. Maybe the company cannot keep up with demand. Each explanation fits the same facts.

Analysis of competing hypotheses is a way to choose between possible explanations like these. You write down every one that could fit, and list your evidence and assumptions together. Then you check each item against every explanation for contradictions. You keep the one the evidence contradicts least, even if another has more facts in its favour.

Where it comes from

In the late 1970s Richards J. Heuer, Jr. built the method in the CIA's Directorate of Intelligence. There he ran the methods team of the political analysis office until he retired in 1979. The method first appeared in public in summer 1981, as an article of the same name in the CIA journal *Studies in Intelligence*.

Heuer set out the full method in Chapter 8 of his 1999 book *Psychology of Intelligence Analysis*. It collects papers he wrote for the Directorate from 1978 to 1986, some on contract after he retired. In it, he calls the method "an eight-step procedure grounded in basic insights from cognitive psychology, decision analysis, and the scientific method." Below, his last two steps are merged into one: report the conclusions and name what to watch. In 2010 Heuer and Randolph Pherson listed the method among fifty techniques in *Structured Analytic Techniques for Intelligence Analysis*.

What it corrects

Heuer calls the mistake satisficing, a decision-analysis term for stopping at the first answer that seems good enough. Early on, you sense the likely answer, then go through everything you have, looking for support. You usually find it, and the more you have, the surer you feel.

Laziness is not the problem. The same facts also fit the explanations you never wrote down. So when they agree with your favourite, you learn almost nothing about whether it is true. Working harder makes this worse. A careful person gathers more evidence, but most of it fits every explanation. You feel surer, yet you have tested nothing.

How it works

1. List the explanations, even unlikely ones, and ask people who see things differently to add more.
2. List the evidence and arguments in one place, with your assumptions about what people want and can do.
3. Build a matrix: a table with explanations across the top and evidence down the side. In each cell, mark whether the item fits that explanation, contradicts it or is irrelevant.
4. Cross out evidence that fits every explanation: however solid, it cannot help you choose. If nothing on the list separates two explanations, merge them.
5. Go down each column and try to disprove that explanation. Rank the explanations by the evidence against them.
6. Test the two or three items your ranking rests on. What if one is wrong, misleading or open to another reading?
7. Report where each explanation stands and what could change the ranking later.

Worked example

In 2000 Iraq's Military Industrialization Commission ordered 60,000 aluminium tubes of alloy 7075-T6. Each was 81 mm across the outside and 900 mm long, with 3.3 mm walls. A first batch was stopped in Jordan in 2001.

The CIA's weapons centre believed the tubes were rotors for the centrifuges that enrich uranium. Analysts at the Energy Department's laboratories and the State Department's intelligence bureau disagreed. They believed the tubes were combustion chambers for the Nasser-81 rocket. Iraq had begun copying it from the Italian Medusa rocket in 1986 and declared it to inspectors in 1996. Their objections were printed in the October 2002 National Intelligence Estimate.

Most of the evidence fitted both explanations. The alloy was no help: the Medusa and the American Hydra-70 rocket are both made of 7075-T6. The nuclear explanation's strongest point was the tolerances, how exact each tube had to be: more exact than a rocket body seemed to need.

Inspectors from the IAEA, the UN's nuclear agency, started inspections in Iraq again on 27 November 2002. They looked for a fact that would separate the two explanations. On 27 January 2003 Mohamed ElBaradei, the agency's director general, reported to the UN Security Council. He said the tubes "would be consistent with the purpose stated by Iraq" and, "unless modified, would not be suitable for manufacturing centrifuges."

The deciding fact came when inspectors found and questioned the Iraqi who had set the tolerances. He had tightened them so the rocket would not stick in its launch tube. In a test, one had stuck and exploded. On 7 March 2003 ElBaradei told the Council that inspectors had learned from documents and interviews why the tolerances changed. That same day Colin Powell, the US Secretary of State, still cited them to the Council. He called them "more exact, by a factor of 50 percent or more, than those usually specified for rocket motor casings."

After the 2003 invasion, the Iraq Survey Group searched Iraq for banned weapons and sided with the inspectors. The nuclear case rested almost entirely on evidence that fitted both explanations.

In a Business Case Weekly case

In the Nokia case it is 2006 and you sit on Nokia's executive board. Nokia sells more phones each year. Rivals envy its low logistics costs. It holds over a third of the world market. The obvious reading: the company is strong. But a small team inside Nokia disagrees. The team says these numbers measure the phone business you built, not the one software is about to create.

Here, competing hypotheses means checking these numbers against both readings. Every number fits both, so none helps you choose. Instead, ask which single observation would look different under each reading. Does the company already track it? Which reading is right is still your call.

In your answer

- "Two explanations fit these facts: ... and ... ."
- "This number fits both, so I am not counting it as support."
- "Evidence against my own explanation would be ..."
- "Next I would look for ...: it points one way under the first explanation and the other way under the second."

Common misuse

The usual mistake is to turn the matrix into arithmetic: score every cell, add the columns, hand in the lowest total. When you add up scores, you skip the two judgements the method is for. Is this source reliable? Does this item really tell the explanations apart? The second

mistake is a rival added for show, so weak that it never had a chance. You can catch both with one check. Change the mark in a single cell. If that makes you rethink the ranking, the matrix is working. If it only changes a total, you have built a spreadsheet. If no mark in a rival's column could ever lift it in the ranking, the rival is just for show.

References

- Richards J. Heuer, Jr., *Psychology of Intelligence Analysis*, Center for the Study of Intelligence, CIA, 1999. Chapter 8 is the method: about twenty pages, an evening's read.
- David Albright, *Iraq's Aluminum Tubes: Separating Fact from Fiction*, Institute for Science and International Security, 5 December 2003. Both readings of the tubes, point by point.
- Randolph H. Pherson and Richards J. Heuer, Jr., *Structured Analytic Techniques for Intelligence Analysis*, CQ Press, 2010. No stable public copy found.

Survivorship bias

R Relevant evidence · M Manage uncertainty · 5 min

When you need it

The argument rests on a famous winner's example

Before you learn from the winners, count everyone who tried, including those who failed.

How to use it

1. Count everyone who tried the move, not only the winners you can see.
2. Name why the failures went missing.
3. Keep only the lessons that hold up when you compare the winners with the failures.

Try it in your answer

Name the group missing from the comparison, and one conclusion you become less sure of once you count it.

Practise it on a case

Xiaomi question 2 · **Dyson** question 4 · **Y Combinator** question 5

What it is

Imagine ten famous companies that made the same bold move and did well. What about those that tried it and failed? You never hear about them, so you never learn from them. Learning only from the winners you can still see is called survivorship bias.

To correct it, count everyone who tried, as if the results were not yet in. Then you learn two things. One is an honest base rate: how many of those who tried succeeded. The other is whether the failures had the winners' famous trait too. If so, it cannot explain the success.

Where it comes from

The idea has no single author, and nobody knows when the name was first used. The statistician Abraham Wald (1902–1950) was the first to study it carefully. In the Second World War he worked in the Statistical Research Group at Columbia University. In 1943 he wrote "A Method of Estimating Plane Vulnerability Based on Damage of Survivors". It was secret then; the Center for Naval Analyses reprinted it in 1980.

Wald could count damage only on the planes that came home. From them, he estimated the damage across every plane that flew. His reasoning: a plane hit in its weakest spot is the least likely to come home. So if returning planes show few hits in one place, hits there were bringing planes down.

In the popular version of the story, Wald tells the air force to armour the places with no bullet holes. That version came later. The paper itself is a set of equations for making estimates.

The term became common in finance in the 1990s. In 1992 four researchers, Brown, Goetzmann, Ibbotson and Ross, showed what goes wrong when data keeps only the investment funds still in business. In such data, fund results can look predictable when they are not.

What it corrects

Say you are deciding whether to copy another company's move: a pricing method, a hiring standard, a new product line. Your evidence is the companies that made the move and still exist. You study these winners carefully, find what they share and treat those traits as causes.

The mistake is hard to see because the list looks complete. You cannot tell from it that companies were left out because they failed. And success or failure is the very thing you want to explain. Working harder makes this worse. The closer you study the survivors, the more their shared trait stands out. You grow surer, but your answer is no more accurate. You can fix it only with evidence from outside the group of winners.

How it works

1. Turn the claim into a rate: of everyone who tried this, how many got the result?
2. Count everyone who tried, as if the results were not yet in.
3. Name why companies go missing: bankruptcy, a quiet sale, leaving the stock exchange, a cancelled project, or a story too dull to write about.
4. Find what you can about the missing group, even roughly: company registers, yearly counts of new firms, an investor's whole portfolio.
5. Check whether the winners' trait was also common among the failures.
6. Keep only the conclusions that pass that check, and say what you could not find.

Worked example

In October 1982 two McKinsey consultants, Tom Peters and Robert Waterman, published *In Search of Excellence*. They asked partners at their firm and other business people who was doing impressive work, and got 62 names. After testing each company's results, they kept 43. General Electric was on the first list but was dropped: only results counted. The book's eight principles, from "a bias for action" to "stick to the knitting", came from what the 43 had in common.

The authors chose the companies by their results. Any company that followed the principles but had a bad decade failed the tests and was left out. So nobody could tell from the study whether the principles caused those results. Several of the 43 later struggled. One was Wang Laboratories: at its peak in the 1980s, its yearly sales were almost $3 billion. In August 1992 it filed for bankruptcy protection.

That does not make the book worthless. In 2002 Forbes found that after the book came out, shares in the 43 still earned about 14.1% a year on average. Shares in Dow Jones index companies earned 11.3%. But nobody could prove from the book that the principles caused the success: the other half of the evidence was missing.

In a Business Case Weekly case

In the Y Combinator case, you hear first about a few famous startups that Y Combinator funded. In all, it funded 5,668 companies. It put about $1.5 billion into them, and together they are worth roughly $600 billion. One question is whether bigger batches, the groups of startups that go through the programme together, weaken Y Combinator's model.

To correct for survivorship bias, argue from success rates across all 5,668 companies, not from the few graduates everyone can name. Also ask what the median company got from the programme: the one in the middle when all 5,668 are ranked. Then you know which evidence would settle the question. The answer is still yours.

In your answer

- "These cases were picked for how they ended: the list was made after the results were known."
- "About N companies tried the same move. The ones missing from the story are ..."
- "The same trait shows up in the failures I can find, so it cannot explain the success. At most, it is a requirement."
- "I cannot rebuild the missing group here, so I am unsure of this conclusion. I would test it by ..."

Common misuse

A fake fix is to name the bias and carry on. Someone writes "there is obviously survivorship bias in these examples". Then they draw the same conclusion anyway. The opposite mistake is to refuse every lesson because of the bias, as if a successful company proved nothing. Ask yourself: did your base rate change, and by how much? If no number or conclusion changed, you have not done the work.

References

- Abraham Wald, "A Method of Estimating Plane Vulnerability Based on Damage of Survivors" (1943), reprinted by the Center for Naval Analyses, 1980. The original paper; hard reading.

- Stephen Brown, William Goetzmann, Roger Ibbotson and Stephen Ross, "Survivorship Bias in Performance Studies", *Review of Financial Studies* 5(4), 1992, 553–580.
- Mark Carhart, Jennifer Carpenter, Anthony Lynch and David Musto, "Mutual Fund Survivorship", *Review of Financial Studies* 15(5), 2002, 1439–1463. Measures the bias: 0.07% a year in one-year samples, about 1% a year in samples longer than fifteen years.
- Jordan Ellenberg, *How Not to Be Wrong: The Power of Mathematical Thinking*, Penguin Press, 2014. The book opens with the Wald story.

Kind vs wicked learning environments

R Relevant evidence · M Manage uncertainty · 5 min

When you need it

The argument rests on experience: “we have done this before, and it worked”

Before you learn from a result, separate what your choice caused from luck, timing and which cases were picked in the first place.

How to use it

1. Name the feedback the decision will give, and when it will arrive.
2. Separate what your choice caused from luck, timing and which cases were picked in the first place.
3. If you cannot learn clearly from the real result, try a smaller version you can repeat.

Try it in your answer

Write what this result can teach, what it cannot tell apart, and how you would get feedback sooner.

Practise it on a case

Klarna question 4

What it is

Think of learning darts. You throw, see at once where the dart landed, adjust and throw again. The same choice repeats. Each result comes fast and clear, and you can see which throw caused it. That is a

kind learning environment. Now think of buying companies. You do it a few times in a career, and each result comes years later, mixed up with luck. That is a **wicked** one. An environment is also wicked when you only see the results of options you chose before.

Before you trust a result, ask which kind of environment produced it, and so how much it can teach you.

Where it comes from

Robin Hogarth, a psychologist, ran the University of Chicago's Center for Decision Research from 1983 to 1993. In his book *Educating Intuition* (2001), he asked when experience gives us intuition we can trust. His answer: it depends on the environment. Make the same kind of judgement many times with fast, accurate feedback, and you build real skill. With missing, slow or distorted feedback, you build only confidence.

In 2015 Hogarth and two colleagues, Tomás Lejarraga and Emre Soyer, made the idea more precise. An environment is kind, they wrote, when the information you learn from matches the information you need when you act. Soyer and Hogarth then wrote *The Myth of Experience* (PublicAffairs, 2020) for general readers. David Epstein's book *Range* (2019) brought the words "kind" and "wicked" into business writing, where most people meet them now. But the idea is Hogarth's.

What it corrects

A bank turns down a loan and never learns whether the person would have paid it back. A company walks away from a deal and never learns what the deal would have earned.

A good team studies its results, updates its views and repeats what worked. In a kind environment, that is right. In a wicked one, the team sees few results, sees them late, and sees only the results of options it chose. Working harder cannot fix this. The team checks the data that arrived, but the missing part is the data that never arrived. The more carefully it studies this one-sided data, the surer it gets of a wrong belief. In the end the lesson becomes a rule nobody in the room can argue with, because everyone saw it work.

How it works

1. Write down what the result is supposed to prove.
2. Name the feedback that would test that claim, and when it will arrive.
3. Ask whether this decision repeats in similar conditions, or happens only once.
4. List what you cannot separate in the result: luck, timing, which cases were picked in the first place, and your own influence.
5. Get kinder feedback: faster and clearer. Try a smaller version you can repeat, an earlier step you can measure, or a forecast written down before the result. Set the date to check it.

Worked example

The team behind Bing, Microsoft's search engine, changed its learning environment on purpose. Ronny Kohavi, who ran Bing's experiment platform, wrote in August 2013 that "on a typical day, there are over 250 experiments running on Bing". He added that "when ideas are evaluated objectively in a controlled experiment, less than a third move the metrics they were designed to improve". In such a test, some users see a change and others do not.

Before testing, managers judged ideas in a wicked environment. When a new feature went live, its effect on revenue was mixed up with the season, competitors and every other change that week. So nobody learned whose ranking of ideas had been right.

Here is what that cost. An employee suggested longer ad headlines, made by joining the title to the first line of text below it. The program managers rated the idea low, and it sat unbuilt for over six months. Then an engineer ran it as a controlled experiment. Revenue rose 12%. In Harvard Business Review, Kohavi and Stefan Thomke valued the gain at more than $100 million a year in the United States alone.

What changed was the environment. Each idea now got its own result within days. Rankings that had felt reliable could finally be checked. Most ideas did not move what they were meant to, and one idea the managers had rated low lifted revenue by 12%.

In a Business Case Weekly case

In the Y Combinator case, the accelerator, which funds start-ups in batches, backed about 400 companies in Winter 2022. It then cut the next batch by 40%, to 250. Investors in its funds were asking whether the bigger batches had hurt quality.

You cannot answer them with this method. Instead, ask what evidence ever could. A batch's real results take most of a decade to appear. The case reports that four companies hold more than 84% of the stock-market value Y Combinator has created. So its whole record rests on four companies. The Y Combinator name also makes its picks easier to fund, so its choice partly creates their success.

Of its companies, 4.5% became unicorns, start-ups worth over $1 billion. They were picked over twenty years, so you cannot judge this year's picks by them. At the fork, name earlier measurements you

can repeat, to see before 2030 whether batch quality is changing.

In your answer

- "This result supports ..., but it cannot separate that from timing or selection, because ..."
- "The first feedback that would test the claim is ..., and it arrives by ..."
- "This decision happens once, so rather than wait for its outcome I would test a smaller version I can repeat: ..."
- "I am writing down the forecast now (...), so that on ... we can compare the result with it."

Common misuse

The first mistake is to stop at a label. Someone writes "this is a wicked environment", decides that nothing here can be known, and leaves the plan as it was. That costs nothing and changes nothing.

The opposite mistake is to call an environment kind because a number moves fast. The fast number may measure something other than the decision: clicks instead of customers who stay, or signed deals instead of revenue.

You can check for both the same way. Name a measurement, a date and a result that would change what you do. If your diagnosis leads to no such line, you have not used the method.

References

- Robin M. Hogarth, *Educating Intuition* (University of Chicago Press, 2001). Where the idea is first set out.

- Robin M. Hogarth, Tomás Lejarraga and Emre Soyer, "The Two Settings of Kind and Wicked Learning Environments", *Current Directions in Psychological Science* 24(5), 2015, pages 379–385. A later, more precise version.
- Ronny Kohavi, "Large Scale Experimentation at Bing", Bing Search Quality Insights, 8 August 2013. The source of the figures above; fifteen minutes.
- Ron Kohavi and Stefan Thomke, "The Surprising Power of Online Experiments", *Harvard Business Review*, September–October 2017. The headline experiment in full, and how a company builds a kinder feedback loop.

Inversion

A Alternatives · M Manage uncertainty · 5 min

When you need it

Every option comes back as the same safe answer

Ask what you would do to make the plan fail, then remove those causes.

How to use it

1. Describe the failure as a concrete event: what happened, to whom, by when.
2. List what you would do to make that failure almost certain.
3. Take the biggest cause you control and remove it, reverse it or watch it.

Try it in your answer

List three ways to make your recommendation fail. Then show which one you remove with your plan.

Practise it on a case

Substack question 2 · **Novo Nordisk** question 3 · **SEGA** question 4

What it is

Suppose you are planning a product launch. Instead of asking how to make it work, you ask what you would do to make it fail. Then you remove, reverse or watch each cause. That is inversion: turning the question upside down. It helps most when asking how to succeed brings no new ideas.

Where it comes from

The method rests on a saying, "man muss immer umkehren": one must always invert. It is credited to the German mathematician Carl Gustav Jacob Jacobi (1804–1851). He is said to have given this advice to students choosing a research problem. But the story is second-hand. The earliest known record is a 1916 speech by the mathematician Edward B. Van Vleck to the American Mathematical Society. He said: "The great mathematician Jacobi is said to have inculcated upon his students the dictum: Man muss immer umkehren." Van Vleck names no source. His "is said to have" means he was retelling a story, not quoting a text. How Jacobi worked is better recorded. He turned the elliptic integral inside out and got the theory of elliptic functions, published in 1829.

The investor Charlie Munger brought the idea into business. On 13 June 1986 he spoke at the graduation of the Harvard School, a school in Los Angeles. He told the class he could not say how to be happy. So, as the TV host Johnny Carson once did in a speech, he would describe how to be sure of misery. His recipe: mood-altering chemicals, envy, resentment, being unreliable, learning only from your own experience and staying down after a setback. "The great algebraist, Jacobi, had exactly the same approach as Carson," he said. Most business writing on inversion goes back to that talk.

What it corrects

Ask a team how to make a plan work, and people suggest moves that sound reasonable. Things sound reasonable when people in the room already agree with them, so the same short list keeps coming back. Nobody asks how the plan could fail, so nobody names the ways it could.

Working harder does not fix this. Careful people check the plan more closely against the risks already written down, but the dangerous ones were never written down. Turn the question round, and different answers count as good. When the goal is failure, pointing out something the plan quietly relies on is a good answer, not an embarrassment. And those who know the plan's weak spots get a reason to say so.

How it works

1. Write the failure as an event with a date: what happened, to whom, by when.
2. Ask the upside-down question out loud: what would we do to make that failure almost certain?
3. Write the answers as actions, dependencies and rewards, not adjectives. "Weak ownership" becomes "nobody named as owner after launch".
4. Mark each failure condition as inside or outside your control.
5. Take the biggest one you control. Remove it, reverse it or set a tripwire on it: a warning sign agreed in advance.
6. Say what you found this way that your first plan had missed.

Worked example

Bank supervisors made inversion a rule. A normal stress test runs forward: pick a bad scenario, run the bank through it, check whether its capital lasts. The European Banking Authority (EBA) added the reverse test in guidelines EBA/GL/2018/04, issued on 19 July 2018 and applying from 1 January 2019.

In these guidelines, a reverse stress test "starts from the identification of the pre-defined outcome (e.g. points at which an institution business model becomes unviable ...) and then explores scenarios and circumstances that might cause this to occur." Put simply: decide what failure looks like, then look for its causes. Under paragraphs 84 to 91, banks must do this regularly, not once. Paragraph 90 goes furthest. If the risk of the bank's business model failing is unacceptably high, it should plan its response and consider changing the model itself.

What matters is the size of the shock needed to break the bank. In the guidelines' words, "the plausibility of the parameter shifts required to reach the pre-defined outcome gives a first idea about possible vulnerabilities". One bank fails only in a once-in-a-century crisis, another after one bad quarter. You can tell them apart only by asking the upside-down question. In the 2018 guidelines the EBA recalls its 2013 review of national supervisors: "very few competent authorities required reverse stress testing."

In a Business Case Weekly case

In the SEGA case you sit on the board in late 1994. SEGA has three game platforms on sale and a fourth about to launch. SEGA's American and Japanese divisions pull in opposite directions. Ask how to make this work, and the obvious answer is better coordination. Then turn it upside down. What would you set up to make sure the new console fails? Who reports to whom, what are people rewarded for, which budget do they share? Which do you already have? The answers are about how SEGA is organised. The second half of the case's question asks you to change exactly that.

In your answer

- "This decision fails predictably if ..., and the quickest way there is ..."
- "To make it fail, I would ...; our plan already does this in one way: ..."
- "Of these failure conditions, we control ...; we can only watch ..."
- "I would remove ... now, and treat ... as the sign that the one we cannot control has started."

Common misuse

The usual mistake is an ordinary list of risks, written more gloomily. You rank everything that could go wrong by how bad it would feel, and change nothing in the plan. The opposite mistake is trying to remove every condition on the list, even ones no company could control. You are done when a specific decision has changed. To check, point to the sentence in your plan that changed because of the upside-down list. If none changed, you have written pessimism.

References

- Edward B. Van Vleck, "Current tendencies of mathematical research", *Bulletin of the American Mathematical Society* 23(1), 1916, pp. 1–13 — Jacobi's saying on page 3, with the mathematics behind it.
- Charlie Munger, Harvard School commencement address, 13 June 1986 — the method used, not described; short enough to read tonight.
- European Banking Authority, *Guidelines on institutions' stress testing* (EBA/GL/2018/04), 19 July 2018 — section 4.6.5 and paragraphs 84 to 91 make inversion a duty.

Steel-man

A Alternatives · E Execute and learn · 5 min

When you need it

The rejected option is described so that it is easy to reject

Write the strongest honest version of the option you plan to reject, so that its supporters would say “yes, that is my argument”.

How to use it

1. Write the other option's goal and best evidence in its supporters' words.
2. Write what would have to be true for that option to win.
3. Explain exactly why your choice still wins, despite what you give up.

Try it in your answer

Write two fair sentences for the best option you reject. Then defend your choice against them.

Practise it on a case

Costco question 1 · **Ryanair** question 1 · **Spotify** question 1

What it is

Imagine the person who most wants the option you are about to turn down. They read your description of it. Would they say "yes, that is my argument"? If so, you have built a steel man: the strongest honest version of an option you plan to reject.

You build it before you reject anything. It has three parts: the goal that option really serves, its best evidence and what would have to be true for it to win. This is not just about fairness. Your choice has to beat the best rival, not a weak description of it.

Where it comes from

The word has no single author. It was made as the opposite of the much older "straw man": a weak version of someone's argument, easy to knock down. The earliest example in the Word Spy dictionary is a 2003 forum post. Luke Muehlhauser defined it for debate and made it popular. On 24 October 2011, writing as lukeprog, he published "Better Disagreement" on the website LessWrong. In a footnote he defined a steel man as "an improvement of someone's position or argument that is harder to defeat than their originally stated position or argument." In December 2012 Chana Messinger took the idea further in "Knocking Down a Steel Man", on her blog *The Merely Real*.

The habit is older than the word. In *On Liberty* (1859) the philosopher John Stuart Mill wrote that someone who knows only his own side of an argument knows little of it. Of the other side's arguments he wrote: "He must be able to hear them from persons who actually believe them; who defend them in earnest, and do their very utmost for them. He must know them in their most plausible and persuasive form." The game theorist Anatol Rapoport wrote rules for criticism. The first is to restate the other view "so clearly, vividly, and fairly that your target says, 'Thanks, I wish I'd thought of putting it that way'". The philosopher Daniel Dennett reprints the rules in *Intuition Pumps and Other Tools for Thinking* (2013). Philosophers have a related idea, the principle of charity: read an argument in its most sensible form. Neil L. Wilson named it in 1958–59. Willard Quine and Donald Davidson developed it.

What it corrects

In a recommendation you describe the options you turn down in your own words. You judge them by your goal. But each was built for a different goal, so it loses before anyone looks closely. Even honest people make this mistake.

Nothing here feels like cheating, so careful people miss it too. You check the sums, but the comparison was unfair from the start. And the mistake helps you: the weaker the rival looks, the surer your conclusion sounds.

How it works

1. Name the option you plan to reject. If there are several, pick the strongest, not the easiest to beat.
2. Write its goal in its supporters' words. What do they want to protect?
3. Give it its best two or three facts, from the sources you used for yours.
4. Write what would have to be true for it to win.
5. Check that against what you know, and say how you know it.
6. Argue your option against this version. Write down what changed.

Worked example

Charles Darwin argued against himself in print. Chapter VI of *On the Origin of Species* (1859), "Difficulties on Theory", opens: "Long before having arrived at this part of my work, a crowd of difficulties will have occurred to the reader. Some of them are so grave that to this day I can never reflect on them without being staggered." He then sets out the objections at full strength, among them the missing in-between forms and the eye's origin.

Under attack, he kept the habit. In 1871 the biologist St George Jackson Mivart published *On the Genesis of Species*. Mivart argued that natural selection could not explain the early stages of a useful body part. Half an eye sees nothing, so why would it survive? Darwin did not restate the objection in an easier form. In the sixth edition, finished in January 1872, he added a new Chapter VII: "Miscellaneous Objections to the Theory of Natural Selection". Most of it is about those early stages. Because he kept the objection in Mivart's terms, he had to answer it precisely. His finished theory carries the marks of that answer.

In a Business Case Weekly case

In the Rolex case, a question under "Scarcity Meets Its Limits, 2015-2022" is set in 2021. On the resale market, Rolex watches sell for roughly three times the shop price. Rolex could make about 30 per cent more watches, raise shop prices by about 30 per cent or start a programme to certify second-hand watches. You are asked to argue for the option Rolex rejected most firmly.

Without a steel man, it is easy to call one option greedy and another timid. In fact each has a real goal, a sensible explanation of the price gap and a situation where it would be right. Build all three at full strength. Only then choose, and say what you give up.

In your answer

- "The strongest argument for the option I reject is that it protects ..., and its best evidence is ..."
- "Its supporters would say the goal is not ... but ..."
- "That argument wins if ... is true. Here it is not, because the case reports ..."

- "Building that steel man changed one thing in my answer: ..."

Common misuse

The common mistake is a polite paragraph. You give a fair-sounding account of the rival, then write the conclusion you would have written anyway. It looks like a steel man but changes nothing. To check, ask what the steel man changed in your recommendation. A number? A condition? How sure you are? When you would revisit the decision? If nothing changed, and you cannot name evidence that would have changed it, you complimented the rival instead of comparing.

The rarer mistake is the reverse. You improve the rival into a version nobody in the room holds. You defeat it, and the real objection still stands. Stick to the version a real supporter would sign.

References

- John Stuart Mill, *On Liberty*, chapter 2, "Of the Liberty of Thought and Discussion" (1859). Why you should know the other side at its strongest. An evening's read.
- lukeprog (Luke Muehlhauser), "Better Disagreement", LessWrong, 24 October 2011. The post that made the term popular, with the definition in a footnote.
- Charles Darwin, *On the Origin of Species*, chapter VI, "Difficulties on Theory" (1859). The worked example, in Darwin's own words.
- Daniel C. Dennett, *Intuition Pumps and Other Tools for Thinking*, W. W. Norton, 2013, chapter 3, on Rapoport's rules for criticism. No link: no authorised full text is online.

Opportunity cost & trade-offs

A Alternatives · F Focus · 5 min

When you need it

You can say what the choice gains but not what it gives up

Say what you give up with your choice, not only what you gain.

How to use it

1. Name the scarce resource and what it is for.
2. Compare your choice with the strongest realistic other use of the same resource.
3. Write the trade-off in one sentence, and say what would make you swap the ranking.

Try it in your answer

Finish the sentence: “I give up ___ to get ___. I would swap the ranking if ___.”

Practise it on a case

Intel question 1 · **Rolls-Royce** question 1 · **Telegram** question 1

What it is

A factory that spends this week making chairs cannot also make tables. The profit it could have made on tables is the real cost of the chairs. Economists call this opportunity cost: the value of the best thing you give up when you spend a scarce resource.

A trade-off is what you agree to lose for your choice. You might give up speed to lower risk, customers to earn more on each one, or profit now for a stronger position next year. Write down both sides: what

you gain and what you give up. Compare your proposal with the best other use of the resource, not with doing nothing.

Where it comes from

The idea has no single author. The Austrian economist Friedrich von Wieser (1851–1926) built the idea, which is known in English as alternative cost. Using a resource, he said, costs the value of its best other use. He set it out in an 1884 study and developed it in *Der natürliche Werth* (1889). That book came out in English as *Natural Value* in 1893.

In January 1894 the economist David I. Green gave the idea its English name, opportunity cost, in the *Quarterly Journal of Economics*.

But the idea is older than Wieser. In 1850 the French economist Frédéric Bastiat wrote the pamphlet *Ce qu'on voit et ce qu'on ne voit pas*. A weak analyst, he argued, counts only the visible effect of a decision. A good one also counts what the same money would have done elsewhere.

"Trade-off" is everyday business English that no scholar coined. Managers use the two ideas together out of habit, not because of any theory.

What it corrects

A proposal reaches the board, and its numbers beat the lowest return the company accepts. But the team has compared the project only with doing nothing. Good teams do it too, because accounts work that way: money out, money in, time to pay back.

The accounts have no line for the second-best use of the same factory, the same twelve engineers or six months of top managers' attention. Careful teams check the project's numbers more closely, but they never add the missing line. So a company can approve twenty proposals that each make money and still spend its scarcest resource on its third-best use. Nobody asked what else that resource could be used for.

How it works

1. Name the scarce resource the decision uses up: money, factory capacity, engineers' time, top managers' attention, customers' trust or the freedom to change course later.
2. Say what the resource is for, and over what period you will judge the result.
3. Name the strongest realistic other use of the same resource. Keeping it in reserve counts too, if you really can.
4. Score both uses against the same goal over the same period.
5. Write the trade-off in one sentence: what you give up to get the advantage you chose.
6. Say what would make you swap the ranking, and where you would see it.

Worked example

Intel began by making memory chips. In 1970 it put the first DRAM memory chip on sale. By the early 1980s Japanese makers were winning that market on price and volume.

Intel was short of factory capacity, not cash. Each week its plants could start only so many wafers, the silicon discs that chips are made on. Intel's planners shared out that capacity by one fixed rule: earn

the most profit from each wafer. So they weighed each DRAM wafer against what it would earn as microprocessors. Wafer by wafer, microprocessors won. DRAM output shrank, though no meeting decided that it should.

Robert Burgelman, a Stanford professor, later studied the episode. He found that under this rule production had moved to microprocessors before the official strategy changed. On 10 October 1985 Intel announced it was leaving DRAM. The exit was costly: in 1986 Intel cut 7,200 jobs and closed eight older plants. Yet it spent $228 million on research and development, 18% of that year's revenue.

Andy Grove, then Intel's president, later called leaving DRAM "the best business decision we ever made". Intel's own timeline quotes him. The planners' rule gave that answer years before the official strategy changed. It weighed each wafer against its best other use, not against an idle plant.

In a Business Case Weekly case

In the Arm case, you decide in 2025 how much of the finished chip Arm should design itself. Its licensees pay to use its architecture, the basic design of their chips. If Arm sells complete compute subsystems, ready-made blocks of a chip, it earns more per design and helps customers move faster. But it also comes closer to its licensees' products.

With opportunity cost in mind, you rank the options instead of only counting revenue. Engineers who build subsystems are not improving the architecture that licensees pay for. Staying neutral, out of its licensees' markets, is a resource too, and Arm can spend it

only once. Say which advantage is worth that cost, and which customer reaction would make you swap the ranking. Where to stop is still your call.

In your answer

- "The scarce resource here is ___. Its strongest other use is ___."
- "I give up ___ to get ___. The exchange is worth it because ___."
- "Keeping this capacity in reserve has a cost too: ___."
- "I would swap this ranking if ___. I would see it in ___."

Common misuse

The usual mistake is to list every drawback of your choice and call it opportunity cost. A list of risks is not an option you gave up. To tell them apart, ask: can you name a specific other use of the same resource, and would you fund it instead? If you cannot, there is no opportunity cost on the page, only worries.

The second mistake is false precision: a confident number for an option you never tried. If you cannot honestly estimate what you gave up, rank the options. Then say what evidence would change the ranking.

References

- Friedrich von Wieser, *Natural Value*, translated by Christian A. Malloch and edited by William Smart (Macmillan, 1893) — the idea in English: archive.org/details/naturalvalue00wiesiala
- David I. Green, "Pain-Cost and Opportunity-Cost", *Quarterly Journal of Economics* 8(2), 1894, pp. 218–229 — the paper that named it: econpapers.repec.org record

- Robert A. Burgelman, "Fading Memories: A Process Theory of Strategic Business Exit in Dynamic Environments", *Administrative Science Quarterly* 39(1), 1994, pp. 24–56 — the Intel study behind the example: gsb.stanford.edu record
- CORE Econ, *The Economy*, Unit 3: scarcity, work and choice — the idea explained through examples: books.core-econ.org/the-economy/v1/book/text/03.html

Second-order thinking

A Alternatives · M Manage uncertainty · 5 min

When you need it

The first effect is obvious; the reaction to it is not

Keep asking: and then what?

How to use it

1. Write the direct effect of the decision in one sentence.
2. List who will feel that effect, and say what is now in each one's interest to do.
3. Follow the strongest response one step further, and check whether your recommendation still holds.

Try it in your answer

Pick one group your decision touches. Follow their response two steps further, and name the result you would watch.

Practise it on a case

Tesla: merge with SpaceX? question 1 · **Substack** question 1 · **Costco** question 2

What it is

A shop cuts its prices, and sales go up. Everyone sees that first effect. Later, customers learn to wait for the next discount, and the profit on each sale shrinks. These reactions come after the decision is made, and they can undo the gain.

Second-order thinking means following a decision past its first effect. You ask what it does at once, how the people it touches will respond, and what their response does to your goal.

Where it comes from

The idea has no single author. The French economist Frédéric Bastiat put it clearly in 1850. His pamphlet was called *Ce qu'on voit et ce qu'on ne voit pas* ("What is seen and what is not seen"). He wrote: "an act, a habit, an institution, a law produces not only one effect, but a series of effects. Of these effects, the first alone is immediate; it appears simultaneously with its cause; *it is seen*. The other effects emerge only subsequently; *they are not seen*."

In 1936 the sociologist Robert Merton named the problem in the *American Sociological Review*: "The Unanticipated Consequences of Purposive Social Action". He meant what people did not foresee when they acted on purpose. He blamed ignorance, mistakes and caring only about the result you want now.

In 1971 Jay Forrester, who studied complex systems, explained why it happens, in "Counterintuitive Behavior of Social Systems". Systems made of people run on many feedback loops, where an effect comes back round and changes its own cause. We are used to short chains of cause and effect, so we misread such systems.

Business readers mostly know it as "second-level thinking", from chapter 1 of *The Most Important Thing* (2011) by the investor Howard Marks. In *Principles* (2017) another investor, Ray Dalio, calls it weighing second- and third-order consequences.

What it corrects

Say you plan to cut a price or change how you sell. You can measure the direct effect now, but not the reactions. A good analyst builds the argument on what can be measured, because that is where the data is. And the analyst usually gets that first step right. Working harder does not help: the numbers get more exact, but they still describe only that step. The reactions drop out of the forecast for two reasons. The people who react are outsiders, and they act in their own interest, not yours. And they react late, after the decision has been reviewed and called a success.

How it works

1. Write the direct effect in one sentence. Name the number you expect it to move.
2. List who will feel that effect: customers, competitors, distributors, staff, regulators.
3. For each of them, say what is now in their interest to do.
4. Take the strongest response and follow it one step further: what does it do to your number?
5. Say how long the reaction takes to reach your number: weeks, quarters or years.
6. Check that your recommendation holds once the reaction hits your number. Name one early signal of the reaction from outside your own systems.

Worked example

On 20 November 1989 Mexico City introduced Hoy No Circula ("No driving today"). Each car had to stay off the road one weekday a week. The last digit of its number plate set the day. The rules covered

2.3 million vehicles, so about 460,000 stayed off on any weekday. Enforcement was strict, and almost everyone obeyed. The second-order reaction was just as simple as the rule. A family with two cars whose plates end in different digits could drive every day.

In 2008 the economist Lucas Davis studied the result in the *Journal of Political Economy*. He used readings from air-monitoring stations. He checked every pollutant in several ways, and found no sign that the air got cleaner. The number of registered vehicles rose by about 325,000 when the programme started (probably between 51,000 and 597,000). New car sales rose about 15%, but that rise came to under 2% of all the cars on the road. So nearly all the extra cars were used ones, from elsewhere in Mexico and the United States. Older cars pollute more. Metro and bus use did not rise, and petrol sales did not fall. The ban kept each car off the road on its day, but the extra cars cancelled the gain.

In a Business Case Weekly case

In the Nike case, the decision in June 2020 is about a plan called Consumer Direct Acceleration. Nike would drop many wholesale partners, the shops that resell its goods, and sell straight to customers. The first-order argument on management's slide is convincing: more profit per pair and fast online growth in the pandemic.

Think one step further, and you will ask two questions. What will a retailer do with the shelf space Nike gives back? What will a shopper do when the shoe is missing from the store they walk into? Both reactions happen in someone else's shop, and management's own dashboard cannot show them. So the case asks you for tripwires:

warning signs chosen in advance and measured outside Nike's own stores and apps. Even then, you will not know yet whether the plan was right.

In your answer

- "The direct effect is X, and I expect it to last about N months."
- "The one with most to gain from reacting is Y. The cheapest thing Y can do is Z."
- "If Y does Z, the gain falls from A to B. That changes my recommendation like this: ..."
- "The signal I would watch is outside our own channel. It is ..."

Common misuse

The usual mistake is a chain of five guesses. Each sounds likely, but none has a number, and the chain ends in a story nobody will check. Nobody can prove such a chain wrong. So it never overturns a recommendation, and it never supports one. Test each step with one question. Can you name who acts, what they gain, and what you would see when they start? A step without all three proves nothing. Stop at the last step that has all three, and put a number on it.

References

- Frédéric Bastiat, "What Is Seen and What Is Not Seen" (1850), in *Selected Essays on Political Economy*, trans. Seymour Cain (Foundation for Economic Education, 1995) — full text at Econlib. Short enough to finish in an evening, and still the best statement of the idea.
- Jay W. Forrester, "Counterintuitive Behavior of Social Systems", *Technology Review* 73, no. 3 (January 1971): 52–68.

- Lucas W. Davis, "The Effect of Driving Restrictions on Air Quality in Mexico City", *Journal of Political Economy* 116, no. 1 (2008): 38–81.
- Howard Marks, *The Most Important Thing: Uncommon Sense for the Thoughtful Investor* (Columbia University Press, 2011), chapter 1, "Second-Level Thinking".

Outside view

M Manage uncertainty · R Relevant evidence · 5 min

When you need it

The plan rests on a forecast or on “we are different”

Start from what usually happens in similar cases, then adjust for your own.

How to use it

1. Choose a reference class: past cases like yours in what matters.
2. Find the class's base rate and range before you read this case's own forecast.
3. Adjust only for facts that really set this case apart.

Try it in your answer

Name the reference class this decision belongs to, and one fact from the case that justifies moving away from its base rate.

Practise it on a case

Netflix question 1 · **Constellation Software** question 1 · **Nintendo** question 2

What it is

A friend is sure her new restaurant will profit in its first year. Before you judge her plan, ask how most new restaurants do. That question is the outside view.

You forecast from a reference class: past attempts whose outcome depended on the same things as yours. Your estimate starts from the range of its results. What you know about your own plan comes later, as changes you must argue for.

Where it comes from

The idea comes from the psychologists Daniel Kahneman and Amos Tversky. In a 1977 paper they argued that forecasters often ignore the record of similar past cases. They called this record distributional information. Ignoring it may be the biggest cause of forecasting error, they said. So analysts should set up each forecast to use that record.

The names "inside view" and "outside view" came in a 1993 paper by Kahneman and the strategy researcher Dan Lovallo. With the inside view, you treat each case as one of a kind and work from its details. With the outside view, you work from statistics on cases alike in what matters. In July 2003 the two explained the idea to managers in *Harvard Business Review*. They wrote about the planning fallacy, the habit of underestimating time and cost.

Bent Flyvbjerg, a Danish expert on large projects, turned the idea into a procedure, reference class forecasting. He built it with the consultancy COWI for the UK Department for Transport. It became official guidance in August 2004.

What it corrects

You know your own project better than anyone, so you forecast from the plan: its steps, its team, the obstacles you can see. These forecasts go wrong the same way, again and again. Cost and time come out too

low and benefits too high, because a plan assumes nothing unexpected will happen.

Working harder makes this worse. More effort adds detail to the same plan. The missing risks are not in it. They are what happened to others: similar teams that gave up, technology that came late, demand that never appeared. Kahneman and Tversky also found that even people who fully understand these errors keep making them.

Organisations add a second problem. When projects compete for money, their backers have a reason to be hopeful. They can easily argue with a scenario. A record of how similar projects turned out, the base rate, is harder to argue with.

How it works

1. Name what you are forecasting and when you will know the answer.
2. Choose the reference class: past attempts like yours in what matters, wide enough to have data and narrow enough to compare.
3. Collect the results of that class, including abandoned attempts.
4. Take the usual result as your first estimate and the spread as your range.
5. Adjust only where evidence shows your case differs from the class in something its data measures. Write down the size of the change.
6. Write down what would send you back to the base rate.

Worked example

Edinburgh was planning tram Line 2. In October 2004 the engineering firm Ove Arup and Partners Scotland reviewed the line's business case for the Scottish Parliament's Edinburgh Tram Bill

Committee. Transport Initiatives Edinburgh, the project company, had written it. This inside view was careful work. The base cost, £255 million, came from a cost database. It was compared with other UK light rail schemes and checked against earlier estimates. On top came £64 million, about 25 per cent, for unexpected costs and optimism bias. The total was about £320 million.

Arup then applied the Department for Transport's uplifts: extra percentages based on overspending on finished rail projects. For an 80 per cent chance of staying within budget, the uplift was 57 per cent: £255 million times 1.57 gives £400 million. For a 50 per cent chance it was 40 per cent, or £357 million. Arup said even these were probably too low. The uplifts are meant for the decision to build. Line 2 was not even at the outline business case stage. The earlier the stage, the more is unknown, and the larger the uplift should be.

For one project on one day, the careful estimate and the base rate differed by £80 million at the 80 per cent level. Line 2 was never built. Trams did run on a shortened Line 1, from the airport to York Place. It opened in May 2014 and reportedly cost £776 million, £231 million over the budget for the full line to Newhaven. The 2023 public inquiry put the real cost near £836 million. The rail uplifts were based on overspending of this size.

In a Business Case Weekly case

In the SoftBank case, take the 2022 fork. As one of the Vision Fund's limited partners, its outside investors, you are asked to back a newer, bigger idea. The inside view is that idea and the fund's account of what changed after WeWork.

From the outside, first ask which group this bet belongs to. Late-stage tech funds of the same years? Funds of this size? Funds where one person decides? Then ask what those funds earned, counting those quietly closed. Last, ask which fact about this fund justifies a different number. You get no answer this way, only the figure your memo must argue against.

In your answer

- "The reference class here is ..., and the usual result across it is ..."
- "Before reading this company's own forecast, I put the base rate for ... at ..."
- "This case differs from the class in one measurable way, ..., which moves my estimate from ... to ..."
- "If that difference does not hold, I go back to the class figure of ..."

Common misuse

The fake version is a reference class chosen after the answer. Say you pick cases by industry label, or because they flatter the plan. Then any number you wanted comes back, looking like a base rate. To avoid this, name the class and its range of results before reading the case's own forecast. And be able to say which members of the class failed. If none failed, you probably picked only the success stories.

The opposite mistake is to treat the base rate as a verdict. It is only where the estimate starts. Strong evidence about this case can move it far. But you must argue for the move and say how big it is.

References

- Daniel Kahneman and Dan Lovallo, "Timid Choices and Bold Forecasts: A Cognitive Perspective on Risk Taking," *Management Science* 39(1), 1993, 17–31. PDF
- Bent Flyvbjerg, "From Nobel Prize to Project Management: Getting Risks Right", *Project Management Journal* 37(3), 2006, 5–15. The Edinburgh figures and the whole procedure.
- HM Treasury, Green Book supplementary guidance: optimism bias. The Treasury's general uplift tables for UK public projects, for use when there is no better data. The Department for Transport built its rail uplifts, used at Edinburgh, in response to the Green Book.
- Daniel Kahneman and Amos Tversky, "Intuitive prediction: Biases and corrective procedures," 1977; reprinted in *Judgment Under Uncertainty: Heuristics and Biases*, Cambridge University Press, 1982.

Pre-mortem

M Manage uncertainty · E Execute and learn · 6 min

When you need it

The plan feels safe and you cannot say how it would fail

Imagine the plan has already failed, and find out why while you can still change it.

How to use it

1. Announce, as a fact, that by a stated date the plan has failed.
2. Have everyone write causes alone, in silence, before anyone argues about them.
3. Turn each top cause into a change: prevent it, fund a backup plan, or set a tripwire.

Try it in your answer

Name one way the plan could fail, the assumption it comes down to, and the first sign that would show it.

Practise it on a case

Mattel: look for a buyer? question 1 · **Hilton** question 1 · **Xiaomi** question 1

What it is

Your team is about to approve a plan. The leader says: "It is a year from now, and the plan has failed. Write down why." Everyone writes alone, in silence, before anybody speaks. That is a pre-mortem. Teams usually look for the cause of a failure afterwards, in a post-mortem. In a pre-mortem they look before, while the plan can still change.

The exercise is short. You end with a list of specific causes of failure, the most likely first. You also agree on at least one change: to the plan, or to what the team will watch.

Where it comes from

The method comes from Gary Klein, a psychologist who studies how experts decide when time is short. He described it in Harvard Business Review in September 2007, in "Performing a Project Premortem". In January 2021 he wrote in Psychology Today that he had developed it three decades earlier, to cut problems in his company's projects.

The idea behind it is older. Researchers call it prospective hindsight: imagining a future event as if it has already happened. Klein's article traces it to a 1989 paper in the Journal of Behavioral Decision Making. Its authors were Deborah J. Mitchell of the Wharton School, J. Edward Russo of Cornell and Nancy Pennington of the University of Colorado. Klein says the trick made people 30% better at finding reasons for future outcomes. The paper's authors claimed less. They found that people who were sure of an outcome explained it differently from people who were not. Moving the event into the past mattered less.

Later Beth Veinott, Klein and Sterling Wiggins tested the exercise itself. They asked 178 people to review a plan for responding to a pandemic. Some ran a pre-mortem. Others simply criticised the plan, listed pros and cons, or listed only cons. The pre-mortem group lost the most confidence in the plan. Klein sees this as overconfidence falling, and that is the point of the exercise.

What it corrects

A plan has just been presented. The manager behind it clearly believes in it, and the meeting is nearly over. Someone in the room has a concrete doubt. Perhaps the plan depends on something they know is fragile, or a deadline clashes with another promise. They say nothing.

They stay quiet for two reasons. First, the leader asked "Any concerns?" To answer, people have to judge the plan, not picture how it fails. So they say "risky" instead of naming a cause. Second, speaking up costs something. The doubt sounds like criticism of a colleague's work. And if the project goes well, they will look foolish in public.

Working harder fixes neither problem. A careful person studies the plan more closely. But that plan is built on the very assumptions they ought to question. Klein changes the question instead. Once the failure is announced as a fact, nobody has to judge the plan. People only have to explain what happened, and experts do that quickly and in detail. Naming a cause is no longer an attack on a colleague. It is the job everyone in the room was given.

How it works

1. Present the whole plan, so that everyone argues about the same plan.
2. Announce, as a fact, that by a stated date the plan has failed. For example: the launch slipped nine months, and the product was withdrawn.
3. Give everyone two or three minutes to write causes alone, in silence, including the ones they would normally keep to themselves.

4. Go round the table, one cause per person, the leader first. Write every cause down before anyone argues about one.
5. Merge the duplicates. Then rank the causes by how likely they are, not by how dramatic they sound.
6. Turn each top cause into a change. Prevent it, fund a backup plan, or set a tripwire: a named warning sign that one person watches.

Worked example

Klein's 2007 article describes a project to give advanced planning algorithms to military officers who plan air campaigns. One team member had sat silent through a long kick-off meeting. In the pre-mortem he named a cause of failure: one of the algorithms was too much for the laptops used in the field. The software would run for hours, but the planners needed an answer in minutes. So the project would be impractical.

Then the developers admitted that they had already built a shortcut, a simpler version of the algorithm, but had not wanted to mention it. The team swapped it in, and the project succeeded. Nobody discovered anything new in that room. The team already knew about the laptops and the shortcut, and in the exercise people said so. Klein says the exercise takes twenty to thirty minutes.

In a Business Case Weekly case

The Rolls-Royce case ends with a decision about money in mid-2020. Revenue is down by about a third, the loss is heading for £4 billion, and 9,000 job cuts have been announced. The company must choose how to raise money. It can sell new shares to its shareholders at a deep discount (a rights issue). It can borrow more, ask the state for support, or sell parts of the business.

Do not aim the pre-mortem at the pandemic, which has already happened. Aim it at the mix of options you choose. Pick a date eighteen months ahead. Say that by then the company had to ask investors for money a second time, on worse terms. Then name the part of your own mix that caused it. Write down causes that only this company could have. How does it get paid? What will its promises to airlines cost in cash? And in each option, who is left holding the risk?

In your answer

- "Say it is [date] and this decision has failed. The most likely cause is ..."
- "That failure comes down to one assumption: ..."
- "The first sign of it would be ..., and I would expect to see it by ..."
- "So I would change the plan now by ..., and keep ... in reserve."

Common misuse

The usual mistake is the disaster list: a page of general catastrophes that gets written and filed away. People call it a pre-mortem, but it does not help. A cause that could hit any project tells you nothing about this one. You can spot a disaster list with two questions. Could the cause you wrote appear unchanged in a competitor's plan? If so, it is not specific enough yet. Did any line on the list change the plan, the budget or what you watch? If not, the meeting made people anxious and decided nothing.

References

- Gary Klein, "Performing a Project Premortem", Harvard Business Review, September 2007 (reprint F0709A): https://hbr.org/2007/09/performing-a-project-premortem

- Deborah J. Mitchell, J. Edward Russo and Nancy Pennington, "Back to the future: temporal perspective in the explanation of events", Journal of Behavioral Decision Making, volume 2, 1989, pages 25 to 38: https://onlinelibrary.wiley.com/doi/abs/10.1002/bdm.3960020103 — the primary source for prospective hindsight, and short enough for one evening.
- Gary Klein, "The Pre-Mortem Method", Psychology Today, 14 January 2021: https://www.psychologytoday.com/us/blog/seeing-what-others-dont/202101/the-pre-mortem-method — Klein's own account of the procedure and where it has been used.
- Beth Veinott, Gary Klein and Sterling Wiggins, "Evaluating the Effectiveness of the PreMortem Technique on Plan Confidence", Proceedings of the 7th International ISCRAM Conference, Seattle, 2010.

Fragility audit

M Manage uncertainty · 5 min

When you need it

Everything hangs on one customer, supplier or person

Find the one dependency that could ruin your plan, without trying to predict what will go wrong.

How to use it

1. List what must keep working, and say how each thing could fail.
2. Rate the damage of each failure first, then its likelihood.
3. Start with the worst: remove the single point of failure, or buy a back-up before you need it.

Try it in your answer

Name the most dangerous dependency in your answer, and one step that limits the damage even if your forecast is wrong.

Practise it on a case

Nintendo question 1 · **Huawei** question 1 · **LEGO** question 1

What it is

Every plan depends on things that must keep working: a supplier, a customer, a platform, a key person, a price. In a fragility audit you list them and ask what the plan loses if each one fails. You look for the one whose loss would hurt far more than its size suggests.

Forecasting is guessing which future will come. Here you look for the one thing that could ruin the plan in many different futures. You predict nothing. You decide what to risk.

Where it comes from

The name has no single author. The method is built from two older ideas.

The idea of fragility comes from Nassim Nicholas Taleb, a writer on risk. He set it out in *Antifragile* (2012), the fourth book in his series *Incerto*. For Taleb, fragile means that past a certain point, harm grows faster than the shock that causes it. Without predicting the shock, he says, you can see where a plan would be hurt out of proportion. In a 2012 paper, Taleb and the mathematician Raphael Douady put this into formulas. They offer a "fast-and-frugal", "model-free, probability free" test. You measure how harm grows as the shock grows, without forecasting whether it will come.

The steps are older. Engineers call them Failure Mode and Effects Analysis (FMEA). You list the parts of a machine, describe how each can fail and rate the result. The US military wrote this into procedure MIL-P-1629 on 9 November 1949. Later, NASA contractors used it on Apollo and Voyager. In a fragility audit you do the same to a business plan. Like Taleb, you rate the damage before the probability.

What it corrects

Say one customer brings in 40% of your revenue. A careful analyst forecasts whether the customer will leave. They gather evidence, estimate the odds and get a number. The forecast may be good. But one wrong forecast could still ruin the company. Either way, 40% of revenue stays at risk.

There is a second problem. Some dependencies hide in the stable parts of a plan, where nobody writes them down. Think of the firm that handles your card payments, or the visa rules. Think of the one engineer who understands the billing software, or the shelf space a retailer keeps for you. None of them is in the figures, because none has ever changed.

How it works

1. Write the recommendation as a commitment with a date and a size, so you can see what could break.
2. List what must keep working for it: suppliers, customers, sales channels, platforms, regulators, key people, a price or rate that must hold.
3. For each, say in one sentence how it fails: what stops, and how soon.
4. Rate the damage first, then the likelihood, and sort by damage. A rare but ruinous failure ranks above a likely but cheap one.
5. For the top two, ask: could the plan recover, or would this be fatal? Fatal means the strategy is gone by next year.
6. Act on the fatal ones now: remove the single point of failure, limit your losses, or buy a back-up while nobody else wants it.
7. Name the first visible sign of each failure, so you have something to watch.

Worked example

On 17 March 2000, lightning struck a power line in Albuquerque, New Mexico, and started a fire at a Philips chip factory. The fire burned for ten minutes. The factory made radio frequency chips for Nokia and Ericsson phones. Philips needed weeks to get back to full output.

Nokia bought some of these chips from more than one supplier. Its staff noticed the drop in deliveries before Philips reported a problem. Within two weeks Nokia sent about thirty managers across Europe, Asia and the United States. It redesigned some of the Albuquerque chips so other factories could make them, and pushed other suppliers for more. Nokia's production stayed on target.

Ericsson bought the chips from that one supplier alone. Jan Ahrenbring, its marketing director for consumer goods, later told the *Wall Street Journal*: "We did not have a Plan B." Ericsson estimated its lost sales at $400 million or more. By January 2001 Nokia held about 30% of the phone market, up from about 27% a year earlier. Ericsson's share had fallen from about 12% to 9%. On 26 January Ericsson announced it would hand all its phone manufacturing to Flextronics.

Neither company predicted the lightning. But Nokia had checked this dependency. It had paid, in design work and ties with suppliers, for the ability to survive the failure.

In a Business Case Weekly case

In the Nike case, the plan is Consumer Direct Acceleration. From mid-2020 Nike cut back on the shops that resold its shoes and put its own stores and apps first. Direct sales did earn more per pair. In a fragility audit you ask instead: what dependency does the plan create? Casual buyers would meet the brand only in Nike's own stores and apps. The shelf space Nike gives up stays with the shops, which can give it to rivals: On, Hoka or New Balance. The deciding question: if someone else fills that space, can Nike buy it back, at what price, and for how long?

In your answer

- "This recommendation depends on X. If X fails, ... stops."
- "The damage would be fatal, not just expensive, because ..."
- "I am not predicting this failure. I am limiting its cost by ..."
- "The first sign would be ..., so I would check it every month."

Common misuse

The usual mistake is a risk register: twenty rows, each with a likelihood score, an impact score, a colour and a person in charge. The plan goes ahead unchanged. The table looks like an audit, but giving a risk a score does not make it smaller. If you changed no commitment and bought no back-up, you only described the risks.

The opposite mistake is buying back-ups for everything. Second suppliers, spare capacity and standby contracts cost money and add complexity. A dependency whose failure would only be annoying deserves none of them. Look at the damage first, and pay only where a failure would be fatal.

References

- Nassim Nicholas Taleb, *Antifragile: Things That Gain from Disorder*, Random House, 2012. The practical method is in the chapters on the barbell and via negativa.
- Taleb & Douady, "Mathematical Definition, Mapping, and Detection of (Anti)Fragility" (2012). The formal version, with the detection test.
- Almar Latour, "Trial by Fire: A Blaze in Albuquerque Sets Off Major Crisis for Cell-Phone Giants", *Wall Street Journal*, 29 January 2001. The original report behind the worked example; a twenty-minute read.

- Sunil Chopra & ManMohan S. Sodhi, "Managing Risk to Avoid Supply-Chain Breakdown", *MIT Sloan Management Review*, Fall 2004. On rating dependencies without buying back-ups for everything.

Falsifiability

M Manage uncertainty · R Relevant evidence · 5 min

When you need it

You cannot name the news that would make you drop the decision

Name the result that would prove your recommendation wrong, so that it can be tested.

How to use it

1. Write your main claim in one sentence, with its reason.
2. Name a result that would prove it wrong early enough for you to change the plan.
3. Decide now what you will do if that result appears.

Try it in your answer

Add one sentence that begins “I would be wrong if ...” and end it with a result you can measure and a date.

Practise it on a case

Klarna question 1 · **SoftBank** question 1 · **Palantir** question 1

What it is

A friend says, "It will rain tomorrow." Tomorrow you look outside and see whether they were wrong. Another friend says, "Tomorrow the weather will be whatever the season brings." No day can prove that wrong. The first claim is falsifiable: it can be proved wrong, because it rules out a dry day.

In a decision, you write down when you would drop your recommendation. You name four things: the result that would prove it wrong, the threshold, the date and your next step. This way you check that the claim can be tested. Whether it is true, you find out later.

Where it comes from

Karl Popper (1902–1994), a philosopher of science, set out the idea in *Logik der Forschung*. Julius Springer published the book in Vienna in 1934, though the title page says 1935. The English version, *The Logic of Scientific Discovery*, came out in 1959.

Science rests on observation, but some beliefs only look like science. Popper wanted to tell them apart. He found the answer in what a theory forbids: the results that should never appear. In 1953 Popper gave a lecture at Cambridge. It later became the first chapter of his book *Conjectures and Refutations* (1963). There he wrote: "A theory which is not refutable by any conceivable event is non-scientific. Irrefutability is not a virtue of a theory (as people often think) but a vice."

People began to use the idea in business much later, and no single person started it. Popper wrote about scientific theories, not strategy memos. He also insisted on a limit. Whether a claim could be proved wrong is a matter of logic. Deciding that a real result has proved it wrong always takes judgement.

What it corrects

A team commits to a strategy it believes will pay off: controlling the customer relationship, entering a market early, or hiring the best people. The team reads a good quarter as proof. A flat quarter, it says,

means the change needs more time. A bad one means the company invested too little. Every result supports the belief, so no result can test it.

This is not the same as being too confident. The fault lies in how the claim is built. Careful people make it worse. When in doubt, they gather more evidence, and a stretchy claim fits any evidence. The analysis improves, but the belief is never tested. The cost comes later, when the decision can no longer be undone and people still argue about what the results mean.

How it works

1. Write the main claim in one sentence, with its reason: if we do X, Y changes because of Z.
2. Say which results the claim rules out. A claim that rules out nothing cannot be tested.
3. Name the most believable other explanation for the same result. Pick an observation that tells the two apart.
4. Add a number, how you will measure it, and by when.
5. Choose a source your own plan cannot inflate. A signal your team controls is not a test.
6. Write down now what you will do if the threshold is crossed: stop, change the plan, or test more deeply.

Worked example

In 1915 Einstein published his general theory of relativity. He predicted that light from a star passing the edge of the Sun would bend by 1.75 arcseconds, a tiny angle. Newton's physics gave 0.87,

about half as much. So Einstein's theory ruled out Newton's number. The test needed a solar eclipse, when stars near the Sun can be seen, so its date was known in advance.

On 29 May 1919 two British expeditions photographed the eclipse, one at Sobral in northern Brazil, the other on the island of Príncipe. On 6 November 1919 the astronomers Dyson, Eddington and Davidson reported the results. The four-inch telescope at Sobral gave 1.98 ± 0.12 arcseconds, and the photos from Príncipe gave 1.61 ± 0.31. Both were close to Einstein and far from Newton.

There is also a warning here about the method's limits. Sobral also had a larger, 13-inch telescope. Measured one way, its photos gave 0.93 arcseconds, close to Newton. Measured another way, they gave 1.52. They were blurred, perhaps because the Sun had heated the mirror. The astronomers set them aside as unreliable, and historians still argue about that choice. A result can prove a claim wrong only as far as you can trust the measurement.

In a Business Case Weekly case

In the Nike case, you need this method at the June 2020 board meeting on a plan called Consumer Direct Acceleration. The plan rests on one claim: selling direct gives Nike control of demand and of its profit per sale. Almost any result fits that claim. A strong quarter confirms it. A weak one means the switch needs more time.

You are asked to name two tripwires, warning signs measured outside Nike's own channels. Each becomes a test, not just a report, when it has a threshold, a quarter in which you will read it, and a consequence. Agree that consequence before Nike cuts its wholesale accounts, the shops that resell its products. Whether to approve the plan is yours to argue.

In your answer

- "My recommendation rests on the claim that ..., and I would be wrong if ..."
- "I would watch ... outside our own channels. If it falls below ... by ..., I treat the claim as broken."
- "Another explanation for the same numbers is ... The observation that tells them apart is ..."
- "If that result appears, I stop and renegotiate rather than push back the deadline."

Common misuse

The trap here is a fake test: one that cannot go off in time, or at all. It usually appears as a milestone. It might be a measure the team can redefine. It might be a review set for after the factory is built. Or it might be a target so low that any result clears it.

One question tells real from fake. Which number could appear next quarter and force you to change the plan, and who reports it? If nobody can name it, or only your own team can measure it, the memo has a test in name only.

The opposite mistake is treating one bad quarter as proof. After a bad result you know only that something is wrong: the claim, or one of the assumptions behind it. So say which assumption you would give up first.

References

- Karl Popper, "Science: Conjectures and Refutations" (lecture, Cambridge, 1953; chapter 1 of *Conjectures and Refutations*, 1963). Twenty pages that hold the whole argument:

https://www.stephenhicks.org/wp-content/uploads/2018/09/PopperK-Science-as-Falsification.pdf

- Stephen Thornton, "Karl Popper", Stanford Encyclopedia of Philosophy: https://plato.stanford.edu/entries/popper/
- Malcolm Longair, "Bending space–time: a commentary on Dyson, Eddington and Davidson (1920)", *Philosophical Transactions of the Royal Society A* 373:20140287, 2015: https://pmc.ncbi.nlm.nih.gov/articles/PMC4360090/
- Karl Popper, *The Logic of Scientific Discovery*, Hutchinson, 1959; first published as *Logik der Forschung*, Julius Springer, 1934.

Two-way / one-way door

E Execute and learn · M Manage uncertainty · 5 min

When you need it

You are agonising over a small bet or rushing a big one

The harder a choice is to undo, the more evidence you gather and the slower you move.

How to use it

1. Split the decision into the part you can undo and the part you cannot.
2. Make the one-way part of the first step as small as you can.
3. Move fast on the part you can undo, and take your time over the part you cannot.

Try it in your answer

Say whether your recommendation is a one-way or a two-way door. Then change the first step so that learning what you need costs less.

Practise it on a case

Evonik: open the books to BASF? question 1 · **Novo Nordisk** question 1 · **Bloomberg** question 1

What it is

Some doors open both ways. Others lock behind you. A two-way door is a decision you can undo once you have learned something: you walk back out. A one-way door is one you cannot undo, because you

have spent money, signed contracts, used up trust or changed what others do next. The harder a choice is to undo, the more evidence and discussion it needs, whatever its size.

Where it comes from

The doors come from Jeff Bezos, Amazon's founder, in his letter to shareholders for 2015, published in April 2016. He wrote that "some decisions are consequential and irreversible or nearly irreversible – one-way doors – and these decisions must be made methodically, carefully, slowly, with great deliberation and consultation." Most decisions, he went on, "are changeable, reversible – they're two-way doors" and "can and should be made quickly by high judgment individuals or small groups." He called the first kind Type 1 and the second Type 2.

The economics behind it is older and has no single author. In 1974 the economists Kenneth Arrow and Anthony Fisher argued that when harm cannot be undone, uncertainty alone is a reason to wait for information. They called the value of keeping a choice open "quasi-option value". In 1994 Avinash Dixit and Robert Pindyck built this into the standard theory of irreversible investment. The economists gave the reason. Bezos gave the idea a name a team can use at a Monday meeting.

What it corrects

A large company settles on one way of making every decision. The way that survives protects people from blame: slow, many people consulted, lots of evidence. Then the company uses it even for choices it could test in two weeks and undo in a day. No report shows the cost: the experiments nobody ran. More care cannot fix this, because too much care caused it. In a footnote, Bezos says companies

that by habit make Type 1 decisions the light, fast way "go extinct before they get large." So the large firms that survive tend to be over-careful.

Young, fast companies make the opposite mistake. They are proud of moving fast, so they also rush through doors that lock behind them. Bezos says this habit is what stops them growing.

How it works

1. Write the decision as the first concrete step, not as the strategy behind it.
2. Split that step into the part you can undo and the part you cannot.
3. Work out what undoing each part would cost: weeks, money, and whose agreement you need. If you cannot describe a way back, that part is one-way.
4. Make the one-way part as small as you can while still learning what you need.
5. Then decide each part in its own way. For the part you can undo, use a small group and move fast. For the one-way part, take time and hear out people who disagree.
6. Say what evidence would later justify the one-way part, and the date you will check for it.

Worked example

In 2011 Netflix made two decisions usually remembered as one. In July it split DVD-by-mail and streaming into two plans at $7.99 each. Its shareholder letter of 25 July 2011 said what this meant. For members who wanted both, the change was "as much as a 60% increase for them from $9.99 to $15.98." In September Netflix

announced Qwikster: the DVD service under a new name, with its own website and billing. On 10 October, about three weeks later, Netflix dropped Qwikster and kept one site.

The lesson is in telling the two apart. The new name was a two-way door. A name, a website and a DVD queue could all be undone, and Netflix undid them. The price rise was a one-way door. The letter of 24 October 2011 treated them very differently. The rebrand "caused a temporary cancellation surge" whose impact was "relatively minor." The price change had "greatly upset many domestic Netflix members" and had "hurt our hard-earned reputation, and stalled our domestic growth." US subscribers fell from 24.59 million to 23.79 million. It was the first fall in the two years of quarterly figures in the letter. Netflix told investors it was "done with pricing changes." It undid the cheap choice within three weeks. The expensive one could not be undone.

In a Business Case Weekly case

The Sega case stops in January 2001. Sega's Dreamcast console is selling well after a price cut, and the PlayStation 2 is in the shops. The decision is how long to keep making the Dreamcast. You cannot settle that by asking which door it is, but you can split it in two. One more quarter of production is a cost Sega can stop the next quarter. Announcing the end of Sega consoles is different. A platform is made of game developers, shop space and places in the release schedule. They all go elsewhere. Sega cannot get them back by changing its mind later. Price each part separately before you argue for either.

In your answer

- "We could undo ... in ... weeks, for about ..."

- "The part that locks behind us is ..., because it changes ... for people outside the company."
- "So I would move fast on ... and wait on ... until ... is true."
- "If I am wrong, the way back is ... It needs ... to agree, and it costs ..."

Common misuse

The phrase becomes an excuse to rush. People say "it's a two-way door" about a step that can be undone on paper but not in real life. Think of a trial run that a customer will plan their product around. Or a price test after which buyers remember what your product is worth. Do not assume a step can be undone. Ask three questions out loud. Who has to agree to undo it? How long will that take? What will it cost? If you cannot describe the way back, it is a one-way door, whatever people call it.

References

- Jeff Bezos, "2015 Letter to Shareholders", Amazon.com, April 2016. The doors passage and the survivor footnote are in the section on one-size-fits-all decisions.
- Netflix, "Letter to Shareholders", 25 July 2011 (Exhibit 99.1 to Form 8-K). The price decision in Netflix's words, before customers reacted.
- Netflix, "Letter to Shareholders", 24 October 2011 (Exhibit 99.1 to Form 8-K). The next quarter: the subscriber table and two apologies of unequal weight.
- Kenneth J. Arrow and Anthony C. Fisher, "Environmental Preservation, Uncertainty, and Irreversibility", The Quarterly Journal of Economics 88(2), May 1974, pages 312-319. Eight pages: why to wait when damage cannot be undone.

Commander’s intent

E Execute and learn · F Focus · 6 min

When you need it

Others will carry out the plan, and it will change on the first day

Write down the purpose, what must stay true and the end state, so people can act when the plan breaks.

How to use it

1. State the purpose: why the work is done.
2. Name the few things that must stay true, whatever else changes.
3. Describe the end state, not every step to reach it.

Try it in your answer

Add one sentence: your purpose, what must stay true and the end state.

Practise it on a case

OnlyFans question 1 · **Porsche** question 1 · **Rolls-Royce** question 4

What it is

Imagine others carry out your plan and cannot reach you. When things change, what do they need from you to decide alone?

Commander’s intent is a short note you write before people set off. In it you name the purpose: why the work is being done. You name what must stay true, whatever else changes. And you describe the end state: what the finished result looks like. The details of the plan can change. The purpose cannot.

Where it comes from

Helmuth von Moltke the Elder ran the Prussian General Staff from 1857. He set out the idea in his 1869 *Instructions for Large Unit Commanders*. He argued that detailed orders are usually useful only for the present situation. Every move of the troops changes the conditions the orders were written for. So an order should open with the aim. Lower-ranking commanders add the details they can see. They act on their own judgement, without waiting for orders, as long as they serve that aim. In Germany the method later got the name *Auftragstaktik*, roughly "mission tactics". The name came from officers who opposed it.

The English name comes from the US Army. Its field manual on operations, FM 100-5, first included rules close to *Auftragstaktik* in its 1982 edition. In 1986 the Army fully adopted orders of this kind. In 1993 it asked commanders for the "willingness and ability to act independently within the framework of the commander's intent". In today's manual, ADP 6-0 of 31 July 2019, intent is a short, clear statement of the purpose and the desired end state. With it, lower-ranking commanders can act "without further orders, even when the operation does not unfold as planned". Chip and Dan Heath brought the term into business writing in *Made to Stick* (2007).

What it corrects

Picture many people carrying out a decision at once, while things change faster than they can check with the boss. A capable manager answers with a thorough plan, and that is the trap. The more detailed the plan, the more it rests on facts that will have changed within days.

When the facts change, the team splits two ways. Some keep following instructions that no longer serve any purpose, because those are their orders. Others improvise sensibly, each with their own idea of the goal. Working harder does not help. The effort goes into detail, and detail goes out of date first. In a long plan, nobody can tell which lines matter and which are one way of many.

How it works

1. State the purpose in one sentence: why this is being done.
2. Describe the end state in things people can see or measure, so they can tell when it has arrived.
3. Name the two or three conditions that must hold, whatever else changes.
4. Say by name who may decide what, including who may leave the plan.
5. Give a default rule for when people cannot reach you and the plan no longer fits.
6. Cross out every instruction that people on the spot could decide for themselves.

Worked example

On 9 October 1805 the British admiral Horatio Nelson wrote a secret memorandum aboard his flagship *Victory*, off Cadiz. He planned for forty ships: two lines of sixteen and an advanced squadron of eight. His purpose was an end state, not a manoeuvre. He wanted the enemy fleet "captured or destroyed", not driven off. He handed over control plainly: "The second in command will, after *my* intentions are made known to him, have the entire direction of his line." Something must be left to chance, he wrote. So he gave a default rule:

“in case signals can neither be seen nor perfectly understood, no captain can do very wrong if he places his ship alongside that of an enemy.”

Twelve days later almost none of those figures held. At Trafalgar on 21 October the British had 27 ships of the line against 33. No advanced squadron was formed, and the fleet attacked in two columns. Nelson was shot early and died three hours later, at half past four. Collingwood, his second in command, took over. Of the 33 French and Spanish ships, 18 were captured or destroyed. The British lost no ship of the line. The plan’s numbers were wrong before the first shot. But the purpose, the handover and the default rule held, so no captain waited for an order that could no longer come.

In a Business Case Weekly case

The McDonald’s case is set in the 2024 “value war” between fast-food chains over cheap meal deals. About 95 per cent of McDonald’s restaurants belong to franchisees: local owners who run them under the brand. So headquarters cannot simply order a discount. Owners pay for it out of their own profits, and each judges by their own numbers. Here, writing the intent means answering three questions. What must stay true for the whole chain, whichever deal runs? Which end state would tell you that customers see McDonald’s as good value again? When may an owner leave the company’s plan without breaking the chain’s promise of good value? As a franchisee, you still have to vote on the $5 meal deal and name the number that would change your vote. Writing the intent only makes clear what you are voting on.

In your answer

- “The purpose of this decision is …, and it holds even if the method changes.”
- “Whatever else changes, these two things must stay true: …”
- “We will know we have reached the end state when …”
- “If the plan cannot be followed, the default is …, and … decides on the spot.”

Common misuse

The common mistake is a slogan. “Win on value” sounds like a goal but decides nothing. People will quote it to justify a trade-off, but nobody can tell from it which one to make. To test your intent, give it to two capable people separately. Give each the hardest choice they could face on the spot. If they reach the same answer without calling you, you wrote intent. If they call, or come back with different answers, you wrote a slogan. A subtler mistake is an intent with ten conditions that must never change. That is a detailed plan with a purpose written on top. You meant to let people judge for themselves, but with that list there is nothing left to judge.

References

- Julian S. Corbett, ed., *Fighting Instructions 1530–1816* (Navy Records Society, 1905). Nelson’s memorandum of 9 October 1805 in full: two pages, the quickest way to see the method on paper. gutenberg.org/ebooks/16695
- ADP 6-0, *Mission Command: Command and Control of Army Forces* (Headquarters, Department of the Army, 31 July 2019). The definition is in paragraph 1-45. The official link has moved; a copy is at irp.fas.org/doddir/army/adp6_0.pdf

- Michael J. Gunther, *Auftragstaktik: The Basis for Modern Military Command?* (School of Advanced Military Studies, US Army Command and General Staff College, 2012). From Moltke's 1869 instructions to FM 100-5, in an evening: the DTIC record
- Wikipedia, Battle of Trafalgar, for fleet sizes, losses and the hour of Nelson's death.

Decision journal

E Execute and learn · M Manage uncertainty · 5 min

When you need it

You will not know the result for months

Before you know the result, write down what you chose and why, what you assume and expect, and when you will check.

How to use it

1. Write down the decision, today's date and the alternatives you considered.
2. Before you know the result, write down what you expect, how sure you are and what you are assuming.
3. Set a date to reread the entry. Then compare what happened with what you wrote, and do not edit the entry.

Try it in your answer

End your answer with the next step and the assumption you will watch. Then add a tripwire, a warning sign you name in advance, and the date or event when you will come back to the decision.

Practise it on a case

Novo Nordisk question 6

What it is

You decide something today and will not know for months whether it was right. A decision journal is where you write it down now, with the date. Each entry records what you chose, what you turned down

and why, and what you expect and how sure you are. It also lists the assumptions behind the choice and when you will reread it.

A diary records what has happened. A journal entry is written before the result. It states an expectation that events can prove wrong, and a date to check it. So later you can judge how well you reasoned, not whether you were lucky.

Where it comes from

No single person invented the term. In 1975 the psychologist Baruch Fischhoff described the problem the journal solves, in his paper "Hindsight ≠ Foresight". In his experiments, people who were told how an event ended thought the ending had been easier to predict. They did not notice their view change. Fischhoff wrote that not noticing "can seriously restrict one's ability to judge or learn from the past."

Advice to keep records reached business books in 1989. In *Decision Traps*, J. Edward Russo and Paul J. H. Schoemaker listed ten traps. The ninth was "Not Keeping Track". It means trusting experience to teach you, so you never write down what your past decisions led to.

The name "decision journal" spread later, mostly through Shane Parrish's blog Farnam Street. In his template you note how you see the problem, the alternatives and the expected outcome, with probabilities. You also note how you feel in body and mind.

A popular story has the psychologist Daniel Kahneman telling an investor to buy a cheap notebook and write down what he expected and why. But it comes second-hand, through the investment strategist Michael Mauboussin, and is not in Kahneman's published work. Treat it as a story, not a source.

What it corrects

Take a risky decision whose result comes months or years later. Along the way, events nobody controls shape it. When the result comes, you no longer remember how you reasoned. You remember a version edited to fit the outcome. After a success, the winning argument looks obvious and the doubts are gone. After a failure, the warning signs seem to have been in plain view.

Trying hard to be honest does not help. The people in Fischhoff's study, once told the outcome, still believed they were reporting what they had thought before. Nobody can feel their memory change, so even an honest, experienced, senior manager makes the same mistake. Only a note written before the outcome, and left alone afterwards, keeps your earlier belief.

How it works

1. Write the decision in one sentence, with today's date: what you choose and what you give up.
2. List the real alternatives and why you rejected each.
3. Write your prediction. What will you see, by when, and how sure are you, as a number?
4. Name the two or three assumptions the choice depends on. Mark the one you are least sure of.
5. Set a date to reread the entry, or an event that would bring it forward.
6. On that day, reread the entry before you look at what happened. Judge the reasoning and the result separately, and do not edit the entry.

Worked example

The psychologist Philip Tetlock kept such records for forecasts, on a huge scale. Over about twenty years he gathered forecasts from 284 people paid to comment or advise on politics and the economy. Each forecast was a probability with a date, on a question inside or outside the forecaster's field. By 2003 he had 82,361 of them. He also recorded how each forecaster explained a miss, so the excuse could be checked against the original number.

In 2005 he scored the forecasts against what happened and published the results in *Expert Political Judgment*. The experts barely beat well-informed non-experts, and neither group beat simple statistical rules. Anyone judging the experts by the hits they remembered would have missed this. Tetlock saw it because every forecast had a date and could not be changed once the outcome was known.

In a Business Case Weekly case

The Nike case stops in June 2020, when Nike's board is asked to approve Consumer Direct Acceleration. Nike would cut its wholesale accounts, the shops that resell its products, and sell directly to customers. Judged by profit margins, the move looks obvious.

A journal entry will not decide the vote. But it can hold three things. First, what the plan assumes shoppers want. Second, where you expect Nike to stand on other shops' shelves by 2024. Third, at what quarterly figure you would say that assumption had failed. Measure it outside Nike's own shops and website. Then, whatever happens by 2024, you can check your reasoning against it.

In your answer

- "Today I choose X over Y. I reject Y because ..."

- "This rests on the assumption that ... I am least sure about ..."
- "By <date> I expect to see ... I am about <n>% sure."
- "I will look at this decision again when ... I will reverse it if ..."

Common misuse

The first mistake is to write a diary. It tells what you decided and why it felt right, but has no prediction and no review date. You learn nothing from it, because it holds nothing the outcome could prove wrong.

The second mistake is to judge by the result. You reopen the entry, see a good outcome and mark the decision right. A lucky guess gets filed as a sound one. But a decision can be well reasoned and still unlucky. The point is to judge the reasoning apart from the result.

A quick check: find the sentence in your entry that the outcome could have proved false. Then find the date you meant to check it. If either is missing, you have written a diary.

References

- Baruch Fischhoff, "Hindsight ≠ Foresight: The Effect of Outcome Knowledge on Judgment Under Uncertainty", *Journal of Experimental Psychology: Human Perception and Performance* 1(3), 1975, 288–299. Twelve pages, and the whole case for writing decisions down.
- J. Edward Russo and Paul J. H. Schoemaker, *Decision Traps: Ten Barriers to Brilliant Decision-Making and How to Overcome Them*, Doubleday, 1989. Traps eight and nine are on feedback and keeping records.

- Philip E. Tetlock, *Expert Political Judgment: How Good Is It? How Can We Know?*, Princeton University Press, 2005. Twenty years of predictions, scored, and what they say about experts.
- Farnam Street, "How a Decision Journal Changed the Way I Make Decisions". The popular template.

Decision Quality: six links

✓ Before you send · F Focus · R Relevant evidence · A Alternatives · M Manage uncertainty · E Execute and learn · 5 min

When you need it

The memo reads well, and you need its weakest point

Check the six links of a decision and fix the weakest: frame, values, alternatives, information, reasoning and commitment.

How to use it

1. Check that the frame and the goal describe the decision that really matters.
2. Check the alternatives, the information and the reasoning: is a link missing or weak?
3. Commit only when the weakest link is strong enough for what is at stake.

Try it in your answer

Give the six links a rough score, then improve the weakest one, not the strongest part of your answer.

Practise it on a case

Booking.com question 2

What it is

A chain breaks at its weakest link, however strong the other links are. A decision is the same.

In Decision Quality, you check a recommendation as a chain of six links. They are the right frame (what exactly is being decided), clear values and trade-offs, creative alternatives, reliable information, sound reasoning and commitment to action. The decision is only as good as the weakest of the six. So you do not fill in the six links like a form. You find the weakest one and spend the rest of your effort there, instead of polishing what you have already done well.

Where it comes from

The method grew out of decision analysis, a step-by-step way to think through hard choices. Ronald A. Howard of Stanford named the field in the 1960s. He set it out in his 1966 paper "Decision Analysis: Applied Decision Theory".

No single person invented the six-link chain. Consultants worked it out at Strategic Decisions Group (SDG), the firm Carl Spetzler and Howard founded in 1981. Today SDG says the method comes from more than 35 years of use inside large organisations. Carl Spetzler, Hannah Winter and Jennifer Meyer give the fullest written account. Their 2016 book from Wiley is *Decision Quality: Value Creation from Better Business Decisions*. The Decision Education Foundation teaches the same six elements to school pupils as "the decision chain". The wording differs a little from source to source. You will see "framing" for frame, and "useful information" for reliable information. The six links are the same.

What it corrects

Imagine you are reviewing good work. An analyst brings you a financial model, a market estimate with ranges and a clear favourite option. You judge the work by how deep it goes. You check the model, test the estimate against the worst cases and tighten the writing.

Even a careful reviewer checks only the links that are on the page. That is the mistake. The two links that fail most often are alternatives and values. When they fail, there is nothing on the page to check. When an alternative is missing, you see a confident comparison of one plan with doing nothing. When nobody names the goal, you just see the word "best". The analyst quietly assumes that everyone wants the same thing. Working harder does not fix this, because you only work on the links you can see. Reading the model twice makes the strongest link stronger. You still never see the option nobody wrote down.

How it works

1. Set out the frame: who decides, what choice is open, what is off the table, and over what period.
2. Write the goal and the trade-off as "more of X while protecting Y", not just the word "best".
3. List alternatives that are really different, not three versions of one plan.
4. Separate what you know from what you assume. Give a range for each assumption that matters.
5. Link the facts to the recommendation with reasoning that someone could argue against.
6. Name the action, who owns it, the first money or people you commit, and the date you will review it.
7. Score the six links quickly, then improve the weakest. Stop when more work on it would cost more than the decision could gain.

Worked example

On the evening of 27 January 1986, engineers at Morton Thiokol advised against launching the Space Shuttle Challenger. Their company made the shuttle's rocket boosters. Their worry was the cold. The seals in the booster joints had never flown below 53°F (about 12°C). The next morning at Kennedy Space Center was forecast to be near freezing. George Hardy of NASA said he was "appalled" by their advice. Thiokol's managers talked privately for about thirty minutes, then overruled their own engineers. Challenger was launched on 28 January and broke apart 73 seconds later.

Now read it as a chain. The information link held. Thiokol had flight data on how the seals behaved in the cold, and its engineers showed it. The link that broke was the frame: the question the meeting was answering. Weeks later one of the engineers, Roger Boisjoly, told the Presidential Commission on the accident: "This was a meeting where the determination was to launch, and it was up to us to prove beyond a shadow of a doubt that it was not safe to do so. This is in total reverse to what the position usually is." Normally, before a flight, someone must prove the shuttle is ready. That night the engineers had to prove it would fail, and nobody could. However good the information, the decision could be no better than the link that broke.

In a Business Case Weekly case

In the Booking.com case, it is mid-2002 and the founder, Geert-Jan Bruinsma, has three options. He can sell most of his shares to a Dutch investor group. Or he can keep them, and either hire a chief executive or run the company himself. The growth data is the easy link here, but it is not the weakest. Going through the six links, you must first say whose decision this is, because the founder's goal and the company's are not the same. Then you write down the trade-off.

On one side is money that would make his family secure. On the other are his shares, a bet on European travel that might not pay off. Only then can you see whether the list of alternatives is complete.

In your answer

- "The decision belongs to …, who must choose … before …; it does not cover …"
- "The goal is … while protecting …, and I will give up … to get it."
- "The weakest link here is … because …, so the next work goes there, not into the forecast."
- "I commit to …, paid for by …, and I will reverse it if … by …"

Common misuse

The common mistake is to turn the six links into six headings. You write a paragraph under each, call every link good enough, and the recommendation ends where it started. That way you get nothing from the method. The point is to rank the links, not to cover them all. Here is a quick test. Does your answer name one link as the weakest? Did that change anything in the recommendation? If not, you have written a table of contents, not checked the decision.

References

- Carl Spetzler, Hannah Winter and Jennifer Meyer, *Decision Quality: Value Creation from Better Business Decisions*, Wiley, 2016. ISBN 978-1-119-14467-0.
- Strategic Decisions Group, "Decision Quality" — the consultants' own account of the six links and the weakest-link rule.

- Decision Education Foundation, "The Decision Chain" — the same six elements in plain words, with worked decisions from school and everyday life.
- Report of the Presidential Commission on the Space Shuttle Challenger Accident, Volume I, Chapter 5 (1986) — the launch teleconference in testimony, readable in an evening.

Red team

Before you send · A Alternatives · M Manage uncertainty · 5 min

When you need it

Nobody who would argue against the memo has read it yet

Hand the plan to a separate group whose only job is to break it, and let them succeed.

How to use it

1. Write down the plan and give the challenger one clear job: break it.
2. Attack the plan's assumptions, what it rewards people for, what it depends on and its biggest claim about cause and effect.
3. Answer the attack with evidence, change the plan or say why you accept the risk.

Try it in your answer

Write the strongest attack an opponent would make on your answer. Then reply to it without making it weaker or sillier than it is.

Practise it on a case

Telegram question 2 · **Unitree** question 3 · **SEGA** question 3

What it is

Your team is about to approve a plan. First you hand it to a separate group with one job: break it. That group is a red team. Its members did not write the plan. They argue the other side, play the competitor or the regulator and hunt for facts the authors set aside. They bring back a short list of ways the plan fails, each with its evidence.

Where it comes from

Red teaming has no single inventor. It has three roots.

The name comes from war games. In 1812 Georg Leopold von Reisswitz built an early version of Kriegsspiel, a Prussian table-top war game. The Prussian army adopted the game in 1824. It started the habit of having someone play the enemy, and play to win. In American military and intelligence writing, your side is "blue" and the enemy "red". A 2009 US government handbook, *A Tradecraft Primer*, says analysts normally work from the "blue" side, the friendly forces. A red team "attempts to work in the environment of the hostile forces."

The written method comes from intelligence work. The same handbook describes two techniques that businesses now blend. In Devil's Advocacy (page 17), when everyone agrees, someone builds the best case for an alternative. In Red Team Analysis (page 31), analysts think like the enemy to avoid "mirror-imaging": assuming an opponent weighs a decision the way you would.

Training came later, after failure. The Pentagon reviewed the wars in Iraq and Afghanistan and found that officers were not questioning their own assumptions. So in 2004 the US Army opened the University of Foreign Military and Cultural Studies at Fort Leavenworth to train red teams. The Army later said it would stop funding the school from 1 October 2021 and move its $2.5 million budget elsewhere. The main book on red teams is Micah Zenko's *Red Team* (2015). He follows them through the military, intelligence services, the New York police and private firms.

What it corrects

Picture a team that has worked on one decision for four months. They are not careless. They know the objections they thought of and have an answer to each. But they read each new fact for one thing: does it fit the plan? Facts that fit count for more than facts that do not. The handbook's authors call this a dominant mind-set. They warn that people then tend to dismiss contrary facts or give them too little weight.

Asking the team to try harder does not fix this. They would only dig deeper into questions they already asked. Someone has to be given the job of finding where the plan fails, and be allowed to succeed.

How it works

1. Write down the recommendation and, under it, the assumptions it needs.
2. Pick a challenger from outside the team that wrote the plan. Their job: break it.
3. Choose the weakest-looking assumptions, written down or not. Rate the evidence for each: strong, thin or missing.
4. Play one opponent: a competitor, a regulator or a lender. Write their move in their words.
5. Report one finding: how the plan fails, and which risk was treated as too small.
6. Answer it: keep, change or drop the plan. Note which evidence decided it.

Worked example

In summer 2002 the US military held a joint war game, Millennium Challenge 2002. It lasted three weeks, cost $250 million and involved 13,500 people. Paul Van Riper, a retired Marine lieutenant general, led the red force. With far weaker technology, he refused to fight the war blue had prepared for. He sent orders by motorcycle courier and light signals, so blue's surveillance missed them. When the blue fleet came within range, he struck first with missiles from ground launchers, commercial ships and low-flying planes. Speedboats packed with explosives joined in. In the game, nineteen ships sank, the aircraft carrier among them. "The whole thing was over in five, maybe ten minutes," Van Riper said. He had found how the fleet could lose: masses of cheap weapons against a force that relied on its communications network.

The rest shows what happens when a red team is not allowed to win. The organisers refloated the sunken ships. They banned red from firing at V-22 Ospreys and C-130 planes, made it move its air defences and ruled out chemical weapons. Six days in, Van Riper stepped down. Blue won, as scripted.

In a Business Case Weekly case

The sharpest decision in the SoftBank case puts you on SoftBank's board in 2016. Its founder, Masayoshi Son, wants to buy the chip designer Arm for $32 billion, far above Arm's share price. That same year he is raising a $100 billion fund. A red team would not just say the price is high. It would name what SoftBank is betting on. Arm's strong position has to pay off in the coming computing boom, soon enough for SoftBank to afford the wait. It would ask which rivals must be wrong for that to hold. Then it would play a lender whose

loan to SoftBank falls due before the bet pays off. The case asks what checks you would demand for your vote. The red team shows which failure those checks are for.

In your answer

- "The strongest attack on my recommendation is ..., and it works through ..."
- "That attack succeeds if ... is true, so I need evidence that ..."
- "If I play the competitor, its first move is ..., and my plan answers it only by ..."
- "I go ahead because ... held up under attack / I change the plan because ... changed my mind."

Common misuse

The common mistake is a challenge meeting where people raise only objections the plan already answers. From outside it looks like the real thing: a meeting, a chosen sceptic, a list, a plan that survives. So ask: could the challenger have won? What was the red team allowed to find? Could anything on its list have changed the decision? If nothing could, the game was fixed, as in Millennium Challenge. The second mistake is calling in a red team before there is a recommendation, so it does the thinking instead of testing it.

References

- Micah Zenko, "Millennium Challenge: The Real Story of a Corrupted Military Exercise and its Legacy", War on the Rocks, 5 November 2015. The worked example, by the author of the main book on red teams.

- US Government, *A Tradecraft Primer: Structured Analytic Techniques for Improving Intelligence Analysis* (March 2009). Devil's Advocacy on page 17, Red Team Analysis on page 31. The method itself, an evening's read.
- Micah Zenko, *Red Team: How to Succeed by Thinking Like the Enemy* (Basic Books, 2015). Red teams that worked, and ones that were blocked.
- Bryce Hoffman, "U.S. Army Moves To Close Red Teaming University", 26 October 2020. What the Army built at Fort Leavenworth, and why it stopped funding it.