Survivor Bias

The Complete Report

Why the winners can’t tell you what works, what the arithmetic actually says, and the five questions to ask before you take anyone’s advice.

Chief Master Greg Moody, Ph.D.

August 8, 2026

KarateBuilt Martial Arts · Rev Marketing · Martial Arts Wealth Mastery

Survivorship Bias

How Missing Cases Distort What We Think We Know

A deep report on the statistics, psychology, causal logic, historical examples, and practical methods for detecting decisions built from the survivors rather than the full population.

Core idea: The visible cases aren't automatically representative cases. Whenever survival, success, publication, retention, profitability, graduation, continued participation, or simple visibility determines what reaches you, the selection process itself becomes part of the evidence.

Report focus	Decision-making, business, investing, science, medicine, education, media, organizations, and everyday reasoning
Primary question	Who or what is missing from the sample, and why?
Central warning	Common among winners doesn't mean predictive of winning.
Best defense	Recover the denominator and compare survivors with non-survivors.
Research note	The famous Abraham Wald aircraft example is historically real in substance, but popular retellings simplify the underlying analysis and documentation.

Executive Summary

Survivorship bias is the systematic error that occurs when analysis concentrates on cases that passed through a selection process while overlooking cases that didn't. The classic examples involve literal survival, but the mechanism is much broader: a company can "survive" by remaining in business, a mutual fund by staying in a database, a research result by being published, a customer by remaining active, an employee by staying with a company, a student by remaining enrolled, and an online story by attracting enough attention to remain visible.

The danger isn't merely that some observations are absent. The danger is that the probability of being observed is often related to the outcome being studied. Once that happens, the observed group can systematically misrepresent the original population. A strategy can look effective because failures disappeared. A behavior can appear safe because people harmed by it are unavailable to report their experience. A trait can look characteristic of success because researchers never measure how common that trait was among people who failed.

The practical lesson is straightforward but demanding: before learning from winners, identify everyone who had a realistic chance to become a winner. Then determine what separated the groups. The strongest analysis isn't "What do successful people have in common?" It is "What do successful people do more often than comparable unsuccessful people, and does that difference remain after plausible alternative explanations are considered?"

> **The fastest survivorship-bias test: Show me the winners. Then show me everyone who started with them.**

1. What Survivorship Bias Is

Survivorship bias is a form of selection bias. It occurs when a sample is defined, directly or indirectly, by having survived a process and conclusions are then generalized beyond that selected group. "Survival" need not mean remaining alive. It can mean remaining observable, eligible, profitable, employed, enrolled, published, listed, active, famous, or otherwise present in the dataset.

A useful formalization starts with a target population, a selection mechanism, and an observed sample. Let S denote selection into the observed sample. If the probability of S=1 depends on variables related to the outcome, then conditioning analysis on S=1 can distort associations that existed in the target population. Modern causal-inference literature treats many such problems as selection mechanisms that can include restriction on a collider or its descendant. (Lu et al., 2022; Munafò et al., 2018)

The probability mistake

A recurring reasoning error is confusing the prevalence of a behavior among successful people with the probability of success among people who exhibit that behavior. If X is a behavior and Y is success, then the useful decision question is often P(Y | X): the probability of success given X. Success stories frequently give us P(X | Y): the probability of seeing X among people who succeeded. These quantities can differ radically.

For example, suppose 80 of 100 successful founders worked more than 60 hours per week. That makes long hours common among successful founders. But if 9,000 of 10,000 failed founders also worked more than 60 hours, extreme hours aren't distinguishing success; in this hypothetical sample they are more common among failures. The observation "80% of winners did X" is therefore incomplete until we know the base rate of X among non-winners.

2. Why the Bias Is So Convincing

Survivorship bias exploits the normal way human attention works. Visible examples are cognitively available. Missing examples are, by definition, harder to imagine, count, interview, photograph, quote, or put on a stage. The surviving cases therefore feel like the population even when they are a filtered residue of a much larger process.

Several psychological tendencies magnify the problem. Availability makes memorable success stories easier to retrieve. Hindsight makes the path to success appear more coherent after the outcome is known. Outcome bias causes us to evaluate the quality of a decision partly by how it turned out. Narrative bias encourages us to convert complicated, probabilistic histories into simple causal stories with protagonists, turning points, and lessons.

Success narratives also tend to be generated retrospectively. A successful leader may sincerely believe one principle caused the outcome while underweighting timing, competitors, market conditions, inherited advantages, network effects, fortunate introductions, regulation, or random variation. This doesn't make the leader dishonest. It means autobiographical explanation is evidence about experience, not automatically evidence about causation.

3. Abraham Wald and the Aircraft Example

The best-known illustration comes from Abraham Wald and the Statistical Research Group at Columbia University during World War II. Wald authored a series of memoranda on estimating aircraft vulnerability from damage observed on surviving planes. The original 1943 work was later reprinted by the Center for Naval Analyses, and a 1984 Journal of the American Statistical Association article by Marc Mangel and Francisco Samaniego reviewed and explained Wald's aircraft-survivability work. (Wald, 1943/1980; Mangel & Samaniego, 1984)

The intuitive problem is familiar. Returning aircraft show bullet damage in some areas more often than others. If we inspect only the returning aircraft, it is tempting to reinforce the areas with the most holes. But the fact that those aircraft returned demonstrates that damage in those regions was compatible with survival. Areas with relatively little observed damage may be precisely the areas where hits were more likely to prevent an aircraft from returning, so the missing aircraft contain information that the survivors can't directly display.

> **Wald's deeper lesson isn't "armor the spots without holes." It is: infer the selection mechanism that produced the sample before interpreting the pattern inside the sample.**

Historical nuance

The modern internet version often presents the episode as a neat one-step confrontation in which military officers wanted armor on the bullet holes and Wald instantly reversed the recommendation. The historical record supports Wald's work on vulnerability using data from survivors, but later historians of mathematics have cautioned that the familiar meme compresses a much richer technical and institutional story. The American Mathematical Society has explicitly described its treatment as an investigation of "what we do and don't know" about the legend. [3–5]

That nuance strengthens rather than weakens the lesson. The important fact is that Wald developed methods for inference from selectively observed survivors. The popular picture is useful pedagogy; the underlying achievement was a statistical approach to missing information.

4. The Mechanics: Selection Changes the Meaning of the Data

Imagine a population of 10,000 attempts. A process removes 9,500 of them from observation. We now study the remaining 500. If removal is random with respect to the variables we care about, the damage may be limited. If removal is related to performance, risk, treatment response, profitability, or another relevant factor, the survivors become a systematically selected group.

The key analytical question isn't merely "Is data missing?" It is "Why is data missing?" Cochrane's risk-of-bias guidance makes this point in clinical research: missing outcome data can bias intervention estimates, and the magnitude depends not just on how much is missing but on the relationship between missingness and the unobserved outcomes. (Higgins et al., 2024)

A numerical illustration

Group	Used strategy X	Did not use X	Total
Succeeded	80	20	100
Failed	800	100	900

Among winners, 80% used strategy X. A survivor-only analysis sounds impressive. Yet the success rate among users of X is only 80 / 880 = 9.1%, while the success rate among non-users is 20 / 120 = 16.7%. In this constructed example, the very behavior that appears dominant among winners is associated with a lower probability of success. The denominator changes the conclusion.

5. Survivorship Bias and Its Close Relatives

Concept	How it relates
Selection bias	The broad family: inclusion in the analyzed sample is systematically related to variables that affect the estimate.
Attrition bias	Participants drop out, disappear, or are lost to follow-up in ways related to treatment, exposure, or outcome.
Publication bias	Studies with positive or statistically significant results are more likely to be published and therefore visible.
Outcome/reporting bias	Within a study, favorable outcomes or analyses may be reported more completely than unfavorable ones.
Look-ahead bias	An analysis inadvertently uses information that wouldn't have been available at the historical decision point.
Outcome bias	The quality of a decision is judged from the result rather than from information and process available when the decision was made.
Availability bias	Memorable and readily retrievable examples are treated as more representative or probable than they are.
Collider bias	Conditioning on a common effect of two variables can induce a spurious association between them; selection into a survivor sample can create this structure. (Lu et al., 2022; Munafò et al., 2018)

These terms should not be treated as interchangeable. Survivorship bias describes a particular selection pattern. Attrition and publication are mechanisms that create selectively visible evidence. Collider bias is a causal-structure problem that can arise through selection but can also arise through statistical adjustment. Keeping the distinctions clear helps identify the correct remedy.

6. Business: Why Studying the Companies Still Standing Can Mislead

Business analysis is unusually vulnerable because failed firms disappear from ordinary observation. Websites are taken down, founders stop posting, employees scatter, records become harder to obtain, and the company no longer appears in directories of active firms. Meanwhile, surviving companies continue generating interviews, case studies, conference talks, books, and consulting frameworks.

U.S. Bureau of Labor Statistics Business Employment Dynamics data illustrate the filtering process. In selected cohorts, five-year startup survival rates ranged from 49.8% for establishments born in 2006 to 57.3% for those born in 2018. In other words, after five years the visible population is already substantially different from the population that started. (U.S. Bureau of Labor Statistics, 2024)

The “great companies do X” trap

Suppose researchers interview 50 companies that have operated successfully for 25 years and discover that most emphasize culture, customer service, disciplined cash management, and long-term thinking. Those may be excellent practices. But survivor-only data can’t tell us whether those traits differentiated the survivors from firms that failed, were acquired, pivoted into another identity, or exited for reasons unrelated to management quality.

A stronger design starts with a cohort at a common historical point and follows both survivors and non-survivors. It measures candidate practices before outcomes are known, then tests which variables predict continued operation, profitability, growth, or another clearly defined outcome. That is less glamorous than interviewing winners. It is also much closer to answering the causal question.

Customer success stories

Testimonials establish possibility, not probability. “This customer increased revenue 64%” demonstrates that the outcome occurred. It doesn’t show the expected effect for a new customer. To estimate expectation, decision-makers need the number who started, the number who completed implementation, median and mean outcomes, dispersion, cancellations, failures, and the characteristics that distinguished high performers from low performers.

> **A case study answers “Can this happen?” A representative outcome distribution answers “How often does this happen?” Do not confuse the two.**

7. Investing and Finance: Survivorship Bias in Databases and Backtests

Finance provides unusually clean empirical demonstrations because funds and securities enter and leave databases. Poor-performing mutual funds are more likely to merge or disappear, so a database containing only funds that remain alive at the end of a sample period can overstate historical performance. Elton, Gruber, and Blake explicitly tracked

funds that existed at the end of 1976 to estimate the bias created when disappearing funds are omitted. (Elton et al., 1996)

The logic extends to trading strategies and stock backtests. If a researcher takes today's successful companies and runs their performance backward, the portfolio has been built with future knowledge: it excludes businesses that later went bankrupt, delisted, shrank, or otherwise failed to qualify for today's survivor set. That can make historical results look much better and safer than a real investor could have achieved.

A sound historical backtest requires point-in-time membership, delisted securities, realistic transaction costs, corporate actions, and rules that use only information available at each historical decision date. Survivorship bias and look-ahead bias frequently reinforce one another, producing an attractive equity curve that belongs more to the researcher's database than to a strategy that could actually have been traded.

8. Medicine and Clinical Research: The Patients Who Disappear Matter

In treatment research, analyzing only participants who complete therapy or provide final outcome data can create an overly favorable picture if those who discontinue differ systematically from completers. A person may leave because of side effects, burden, lack of benefit, worsening symptoms, logistical problems, or unrelated events. The direction of bias depends on why the data are missing.

Cochrane provides a striking numerical example: even 10% missing outcome data can substantially alter an estimated mortality rate if the missing participants have a different underlying risk. Its guidance therefore emphasizes reasons for missingness, between-group differences in missing data, and sensitivity analyses rather than simple thresholds. (Higgins et al., 2024)

The principle is broader than randomized trials. Registries based only on patients who return for follow-up can misrepresent long-term outcomes. Studies based only on people who reach a hospital can distort risk-factor relationships if hospital admission itself is influenced by multiple causes. Screening programs, longitudinal cohorts, psychotherapy outcome studies, and public-health surveillance all face versions of the same problem.

9. Publication Bias: When Evidence Must "Survive" to Enter the Literature

Scientific evidence passes through another selection mechanism: publication. Studies with statistically significant, positive, or striking findings are often more likely to be submitted, accepted, published quickly, cited, and incorporated into reviews. Null and unfavorable findings can disappear into file drawers, conference abstracts, abandoned manuscripts, or unreported outcomes.

A Cochrane methodology review found that trials with positive findings had nearly four times the odds of publication compared with negative or null findings (odds ratio 3.90, 95% CI 2.68–5.68). Positive findings also tended to be published sooner. (Hopewell et al., 2009) Cochrane's current handbook summarizes broader evidence that selective publication can distort systematic reviews and effect estimates. (Cochrane, 2024)

Publication bias isn't identical to survivorship bias, but it is structurally similar: the literature available to the reader is a selected subset of all conducted studies. If selection depends on the result, the visible evidence base can become more favorable than the underlying research record.

10. Education, Careers, and the Celebrity-Dropout Error

The famous-college-dropout argument illustrates denominator neglect. Pointing to Bill Gates, Steve Jobs, or Mark Zuckerberg establishes that exceptional entrepreneurial success is possible without completing college. It doesn't estimate the probability of exceptional success among people who leave college, nor does it compare outcomes for otherwise similar people who complete degrees.

The celebrity cases are memorable because they are unusual. The relevant dataset would include ordinary dropouts, unsuccessful founders, salaried employees, graduates who became founders, graduates who never founded companies, and differences in background, ability, capital, timing, and opportunity. A visible exception can't substitute for the distribution.

The same error appears in career advice: "I never used my degree," "I ignored every critic," "I worked for free until someone noticed me," or "I moved to a major city with no backup plan." These stories can contain useful tactics. They can't establish the risk-adjusted value of the tactic without data on the people who tried the same thing and were never featured in the success story.

11. Safety and Everyday Risk: "I Did It and I'm Fine"

Personal anecdotes become especially unreliable when the outcome affects whether a person is available to give the anecdote. "We never wore helmets and we survived," "I drove that route for years without a seat belt," or "My grandfather smoked until 95" select precisely the people whose experience didn't remove them from the conversation.

This doesn't mean the anecdote is false. It means the anecdote has been selected on the outcome. Risk assessment requires rates among all exposed individuals, comparison groups, severity distributions, and time at risk. Exceptional survivors are often the least useful cases for estimating ordinary risk.

12. Social Media: Survivorship Bias With an Attention Algorithm

Social media adds a second selection layer on top of outcome selection. First, unusually successful people are more likely to post a result because the result is notable. Second, unusually dramatic posts are more likely to receive engagement and algorithmic distribution. The user therefore doesn't see a random sample of attempts; the user sees a selected sample of selected samples.

Consider 50,000 people trying a marketing tactic. Perhaps 200 experience extraordinary results, 1,800 experience modest results, and 48,000 experience nothing worth posting. If 50 of the extraordinary cases create posts and 10 become viral, the feed may contain ten compelling demonstrations and almost no visible denominator. Repeated exposure can create a strong intuitive belief that the tactic is common, easy, or reliable.

The same mechanism operates in wealth, fitness, dating, publishing, investing, parenting, productivity, and health content. Extraordinary outcomes aren't only more visible; they are

more monetizable. The business model of attention can therefore amplify the exact observations most likely to distort base-rate judgment.

13. Architecture, Products, and “They Don’t Make Them Like They Used To”

Old objects that remain visible are often unusually durable members of their original cohort. A house built in 1900 and still occupied in 2026 has passed through 126 years of weather, maintenance decisions, redevelopment pressure, economic change, and structural wear. Inferior buildings from the same era are disproportionately absent.

Comparing the best surviving century-old buildings with the full distribution of new construction therefore mixes cohorts asymmetrically. The correct historical comparison would require information about buildings that were demolished, abandoned, condemned, destroyed, or uneconomical to maintain. Similar logic applies to antique tools, furniture, machinery, books, and consumer products. The museum isn’t a random sample of the factory.

14. Organizations: Retention, Employees, Customers, and Students

Organizations routinely ask current members what makes the organization valuable. That is useful for understanding current satisfaction, but it can be a poor way to understand retention because former members are absent by definition.

A company that surveys only employees who have stayed ten years may conclude that its culture rewards independence. Departed employees might reveal that the same environment felt unsupported. A subscription service that surveys active customers may discover strong product satisfaction while never learning that canceled customers found onboarding confusing. A school may study long-term students and identify traits associated with commitment, but those traits may have developed after commitment rather than caused it.

The solution is cohort analysis and exit analysis. Track people from entry, measure experiences over time, identify the point at which dropout risk increases, and compare similar people who diverge. The 30–90 days before cancellation, resignation, or withdrawal often contain more actionable information than a retrospective interview with long-term survivors.

15. Collider Bias: The More Technical Version of the Problem

A collider is a variable caused by two other variables. In a simple causal diagram, $A \rightarrow S \leftarrow B$, the arrows “collide” at S. If A and B are otherwise unrelated, conditioning on S can create an association between them. Selection into a survivor group can effectively condition on S, producing correlations that don’t exist in the original population. (Lu et al., 2022; Munafò et al., 2018)

A classic intuition comes from selective admissions. Imagine admission depends on both test scores and exceptional extracurricular achievement. Among all applicants, those two traits could be weakly related or unrelated. Among admitted students, however, a lower test score may tend to be paired with stronger extracurricular achievement because either

strength can help secure admission. Conditioning on admission creates a negative association inside the selected group.

Survivor samples can behave the same way. If survival depends on both skill and luck, then among survivors a person with less skill may need unusually favorable luck, while a highly skilled person can survive with less luck. Looking only at survivors can therefore make skill and luck appear negatively related even if they were independent in the starting population. This is one reason survivor-only data can produce counterintuitive or even reversed relationships.

16. Extreme Success, Skill, and Luck

Recognizing survivorship bias doesn't imply that success is merely luck. High performance can require substantial skill, effort, judgment, capital, and persistence. The analytical point is that selection on an extreme outcome can also select for favorable randomness, particularly when many competent people are competing and the outcome has a stochastic component.

At the extreme tail, nearly everyone may already possess high skill. Small differences in timing, opportunity, network position, market shocks, injuries, platform algorithms, or chance encounters can then have large effects on rank. Retrospective stories often convert that mixture into a clean formula because formulas are easier to teach than distributions.

A productive interpretation is neither "winners know the secret" nor "winners were lucky." It is: identify which variables are repeatable, controllable, and predictive across both successful and unsuccessful cases, and distinguish them from one-time favorable conditions that can't be reliably reproduced.

17. Do Not Overcorrect: Survivors Still Contain Valuable Information

Once people learn about survivorship bias, they sometimes dismiss all success stories. That is another mistake. Survivors are evidence. They can reveal feasible strategies, operational details, rare combinations, failure-resistant designs, and hypotheses worth testing. The problem is treating them as a representative sample or as sufficient proof of causality.

The correct move is comparative. Use survivors to generate hypotheses, then examine non-survivors to test whether the proposed difference actually distinguishes outcomes. A successful company's process may be worth copying if comparable failed companies didn't use it and if alternative explanations are reasonably controlled. The survivor story becomes stronger when embedded in the full population.

18. A Six-Question Diagnostic for Survivorship Bias

1. What was the original population? Define who or what had a chance to enter the outcome. Avoid starting from the cases that are currently visible.

2. What was the selection process? Identify the filters between the original population and the dataset: failure, dropout, closure, publication, acquisition, algorithmic attention, nonresponse, delisting, death, or another mechanism.

3. Who or what disappeared? Name the missing categories. If you can't name them, you probably don't understand the dataset yet.

4. Is disappearance related to the outcome or predictors? If selection is related to success, treatment response, risk, performance, or variables that also affect the outcome, bias becomes plausible.

5. What is the denominator? Count everyone who entered or was exposed, not merely the people who completed, responded, stayed, or succeeded.

6. What would I expect to observe if my preferred explanation were false? This counterfactual question prevents a compelling survivor story from becoming causal evidence too quickly.

19. An Evidence Ladder for Success Claims

Level 1. Anecdote: A successful person did X. Useful for generating an idea; very weak for estimating probability.

Level 2. Survivor pattern: Many successful people did X. Better description of winners, still missing the comparison group.

Level 3. Comparative association: Successful people did X more often than unsuccessful people from the same relevant population.

Level 4. Adjusted prediction: X predicts outcomes after accounting for major differences that could plausibly explain the association.

Level 5. Causal evidence: Changing X, through randomized, quasi-experimental, or strong longitudinal designs, changes outcomes in the expected direction.

Most motivational advice lives at Levels 1 and 2. Many useful business case studies live at Level 2 or 3. Strong strategic claims should aspire to Levels 3–5, depending on the cost of being wrong.

20. How to Design Data Collection That Resists Survivorship Bias

- Create cohorts at the beginning, not at the end. Define entrants when they enter the system, then follow them through success, failure, dropout, and censoring.
- Preserve failure records. Do not delete closed accounts, failed products, terminated campaigns, delisted investments, or discontinued experiments from analytical history simply because they are no longer operational.
- Record exit reasons. “Inactive” isn’t enough. Distinguish price, dissatisfaction, relocation, side effects, competing priorities, acquisition, bankruptcy, graduation, death, and loss to follow-up when appropriate.
- Measure variables before the outcome. Retrospective winner interviews are vulnerable to hindsight reconstruction. Baseline measurements let researchers test whether a characteristic preceded success.
- Use matched or comparable groups. A good comparison asks why similar cases diverged, not why a billionaire differs from the average citizen.
- Report distributions, not hero outcomes. Median, percentiles, completion rates, failure rates, confidence intervals, and worst-case outcomes often matter more than the top result.
- Run sensitivity analyses. Ask how strong or different the missing cases would need to be to reverse the conclusion.
- Keep point-in-time datasets for historical analysis. Especially in investing and operations, preserve what was known at each decision date.

21. Practical Applications by Domain

Business consulting. Before copying a "best practice," compare firms that adopted it with firms that didn't, including firms that later closed. Ask whether the practice preceded improvement or appeared after the company was already successful.

Marketing. Do not optimize only from winning ads. Preserve failed creatives, audiences, and offers. Otherwise the organization loses the denominator that reveals which patterns actually distinguish winners.

Sales. Studying only closed deals can produce false rules about buyer behavior. Compare won, lost, no-decision, and stalled opportunities from the same period.

Customer retention. Interview churned customers and non-adopters, not just promoters. Examine the sequence before churn rather than relying solely on exit explanations.

Hiring. Studying top employees can identify traits that are consequences of tenure. Include people who left, were terminated, or never reached expected performance.

Education and training. Track all entrants through stages and identify where dropout occurs. Compare similar learners who persist versus leave at each transition.

Investing. Use databases that include dead funds and delisted securities. Avoid building historical portfolios from today's constituents.

Clinical practice and research. Track non-completers, missed follow-ups, adverse effects, and reasons for discontinuation. Completion-only outcomes can be misleading.

Publishing. Do not infer a universal formula from bestselling books without a comparison set of books that used similar tactics and sold poorly.

Personal decision-making. When a friend presents an impressive exception, ask how many comparable people attempted the same path and what happened to them.

22. Red-Flag Phrases That Should Trigger a Denominator Check

- "Every successful person I know…"
- "The best companies all…"
- "Nobody I know has ever had a problem with…"
- "I did it for years and I'm fine."
- "Look at these ten people who became millionaires."
- "Our top customers all use this feature."
- "All of our longest-tenured employees say…"
- "The funds with the best 20-year records…"
- "This treatment works for the patients who stick with it."
- "Every bestseller does this."
- "We surveyed our current members."

None of these statements is automatically wrong. Each is simply incomplete until the selection mechanism and missing comparison group are understood.

23. Worked Example: From Success Story to Causal Question

Claim: "Our highest-retention customers attend the onboarding webinar."

Weak conclusion: "The webinar causes retention."

Survivorship-aware questions: Were customers who were already more committed more likely to attend? Did churned customers receive the invitation? Was attendance measured before the period in which retention was assessed? What is retention among invited non-attendees? What is retention among customers who wanted to attend but could not? Did implementation quality differ?

Better analysis: Define a cohort of all new customers. Measure webinar invitation, registration, attendance, baseline engagement, company size, acquisition source, implementation milestones, and churn over a fixed horizon. Compare retention across attendance groups while accounting for pre-existing differences. If feasible, randomize enhanced onboarding or use a natural experiment. The goal is to transform “survivors tend to do X” into “X changes the probability of survival.”

24. A Survivorship-Bias Audit for an Organization

Audit area	Question
Data retention	Do we preserve records for failures, cancellations, closed accounts, former employees, discontinued campaigns, and abandoned products?
Metrics	Do dashboards show only active populations, or do they include entry cohorts and attrition?
Case studies	Are exceptional successes labeled as examples rather than typical outcomes?
Decision reviews	Do we evaluate the quality of decisions made with the information available at the time, rather than only by outcome?
Exit intelligence	Do we systematically collect and categorize reasons for churn, resignation, dropout, and lost deals?
Historical analysis	Can we reconstruct what the organization knew at a past decision point without using future information?
Research culture	Are negative tests, failed experiments, and null results preserved and searchable?
Executive narratives	When leadership identifies a “success factor,” do analysts test whether failures also possessed it?

25. The Deeper Lesson: Absence Is Data

Most analytical habits focus on what is present: observations, interviews, transactions, survey responses, published papers, surviving businesses, active customers. Survivorship bias teaches a complementary discipline: study the process that determines what becomes present.

The missing aircraft in Wald's problem weren't merely unavailable observations. Their absence changed the interpretation of the damage on the aircraft that returned. The closed business, vanished fund, discontinued patient, unpublished experiment, former employee, canceled customer, and failed product play the same role. Each absence may be a clue about the selection mechanism.

That leads to a durable decision rule: Never ask only what winners have in common. Ask what separates winners from comparable non-winners, how many of each existed, what disappeared before the analysis began, and whether the act of becoming observable changed the relationships you are trying to understand.

> **Survivorship bias is the error of learning from the cases that remain visible without accounting for the cases that disappeared. The defense isn't cynicism. It is better denominators, better comparison groups, and better records of failure.**

Part II. Applied Business Analysis

Part I explains how survivorship bias works. Part II applies it where it costs the most money and gets named the least: the market for business advice. Owners almost never meet this problem as a statistical idea. They meet it as a confident, successful person across a table who is sincerely wrong about why they succeeded.

The sections run from documented cases, through why bad advice keeps getting produced, to what to change in an organization so it stops fooling itself.

B1. What Happened to the Business Books Built This Way

The clearest demonstrations are the business books that were built on survivor samples and then aged in public. All of them follow the same recipe: find companies or people who already won, write down what they share, and present the shared traits as the reason.

In Search of Excellence (1982)

Peters and Waterman selected 43 excellent American companies and derived the attributes they shared. Within about two years BusinessWeek ran a cover story titled "Oops! Who's Excellent Now?" reporting that a large share of the 43 had run into serious financial trouble. Atari had effectively collapsed. (Peters & Waterman, 1982; Oops! Who's excellent now?, 1984)

Twenty years on, Tom Peters published "Tom Peters's True Confessions" in Fast Company and described how the list was actually assembled: colleagues at McKinsey and other smart people were asked which companies were doing interesting work, and the quantitative

measures were built afterward. His line was “Okay, I confess: We faked the data.” He later said the phrasing was harsher than he intended and that varying research measures isn’t the same as fabricating numbers, which is a fair correction. The selection was still a room of people naming companies they admired. (Peters, 2001)

Built to Last (1994)

Collins and Porras surveyed hundreds of CEOs to identify 18 visionary companies and catalogued the habits those companies shared. Within roughly a decade a large share of the list had slipped badly, including Motorola, Ford, Sony, Disney, Boeing, Nordstrom and Merck. The companies qualified for study because they had already won. (Collins & Porras, 1994)

Good to Great (2001)

Collins identified 11 companies whose stock returns had already been exceptional over a defined window, then catalogued the leadership and cultural traits they shared. Circuit City filed for bankruptcy in 2009. Fannie Mae went into federal conservatorship in 2008. (Collins, 2001; Federal Housing Finance Agency, 2008)

The problem is in the first step, not in the later failures. The sample was defined by the outcome. Companies with the same humility, discipline and focus that failed were never eligible, because failing removed them before the research started. The study can tell you what great performers had in common. It can’t tell you that those traits are what separated them from everybody else.

The Millionaire Next Door (1996)

Stanley and Danko interviewed hundreds of millionaires and reported the habits they found: live below your means, drive a used car, invest steadily. The advice is sound. What the study never counted is the far larger group who did all three and never got there. Taleb pushed it further in Fooled by Randomness, noting that the sample also caught people who invested through one of the strongest bull markets on record. Two filters were running, not one. (Stanley & Danko, 1996; Taleb, 2001)

The 10,000-hour rule (2008)

Gladwell popularized it from research on elite performers, and the underlying finding holds: people at the top of demanding fields have accumulated enormous practice. What it can’t tell you is how many people accumulated the same hours and never got near the top, because the research started with the ones who did. Macnamara, Hambrick and Oswald ran a meta-analysis across many fields in 2014 and found deliberate practice accounted for roughly 26% of the variation in games, 21% in music, 18% in sports, 4% in education, and under 1% in professions. (Gladwell, 2008; Macnamara et al., 2014)

That last figure is the one to sit with. In professional work, practice hours explained less than one percent of the difference between people. Practice still matters and should still be done. It is nowhere near the whole story, and the reason a generation believed it was the whole story is that the research only ever looked at winners.

Enormous budgets, careful authors, real data, and the same broken first step every time. If it can happen to McKinsey partners and Stanford professors with a research team, it will happen to an owner with a legal pad at a Sunday seminar.

Rosenzweig named the error in The Halo Effect: the delusion of connecting the winning dots. He also answers the obvious defense, which is that these authors gathered enormous

amounts of data. If the sample is chosen by the outcome, more data doesn't fix it. Denrell showed the same thing formally: because failed organizations are undersampled in the stories managers learn from, observers consistently overrate risky, distinctive strategies, which are exactly the strategies that stand out among survivors. (Rosenzweig, 2007; Denrell, 2003)

B2. Why the Advice Market Produces Bad Advice

Advice isn't a random sample of experience. Four filters sit between what happened in the world and what reaches an owner's ears, and none of them filter for whether the advice is any good.

Filter	Who it removes	What it does to what is left
Survival	Anybody whose business closed	Their advice is unavailable at any price. Roughly half of establishments never reach year five. (U.S. Bureau of Labor Statistics, 2024)
Choosing to teach	Operators who succeeded and never speak, write, sell or present	Willingness to teach tracks personality and income model. It doesn't track whether the system works.
Audience and platform	Advice that is accurate but ordinary	Distribution rewards certainty, novelty and dramatic numbers. Correct, boring, heavily-qualified advice doesn't travel.
Memory	Whatever the advisor no longer remembers as important	Looking back at your own success is honest testimony about your experience. It isn't automatically evidence about cause.

Put those together and the odds an idea is sound, given that you heard it, aren't the odds the idea is sound. What reaches an owner has passed four filters, none of which selected for being right and at least two of which selected against it.

The consequence is uncomfortable and worth saying plainly. The most confident, most polished, most widely shared advice in any industry comes from the most heavily filtered end of the pool. Confidence is what survives in the advice market. It tells you nothing about whether the advice is true.

Ask an advisor how many people they have taught this to and what happened to all of them. If they won't answer, you have your answer.

B3. Why Borrowed Parts Fail Faster Than One Bad Idea

Collecting practices from several successful operators and assembling them feels safer than betting on one person. It is the opposite. Two failures stack on top of each other.

Every part brings conditions you can't see

A practice that worked for one operator worked inside that operator's market, staff, pricing, capital and personal ability. Adopting the practice doesn't import any of that. If a borrowed practice only transfers when your conditions happen to match, then borrowing several requires several separate matches at the same time.

Put a generous 0.6 chance on each single match. Five borrowed practices give 0.6^5, which is about 7.8%. The number is illustrative rather than measured, and the point behind it isn't. Assembling multiplies the risk instead of averaging it. Instinct says a basket of good ideas is safer than one idea. The arithmetic says the reverse.

Nobody built the connections

Most of what an operating system is worth sits in how the pieces hand off to each other. Pricing assumes a retention model. Retention assumes a curriculum sequence. Curriculum assumes a staffing ratio. Staffing assumes the revenue that pricing was supposed to produce. Every borrowed piece was tested against different neighbours than the ones it now has.

That is why the failure is so hard to diagnose and why owners stay in it for years. Each individual part can be shown to work somewhere, so the owner concludes the parts are fine and the problem must be effort, market or staff. The parts are fine. Nothing was built to connect them.

B4. Why Repeating a Result Is the Only Practical Proof

One successful business is a single observation, chosen because it succeeded, with the operator, the location, the market, the timing and the system all moving together. Nothing in it isolates the system.

Repeating the result is the closest thing to a controlled test an owner can actually get. Same operator, different market, different staff, different year. That deliberately varies the things that otherwise travel invisibly with a single site. It isn't a credential. It is a rough experiment, and a rough experiment beats a great story.

What the base rate says

BLS Business Employment Dynamics tracks five-year survival by birth cohort. Across the cohorts published in the 2024 spotlight it ran from 49.8% for 2006 births to 57.3% for 2018 births. (U.S. Bureau of Labor Statistics, 2024)

Birth cohort	**5-year survival**
1994	54.3%
2001	52.3%
2003	55.3%
2006	49.8%
2010	56.0%

2018	57.3%

Surviving five years is close to a coin flip, so it tells you almost nothing about the quality of a system. Assume independence, which is deliberately too generous, and three separate sites all reaching five years at p ≈ 0.53 happens by chance about 15% of the time. Three sites running at top-tier performance at once is far rarer.

A common owner correlates the outcomes and breaks the independence assumption, so treat the figure as direction rather than measurement. The direction is the whole point: repeating a result pushes it out of the range luck explains comfortably. A single result never leaves that range.

One result is a story. The same result in different conditions is a system. The difference isn't how impressive the number is. It is how much luck the evidence can absorb.

B5. Service Businesses Are the Hardest Case

In a service business the operator is part of what the customer buys. Delivery is performed by people whose skill, presence and relationships can't be separated from the result, so the system and the person running it get measured together in every single observation.

That produces a specific and expensive error. An owner with unusual personal ability can carry a weak system for years and sincerely credit the system. A second operator adopts the system without the ability and gets a different result. Both walk away believing something false: the first that the system works, the second that they are the problem.

Three observations pull the person apart from the system, in ascending order of strength:

- Does the result hold through a long absence, travel or illness, with the owner off the floor?
- Does it hold through staff turnover, especially losing a strong performer?
- Has a different operator reproduced it in a different market, working from documentation rather than from mentorship by the person who invented it?

The third is the strongest evidence available short of a designed trial, because it changes the operator and holds the system still. It is also the one almost never offered, for the obvious reason.

B6. The Plateau Owner as an Advisor

An owner who has held flat results for a decade is a survivor. Survival is what makes them visible, reachable and believable, because tenure reads as expertise. Their central conclusion comes from a single case that was selected for not growing, and that case is themselves.

It usually arrives as a ceiling claim: this is what the business pays, this is what the market supports, this is as good as it gets. Inside their own experience the claim is very well supported. Their experience is the entire sample.

The part that makes systems look irrelevant

Section 15 covered what happens when you restrict a group by something two different causes can produce. Ten-year tenure is exactly that. A business survives a decade on operational strength, or on favorable fixed conditions: low rent, a protected location, no real competitor, a second household income absorbing the shortfall. Either one is enough.

Look at owners who are still open after ten years and you find excellent operators grinding through hard markets next to weak operators sitting on easy ones. Put those two side by side and it looks like systems barely matter. They matter enormously. The group was assembled by survival, and survival accepted either ticket.

Arm an owner against this one specifically, because it is the observation a plateaued owner reaches for when defending the ceiling, and it is persuasive because the examples are real.

B7. Five Worked Examples

Each one is built so the arithmetic can be checked. In every case the survivor-only reading is reasonable on its own terms and the full picture reverses it.

Example 1, 80% of winners do X

The claim is true in the table below. It is also useless without the second row.

Group	Used X	Did not use X	Total
Succeeded	80	20	100
Failed	800	100	900
Total	880	120	1,000

Success rate among users of X: 80 / 880 = 9.1%. Among non-users: 20 / 120 = 16.7%. The dominant behavior among winners comes with roughly half the success rate. There is more evidence for NOT doing X than for doing it. Nothing in the original claim was false. The denominator was missing, and the denominator was the answer.

Example 2. Everybody successful works 70-hour weeks

This is the one that costs owners their health, so it is worth the arithmetic. Same 1,000 owners, counting hours instead of strategies.

Group	Worked 70+ hours	Worked under 70	Total
Succeeded	85	15	100
Failed	810	90	900
Total	895	105	1,000

Read the top row alone and the case is closed: 85 of the 100 successful owners are working 70-hour weeks. Read the second row and 90% of the owners who failed were working the same hours. Success rate among the 70-plus group: 85 / 895 = 9.5%. Among those under 70: 15 / 105 = 14.3%.

The hours didn’t separate the groups because both groups supplied them. A behavior present in winners and losers alike can’t be what split them apart, and another ten hours a week won’t find whatever did. The advice wasn’t a lie. It was a number with the denominator removed, delivered by somebody who believed it.

Example 3. Twelve success stories

Twelve is a large number of testimonials and a small number of outcomes. What matters is what twelve is a fraction of.

Enrolled	12 successes represent	What it means
40	30%	Strong. Worth investigating.
300	4%	About what a program doing nothing would produce
2,000	0.6%	The testimonials are measuring the clients, not the program

The advertisement is identical in all three rows. The twelve stories are equally real and equally verifiable. Only the denominator separates a program that works from a filter that collects people who were going to succeed anyway. Ask for the enrolled count first, because it is the number nobody volunteers.

Example 4. The ten-year perfect record

Start with 1,024 advisors and assume every one of them is guessing. Each year half happen to be right.

After year	Perfect records remaining
Start	1,024
1	512
2	256
3	128
5	32
7	8
10	1

At year ten, one advisor holds a flawless record produced entirely by chance. He isn't lying and the record checks out. It is still worthless as evidence of skill, because a room of pure guessers was guaranteed to produce roughly one of him. The error isn't believing the record. It is never asking how many people entered the contest.

This is why historical fund analysis requires survivor-bias-free databases. Poor performers merge or liquidate and leave the record, so a database of surviving funds overstates how the category actually did. (Elton et al., 1996; S&P Dow Jones Indices, n.d.)

Example 5. The feature your best customers all use

Claim: our longest-retained customers all use feature Z, so pushing adoption of Z will improve retention. Two different worlds produce identical data.

What is really happening	Why the numbers look the same	What happens if you force adoption
Z causes retention	Using Z makes the product more valuable, so people stay	Retention improves

Commitment causes both	Already-committed customers explore more and find Z; commitment is what keeps them	Nothing changes. You pushed a marker, not a cause.

Survivor-only data can't tell these apart, because the customers who would settle it have already churned out of the dataset. Separating them takes a cohort defined at entry, adoption measured before the retention window, and a comparison against people who were invited and didn't adopt.

B8. Building an Organization That Stops Fooling Itself

This is an operating habit, not an analytical skill. The controls below are ordered cheapest first.

Control	What to do	What it protects
Keep a loss file	Preserve failed campaigns, dead accounts, discontinued offers and cancelled programs instead of deleting them	The denominator. Deleting losers is the most common way an organization destroys its own evidence.
Report on cohorts	Report on everyone who entered in a period, not on everyone still active	Retention and churn numbers, which are otherwise calculated on a group that excludes the thing being measured
Record why people leave	Categorize reasons for cancellation, resignation and withdrawal. "Inactive" isn't a reason	Why you are losing people, which can't be recovered after the fact
Write down the expectation first	Note the expected result before launching an initiative and keep the note	Against rewriting history, which turns mixed results into confirmations
Keep point-in-time records	Preserve what was known at each decision date	Historical analysis, against judging past decisions with information nobody had
Review decisions, not just outcomes	Judge a decision against what was knowable at the time, separately from how it turned out	The quality of decision-making, which outcome-only review corrupts

B9. The Five Questions

Applicable to any advisor, consultant, coach, program or peer. They are diagnostic, not hostile. A strong advisor answers all five easily and will have volunteered two or three already.

#	Question	What a weak answer

		sounds like
1	How many times have you produced this result, in how many places, with different people running it?	Once, here, with me running it, offered as though repeating it were beside the point
2	How many people have you taught this to, and what happened to all of them?	A list of successes and no enrolled count
3	Did you do this before the result, or notice it afterward?	The practice is dated to the success rather than to a decision that came before it
4	What has to be true about my situation for this to transfer?	“It works for everybody,” which claims the practice is independent of market, staff and operator, and almost nothing is
5	What would I see if this advice were worthless?	No answer available, because the claim was built so that nothing could contradict it

Question five is the strongest and the least asked. An explanation that can’t be wrong isn’t a strong explanation. It will absorb whatever result the adopter produces, it will never be revised, and the adopter will spend years assuming the failure was theirs.

B10. Simpson’s Paradox Is Not the Same Thing

The admissions illustration in Section 15 gets confused with the most famous admissions dataset in statistics. They are different problems that happen to share a setting.

Bickel, Hammel and O’Connell examined Berkeley graduate admissions for fall 1973 and found an aggregate gap favoring men that largely disappeared, and slightly reversed once pooled correctly, after accounting for which departments women applied to. Women applied more often to departments that were harder for anyone to enter. (Bickel et al., 1975) That is Simpson’s paradox, and disaggregating fixes it.

Section 15 describes something else. There, restricting attention to admitted students creates a link between two traits that were unrelated among all applicants. Simpson’s paradox comes from combining groups. This comes from restricting to one. Both are worth knowing, and calling either by the other’s name points you at the wrong fix.

Quick Reference: The Survivorship-Bias Checklist

- Define the original population.
- Identify every major filter between entry and observation.
- Count non-survivors, non-completers, and non-responders.
- Ask whether selection depends on the outcome or its causes.
- Compare winners with comparable non-winners.
- Use point-in-time data for historical claims.
- Report medians and distributions, not only standout cases.

- Preserve negative results and failed attempts.
- Distinguish possibility from probability.
- Distinguish association from causation.
- Ask what evidence would exist if the preferred story were false.
- When in doubt, look for the graveyard.

References

Bickel, P. J., Hammel, E. A., & O'Connell, J. W. (1975). Sex bias in graduate admissions: Data from Berkeley. Science, 187(4175), 398–404. https://doi.org/10.1126/science.187.4175.398

Casselman, B. (2016). The legend of Abraham Wald. American Mathematical Society Feature Column. https://www.ams.org/publicoutreach/feature-column/fc-2016-06

Cochrane. (2024). Cochrane handbook for systematic reviews of interventions, Chapter 7: Considering bias and conflicts of interest among included studies. https://www.cochrane.org/authors/handbooks-and-manuals/handbook/current/chapter-07

Collins, J. (2001). Good to great: Why some companies make the leap… and others don't. HarperBusiness.

Collins, J., & Porras, J. I. (1994). Built to last: Successful habits of visionary companies. HarperBusiness.

Denrell, J. (2003). Vicarious learning, undersampling of failure, and the myths of management. Organization Science, 14(3), 227–243. https://doi.org/10.1287/orsc.14.2.227.15164

Elton, E. J., Gruber, M. J., & Blake, C. R. (1996). Survivor bias and mutual fund performance. The Review of Financial Studies, 9(4), 1097–1120. https://doi.org/10.1093/rfs/9.4.1097

Federal Housing Finance Agency. (2008). Conservatorship of Fannie Mae and Freddie Mac. https://www.fhfa.gov/conservatorship

Gladwell, M. (2008). Outliers: The story of success. Little, Brown.

Higgins, J. P. T., Savović, J., Page, M. J., Elbers, R. G., & Sterne, J. A. C. (2024). Assessing risk of bias in a randomized trial. In Cochrane handbook for systematic reviews of interventions (Chapter 8). https://www.cochrane.org/authors/handbooks-and-manuals/handbook/current/chapter-08

Hopewell, S., Loudon, K., Clarke, M. J., Oxman, A. D., & Dickersin, K. (2009). Publication bias in clinical trials due to statistical significance or direction of trial results. Cochrane Database of Systematic Reviews, MR000006. https://doi.org/10.1002/14651858.MR000006.pub3

Lu, H., Cole, S. R., & Howe, C. J. (2022). Toward a clearer definition of selection bias when estimating causal effects. Epidemiology, 33(5), 699–706. https://doi.org/10.1097/EDE.0000000000001516

Macnamara, B. N., Hambrick, D. Z., & Oswald, F. L. (2014). Deliberate practice and performance in music, games, sports, education, and professions: A meta-analysis. Psychological Science, 25(8), 1608–1618. https://doi.org/10.1177/0956797614535810

Mangel, M., & Samaniego, F. J. (1984). Abraham Wald's work on aircraft survivability. Journal of the American Statistical Association, 79(386), 259–267. https://doi.org/10.1080/01621459.1984.10478038

Moody, G. (2026a). The two pieces of advice that keep school owners broke. today.mastermoody.com. https://today.mastermoody.com/advice-errors-school-owners-make-2026-07-20.html

Moody, G. (2026b). The plateau lifecycle of a martial arts school. today.mastermoody.com. https://today.mastermoody.com/plateau-lifecycle-framework-2026-04-20.html

Munafò, M. R., Tilling, K., Taylor, A. E., Evans, D. M., & Davey Smith, G. (2018). Collider scope: When selection bias can substantially influence observed associations. International Journal of Epidemiology, 47(1), 226–235. https://doi.org/10.1093/ije/dyx206

Oops! Who's excellent now? (1984, May 11). BusinessWeek.

Peters, T. (2001, December). Tom Peters's true confessions. Fast Company, 53. https://www.fastcompany.com/44077/tom-peterss-true-confessions

Peters, T. J., & Waterman, R. H. (1982). In search of excellence: Lessons from America's best-run companies. Harper & Row.

Rosenzweig, P. (2007). The halo effect… and the eight other business delusions that deceive managers. Free Press.

S&P Dow Jones Indices. (n.d.). SPIVA U.S. scorecard. https://www.spglobal.com/spdji/en/research-insights/spiva/

Stanley, T. J., & Danko, W. D. (1996). The millionaire next door: The surprising secrets of America's wealthy. Longstreet Press.

Taleb, N. N. (2001). Fooled by randomness: The hidden role of chance in life and in the markets. Texere.

U.S. Bureau of Labor Statistics. (2024). Business employment dynamics twentieth anniversary! Spotlight on Statistics. https://www.bls.gov/spotlight/2024/business-employment-dynamics-twentieth-anniversary/home.htm

Wald, A. (1980). A method of estimating plane vulnerability based on damage of survivors (CRC 432). Center for Naval Analyses. (Original work published 1943)

Appendix A: Twenty Fast Examples

Entrepreneurship. Interviewing only founders whose companies still exist and calling their habits the formula for survival.

Mutual funds. Calculating long-run average performance using only funds still listed today.

Stocks. Backtesting today's index members as if the investor knew decades ago which companies would remain in the index.

Clinical treatment. Reporting outcomes only for patients who completed the full protocol.

Therapy. Judging a program by clients who stayed through termination while ignoring early dropouts.

Education. Studying black-belt students to infer what creates retention without comparing students who quit.

Hiring. Asking only high performers which interview traits predicted their success.

Leadership. Studying CEOs who made giant bets and survived, while the CEOs whose bets destroyed the company are absent.

Books. Reverse-engineering bestsellers without sampling books with similar design and marketing that sold poorly.

Restaurants. Interviewing restaurants open for 20 years to infer the secrets of restaurant survival.

Advertising. Analyzing only campaigns that exceeded target and deleting weak campaigns from the reporting system.

Sales. Looking only at closed deals to determine the "ideal" buyer journey.

Customer success. Surveying active subscribers to understand why people like the product, then using those results to explain churn.

Fitness. Inferring the safety or effectiveness of an extreme routine from visible people who tolerated it.

Safety. Using a healthy long-lived smoker as evidence that smoking risk is exaggerated.

Architecture. Using surviving 100-year-old houses to characterize the average quality of all houses built 100 years ago.

Technology. Studying platforms that achieved network effects and concluding that the observed launch tactic causes network effects.

Creative careers. Studying famous actors who moved to Los Angeles with no backup plan while ignoring the much larger invisible cohort who did the same.

Research. Synthesizing published studies when significant findings are more likely to reach publication.

Social media. Estimating the probability of a tactic succeeding from viral posts created by the few people with extraordinary results.

Appendix B: The Five-Question Worksheet

One page, run against a single advisor, program or peer before adopting anything they recommend. Fill in every field. A field you can't fill is itself the finding.

Advisor / source	
What is being recommended	
Times they have produced this result	
Different locations or markets	
Different operators running it (Y/N)	
Total taught, the denominator	
What happened to everyone taught	
Did the practice come before the	

result?	
What has to be true here for it to work	
What would prove this advice worthless?	

Scoring is blunt on purpose. Fewer than two independent repeats, or no denominator available, puts the recommendation at Level 1 or 2 on the ladder in Section 19. That doesn't make it wrong. It makes it something to test small, not something to fund.

Appendix C: Four Quick Calculations

Each one turns a survivor claim into a number you can check.

The claim	The calculation	What it tells you
"N% of winners did X"	successes with X ÷ (successes with X + failures with X)	The success rate among people who did X, which is the number that matters
"We have N success stories"	N ÷ total enrolled	The program's actual success rate, to compare against what the same people would have done alone
"N-year track record"	starting population × 0.5^n	How many equally perfect records pure chance would have produced
"Our best customers all do Z"	retention among invited non-adopters vs. adopters, same entry cohort	Whether Z causes anything or just marks people who were already committed

Document History

v1. Original deep report: statistics, psychology, causal logic, historical examples, detection methods (Sections 1–25, Quick Reference, Appendix A).

v2. August 8, 2026. Added Part II (B1–B10): the business books built this way including Peters's 2001 confession, Built to Last, The Millionaire Next Door and the 10,000-hour rule; why the advice market produces bad advice; why borrowed parts fail faster; repetition as proof; service businesses and the owner effect; the plateau owner as advisor; five worked examples including the 70-hour-week case; operating controls; the five questions; and a Simpson's paradox clarification. Added references (Peters & Waterman, 1982)–(Macnamara et al., 2014) and Appendices B and C. Part I is unchanged.

v2 prose rewritten in Dr. Moody's voice following his August 8 edit pass on the companion article.