PARTNER IN GROWTH
LONDON · UK & EUROPE

THE DIAGNOSTIC · EDITION 01

The SaaS Growth Audit *Framework.*

A structured method for diagnosing growth blockers across your funnel — from acquisition to expansion — and for finding the *one* constraint that is actually capping your growth rate.

6 LAYERS | 48 DIAGNOSTICS | 80+ BENCHMARKS | 90-DAY PLANS

Hilal Tasdan
FRACTIONAL CMO · PARTNER IN GROWTH
PARTNERINGROWTH.CO.UK

Most growth problems are misdiagnosed before a single pound is spent fixing them.

A founder tells me pipeline is down, so we should spend more on ads. Four questions later it turns out the ads work fine — 71% of last year's churn came from a segment the company was never built to serve. More ads would have made the problem worse, faster.

That is the entire reason this document exists. Not to give you tactics. To give you a way of finding out **which** tactics are worth anything to you right now, and which are an expensive way of avoiding the real problem.

WHO THIS IS FOR

- B2B SaaS founders running between £500K and £20M ARR
- Teams where growth has flattened and nobody agrees on why
- Companies about to hire their first marketing leader
- Anyone about to sign a 12-month agency retainer

WHO THIS IS NOT FOR

- Pre-product companies — you need customers, not a scorecard
- Teams looking for 47 growth hacks
- Anyone who wants the answer to be "more content"
- People who will score themselves generously

The honest disclosure. I run a Fractional CMO practice. This document is genuinely useful whether or not you ever speak to me — it is the same diagnostic I run in paid engagements, written out in full. If you work through it and want a second pair of eyes on your scores, that offer is on the last page. If you never do, you still leave with a working diagnosis.

Contents

TIME TO COMPLETE

4 hours

Alone, with your dashboards open. Honest scores only.

BETTER VERSION

Two weeks

Score it yourself, then have three colleagues score it blind. Compare.

WHAT YOU LEAVE WITH

One number

Your constraint layer — and a 90-day plan for it alone.

The method.

Before you score anything, three ideas have to land: why most audits produce a to-do list instead of a diagnosis, why your growth rate is set by one layer and not by all six, and why a benchmark is worthless until you have defined the term it measures.

Why most growth audits produce a to-do list instead of a *diagnosis.*

Growth frameworks split cleanly into two families that most audits treat as one.

FAMILY ONE

Structural

Do the pieces of this business fit together well enough to compound at all?

Balfour's Four Fits. Growth loops. The Racecar framework. These ask whether your market, product, channel and price can hold together. They diagnose. They do not prescribe.

FAMILY TWO

Operational

Given that the pieces fit, how do we squeeze more out of them?

AARRR. North Star metrics. ICE, RICE and PXL. A/B testing programmes. These are excellent tools for improving a machine that already works, and useless for one that does not.

THE FAILURE MODE

Almost every disappointing growth audit applies operational tools to a structural problem.

You end up with a 40-item backlog of landing-page tests for a company whose real problem is that a £30K ACV product is being sold through a self-serve funnel. Every item on the list is individually defensible. Collectively they are a very organised way of not addressing the thing that is actually wrong.

This framework gates that. **Layer 0 has to hold before anything below it is worth your quarter.** If it does not, the rest of the document tells you what you are not yet ready to optimise — which is a finding in its own right, and a cheaper one than eighteen months of tests.

SYMPTOM

"Our conversion rate is low."

Usually structural. A page converts badly most often because the wrong people are landing on it.

SYMPTOM

"Leads are poor quality."

Almost always structural. Lead quality is an ICP definition problem wearing a marketing costume.

SYMPTOM

"CAC keeps rising."

Check the model first. Rising CAC in a channel your ACV cannot fund is not a bidding problem.

The failure has been measured. It is called *premature scaling.*

Startup Genome analysed 3,200+ high-growth internet startups and found a single dominant failure pattern: not weakness in one dimension, but **inconsistency between dimensions** — scaling one part of the company ahead of the others.

~70%
of startups in the dataset scaled prematurely on at least one dimension

74%
of high-growth startup failures attributed to premature scaling

20×
faster growth at scale stage for dimensionally consistent companies

Startup Genome, *Report Extra: Premature Scaling* (2011) — still the largest study of the pattern, though observational and largely self-reported.

DIMENSION	WHAT RUNNING AHEAD LOOKS LIKE	MEASURED MARKER
Customer	Buying traffic before the ICP is validated	2.3× more likely to spend >1 SD above average on acquisition
Product	Building features instead of validating demand	3.4× more code written during the discovery phase
Team	Hiring specialists for a process that isn't repeatable yet	3× larger teams at the same stage
Financials	Raising ahead of proof, then having to deploy it	~3× more capital raised before scaling
Model	Monetising broadly before the value is proven	Up to 3× more customers monetised before value was proven

> **Read this as permission, not criticism.** The finding is not that spending is bad or that hiring is bad. It is that spending, hiring, building and raising each have a *prerequisite*, and the companies that failed had skipped one. The audit that follows is essentially a structured way of checking which prerequisite you have skipped.

AND THE COROLLARY

If dimensional inconsistency is the failure mode, then the useful output of an audit is not "here are 40 things to improve." It is **"here is the one dimension that has run ahead, and here is the one that has fallen behind."** Two findings. That is what a good diagnosis looks like.

The Growth Audit Stack

Six layers. Each one takes the output of the layer above it as its input, which is why a weakness anywhere compounds downward — and why fixing the wrong layer changes nothing.

		OWNS	FAILS AS
0	**Fit** Do your market, product, channel and price actually fit each other?	Positioning, ICP, category, channel–model match	"Our leads are low quality"
1	**Acquisition** Can you reach the right people repeatably, at a cost your ACV can fund?	Channel concentration, CAC, pipeline sourcing	"We need more top of funnel"
2	**Activation** Does a new user reach real value before they give up on you?	Onboarding, time-to-value, the activation event	"People sign up and vanish"
3	**Retention** Does the value hold once the novelty and the onboarding attention stop?	Gross retention, cohort shape, churn reasons	"We're growing but net is flat"
4	**Monetisation** Are you charging for the thing that actually creates the value?	Value metric, packaging, discounting, price reviews	"Everyone negotiates us down"
5	**Expansion** Do existing customers become an engine, or an endpoint?	Expansion ARR, referral, land-and-expand motion	"Growth stops when spend stops"

Layer 0 is a gate, not a stage. Layers 1–5 are sequential stages of a customer's life. Layer 0 is the precondition for all of them. That is why it is numbered zero and why it is scored first — a company that scores badly at Layer 0 will produce misleading scores everywhere else, because it is measuring a funnel that is pointed at the wrong people.

What each layer borrows from.

This framework is a synthesis, not an invention. Where a layer leans on published work I have said so, because you should be able to go and read the original — and because frameworks presented as proprietary usually are not.

ATTRIBUTION MATTERS

Growth Loops is a four-author piece — Brian Balfour, Casey Winters, Kevin Kwok and Andrew Chen — not Balfour alone. RICE is Sean McBride at Intercom. The High-Expectation Customer is Julie Supan's concept, not Rahul Vohra's. The term "North Star Metric" traces to Sean Ellis; the North Star *Framework*, with its input tree, is Amplitude and John Cutler. Small things, but a practitioner who gets them wrong probably has not read the source.

SOURCES BY LAYER

LAYER	PRINCIPAL SOURCE
0 Fit	Brian Balfour, *Four Fits for $100M+ Growth* (2017)
	April Dunford, *Obviously Awesome* (2019)
	Moesta & Spiek, Forces of Progress
	Sean Ellis, *The Startup Pyramid* (2009)
1 Acquisition	Balfour, *Product–Channel Fit* (2017)
	Rachitsky & Hockenmaier, *Racecar Growth Framework*
2 Activation	Elena Verna, the B2B PLG escalator
	Winning by Design, *The Bowtie*, onboarding stage
3 Retention	Winning by Design, *The Bowtie*, right side
	Papp & Petit, RARRA (2017)
4 Monetisation	Price Intelligently, lever study (N=512)
	PricingSaaS / Maxio, pricing-change tracking
5 Expansion	Winning by Design, four expansion mechanisms
	Balfour et al., *Growth Loops are the New Funnels*

THE GAP THIS FILLS

The Four Fits are strategically excellent and have no measurement layer. The Bowtie has a superb measurement layer and assumes data maturity that most pre-Series-B companies do not have. This document connects the two and adds a scoring method that works with the data you actually have.

The Constraint Rule

Your growth rate is not the average of your layers — it is their product. Which means improvement is available anywhere, and it is *cheapest* at the weakest one.

THE LEAK MATH

Take a company where each of the five stage layers converts at 50%, except one that converts at 10%.

`0.50 × 0.50 × 0.10 × 0.50 × 0.50 = 0.625%`

Improve a *strong* layer by 50% — from 0.50 to 0.75.

`result → 0.94%` (+50%)

Improve the *weak* layer by the same 50% — 0.10 to 0.15.

`result → 0.94%` (+50%)

Identical — and that is the point. A percentage improvement anywhere gives the same percentage improvement overall. But a 50% lift on a layer already at 0.50 means reaching 0.75, which is close to the ceiling of what anyone achieves. A 50% lift on a layer at 0.10 means reaching 0.15, which is ordinary. **The weak layer is not more valuable in theory; it is far cheaper in practice.**

RULE ONE

Work the lowest layer only

For one full quarter. Not the lowest three. The discipline is the intervention — most teams already know their weak layer and work on everything else because it is more comfortable.

RULE TWO

Work upward, not downward

If two layers tie, fix the higher-numbered one last. Upstream problems poison downstream measurements; the reverse is not true. Bad-fit customers acquired in January are your churn number in December.

RULE THREE

Re-score, do not re-plan

At the end of the quarter, score again. The constraint will usually have moved to a different layer. That is success. Chasing a new constraint every quarter is the operating rhythm, not a sign of drift.

> **The uncomfortable version.** If your lowest layer is Layer 0, you do not have a marketing problem and you should not be hiring a marketer, an agency, or me. You have a positioning and ICP problem, it is founder work, and no amount of spend substitutes for it. Buying growth on top of broken fit is the single most expensive mistake in this document.

Fix your definitions before you fix anything *else.*

Published SaaS benchmarks for the same metric vary by up to five times. Almost none of that variance is real performance. It is definition and sample composition.

METRIC	PUBLISHED MEDIAN	SAMPLE	WHY IT DIFFERS
Net revenue retention	**82%**	~2,700 B2B SaaS, ChartMogul (2025)	Measured from billing data; long tail of small, low-ACV, self-serve products
	101%	~1,000 B2B SaaS, Benchmarkit (2025)	Self-reported; skews VC/PE-backed and $5M+ ARR
	102%	1,500+ private B2B, SaaS Capital (2023)	Self-reported; a lender's borrower base
	110–120%	~100 growth-stage, ICONIQ (2025)	Late-stage enterprise only; survivorship-selected

Four credible sources. A 38-point spread. None of them is wrong; they are measuring different populations.

THE THREE WORST OFFENDERS

1. Trial-to-paid

ProductLed reports a 9% median. ICONIQ reports 50%. Both are correct. ProductLed counts self-serve signups; ICONIQ counts sales-qualified enterprise POCs, which sit *after* an SQL. Never put them in the same table.

2. SQL-to-closed-won

First Page Sage reports 12%; ICONIQ reports 28%. FPS's "SQL" is a marketing-vetted lead; ICONIQ's is a sales-accepted opportunity. Two different objects with one name.

3. Activation rate

Every company defines its own activation event, so cross-company activation benchmarks are, strictly, uncomparable. Use them to set an ambition, never to declare victory.

WRITE THESE DOWN BEFORE YOU SCORE

- Our activation event is… ?
- An MQL is… ?
- An SQL is accepted by… ?
- Churn is counted at… ?
- CAC includes… ?
- Expansion excludes… ?

If two people in your company would fill these in differently, you do not have a measurement problem — you have six. Fix this first; it takes an afternoon and it makes every number in the rest of this document mean something.

How benchmarks are used in this document. Every published benchmark carries its source and year. Where sources disagree I show the disagreement rather than picking a flattering one. Where a widely-repeated number has no verifiable study behind it — the "3:1 LTV:CAC rule", the "3× pipeline coverage rule", several viral time-to-value statistics — it is flagged as a heuristic, not evidence. See the appendix on page 62.

How to score

Each layer has eight diagnostics. Score each 0–3 using the same anchors, so scores are comparable across layers and across people.

0 — Don't know
Nobody has looked. Or we have an answer but no way to check it.

1 — Opinion
Someone has a view. It is not written down and it is not measured.

2 — Documented
Written down and measured — but it does not consistently change what we do.

3 — Operating
Measured, owned by a named person, reviewed on a cadence, and it drives decisions.

LAYER BANDS

SCORE /24	BAND	WHAT IT MEANS
0–11	Broken	This layer is a constraint. It is probably *the* constraint.
12–17	Fragile	It works when someone is watching it. It will not survive scale.
18–21	Solid	Reliable. Improvements here will be incremental.
22–24	Compounding	This layer is an asset. Protect it; do not optimise it.

The bands are deliberately harsh. A first honest pass at 14–16 per layer is normal and healthy for a Series A company. Scores above 20 across the board on a first attempt almost always mean the scoring was generous.

THREE RULES FOR HONEST SCORING

1 — If you have to explain it, it is a 1
"Well, we sort of know that, it's just not in one place" is a 1. Every time. The explanation is the score.

2 — 3 requires a name
If you cannot say which person is accountable for the number and when they last reviewed it, it is a 2 at most — regardless of how good the number is.

3 — Score blind, then compare
Have your CEO, your head of sales and your most senior IC score independently before anyone sees the others' answers. **The disagreements are more diagnostic than the scores.** A layer where independent scorers differ by three or more points is an alignment problem before it is a performance problem.

◆ **Diamonds mark high-signal questions.** On these, the *quality* of the answer tells you more than its content — how quickly it comes, whether it is specific, whether two people give the same one. If you are short on time, answer the diamonds first: there are twenty-four of them, four per layer, and they carry most of the diagnosis.

0
Fit.

The gate. Before acquisition, activation or retention mean anything, four things have to fit each other: your market and your product, your product and your channel, your channel and your price, and your price and your market. Break any one and every number below it becomes noise.

What breaks at Layer 0

Fit failures never announce themselves as fit failures. They arrive disguised as marketing complaints, which is why they survive for years.

YOU HEAR	IT USUALLY MEANS
"The leads are low quality"	Nobody has written down who a good lead is. Marketing is optimising against a target that does not exist.
"We lose to Competitor X"	Check first: a large share of B2B losses are to *no decision*, not to a competitor. That is an urgency failure.
"Sales needs better collateral"	Every rep is telling a different story because there is no agreed frame of reference.
"We just need more awareness"	Awareness of what, by whom, in place of what? If those cannot be answered, spend will not fix it.
"Our market is getting crowded"	You have inherited a category from a competitor and are now fighting on their terms.

THE TELL THAT BEATS ALL OTHERS

What would your customers do if you did not exist?

Not "which competitor would they buy." What would they actually *do* on Monday morning? For most B2B SaaS the honest answer is a spreadsheet, a junior hire, or nothing at all.

If your team answers with a competitor's name and your win/loss data says otherwise, your entire message is aimed at the wrong opponent. You are arguing feature parity with a vendor while your real rival is inertia — and inertia does not lose on features.

April Dunford, *Obviously Awesome* (2019) — competitive alternatives are component one of positioning for exactly this reason.

THREE FIT FAILURES, PRICED

Channel–model mismatch

A £30K ACV product sold through a self-serve funnel, or a £4K ACV product sold by enterprise AEs. Both are structurally unprofitable and neither shows up as a CAC problem until about eighteen months in.

Cost: a full GTM rebuild, and usually a hiring reversal.

Early adopters read as a market

Early adopters tolerate broken products and leave quickly. In one documented case fewer than 20% were still using the product two years on. Positioning built on their feedback does not transfer to the early majority.

Cost: a roadmap year spent on features the market does not want.

Inherited category

You describe yourself using the market frame of your largest competitor or the founder's previous company. Every deal is then fought on someone else's criteria, where your strengths are, by construction, not central.

Cost: permanently depressed win rates that look like a sales problem.

The Four Fits, tested

Brian Balfour's framing (2017): four fits that must hold *simultaneously*, not sequentially. Most companies have two and are trying to fix the wrong two.

FIT 01

Market → Product

Note the word order. You start from a market with an urgent problem and build product into it — not build product and then hunt for a market.

TEST IT

Run the Sean Ellis survey: "How would you feel if you could no longer use this?" The 40% "very disappointed" threshold is the benchmark — **but segment it.** A 45% aggregate can hide a 70% segment and a 15% one, and the 15% is where your churn lives.

FIT 02

Product → Channel

"Products are built to fit with channels. Channels do not mold to products." Each channel demands specific product attributes.

TEST IT

Virality needs fast time-to-value and network effects. Paid needs a targetable audience and a transactional model that funds spend. Content/SEO needs either a genuine expertise moat or user-generated pages at volume. If your product has none of the attributes the channel requires, the channel will not work no matter how well it is executed.

FIT 03

Channel → Model

The CAC a channel produces has to be compatible with your ARPU. This is the fit that kills the most Series A companies.

TEST IT

Divide fully loaded CAC per channel by monthly ARPA × gross margin — omitting margin understates payback by 20–30%. Past 24 months you are not running a marketing programme, you are running a financing operation. Median B2B SaaS payback is 18 months and rising (Benchmarkit, 2025).

FIT 04

Model → Market

The arithmetic constraint. ARPU × reachable customers × realistic capture rate has to clear the number your funding implies.

TEST IT

Do the multiplication honestly, using *reachable* customers rather than a TAM slide. If the answer does not clear your target, no execution fixes it — you move upmarket, widen the product, or change the funding expectation. Those are the only three options.

> **The diagnostic power is in the pairs.** A company with Market–Product and Model–Market but no Product–Channel fit has a real business that cannot be scaled through any channel it has tried. A company with Product–Channel and Channel–Model but weak Market–Product has an efficient machine for acquiring customers who will churn. These are completely different problems, and both present as "growth has stalled."

The half of your messaging that is almost certainly *missing*.

Bob Moesta and Chris Spiek's Forces of Progress: a switch happens only when push and pull together outweigh anxiety and habit. Most SaaS messaging addresses two of the four.

TOWARDS CHANGE

Push of the situation

Dissatisfaction with how things are now. Almost all change starts here.

Where you address it: the problem section of your homepage. Nearly everyone does this.

TOWARDS CHANGE

Pull of the new

Attraction to the alternative — real only once the buyer can *picture* the better state.

Where you address it: benefits, screenshots, case studies. Nearly everyone does this too.

AGAINST CHANGE

Anxiety of the new

"Will this work with our stack? What if we pick wrong? Who owns this when it breaks?"

Where you address it: almost nowhere. This is the hidden deal-killer.

AGAINST CHANGE

Habit of the present

"The spreadsheet works. Retraining twelve people is worse than the problem."

Where you address it: also almost nowhere. This is why deals go quiet.

A 30-MINUTE AUDIT YOU CAN RUN TODAY

Print your homepage, your sales deck and your first three onboarding emails. Highlight every sentence by which force it addresses. I have yet to see a company cover all four. The output is a list of the objections your buyer is having silently, in a meeting you are not in.

ANXIETY COPY THAT WORKS

- Named migration path with a time estimate
- Security and data-residency answers *above* the fold for regulated buyers
- A public exit path — "here is how you get your data out"
- Proof from a company that looks structurally like theirs

HABIT COPY THAT WORKS

- Explicit "you can keep doing X" — reduce what has to change
- Who else has to be retrained, and how long it took them
- The cost of the current way, quantified by them, not by you
- A first step small enough to not feel like a decision

Why this sits in Layer 0, not Acquisition. Unaddressed anxiety and habit do not reduce your conversion rate by a few points. They change *who* converts — you end up selling only to the small minority of buyers with unusually high risk tolerance, which is a different and much smaller market than the one you thought you had.

Layer 0 — diagnostics

Score each 0–3. ◆ marks a high-signal question: how fast and how specifically it is answered matters more than the answer itself.

01 ◆ If five people on your team wrote down your ICP independently, how similar would the answers be — and is it written anywhere? 0 1 2 3

A 3 requires a written ICP with disqualification criteria that someone has used to turn a deal away in the last quarter.

02 What did your last 20 closed-won deals have in common that your last 20 closed-lost deals did not? 0 1 2 3

A 3 means this has been analysed, not remembered.

03 ◆ What would your customers actually do on Monday morning if you did not exist? 0 1 2 3

A 3 means you have asked them, recently, and the answer is not a competitor's name.

04 ◆ What share of your losses are "no decision" versus a named competitor? 0 1 2 3

High no-decision is an urgency and positioning failure, not a competitive one.

05 Who chose your market category — you, an analyst, a competitor, or a founder's last company? 0 1 2 3

A 3 means the frame was chosen deliberately and puts your strengths at the centre.

06 ◆ Where in your funnel do you address the *anxiety* of switching to you, and the *habit* of the current way? 0 1 2 3

Score 0 if you cannot point at a specific asset for each.

07 If you ran the "very disappointed" survey today, what would the number be for your best-fit segment specifically? 0 1 2 3

A 3 requires having run it, segmented it, and filtered to users who reached the core value.

08 Does ARPU × reachable customers × realistic capture clear the number your funding implies? 0 1 2 3

A 3 means this arithmetic exists in a spreadsheet, not in a TAM slide.

LAYER 0 SCORE

/ 24

0–11 Broken · 12–17 Fragile · 18–21 Solid · 22–24 Compounding

THE GATE

If Layer 0 scores below 12, stop here. Score the other layers for completeness if you like, but treat their results as unreliable — you are measuring a funnel pointed at people who were never going to buy. Take the plan on page 18 and run it before you spend another pound on acquisition. This is founder work. It cannot be delegated to an agency, and it is the highest-return quarter available to you.

Layer 0 — evidence

Fit is the hardest layer to benchmark, because the honest metrics are qualitative. These are the numbers that do exist, and the thresholds worth holding yourself to.

SIGNAL	THRESHOLD / BENCHMARK	SOURCE
Product/market fit survey	≥40% "very disappointed"	Sean Ellis, ~100 startups (2009). Slack scored 51%; Superhuman moved 22% → 58% in three quarters.
Channel concentration	>70% from one channel	Balfour (2017) — high-growth companies are channel-concentrated, not diversified.
Pipeline mix, high-growth cos	Sales 62% · Channel 19% · Marketing 15% · CS 4%	ICONIQ, State of GTM 2026
Pipeline mix, everyone else	Sales 45% · Channel 27% · Marketing 20% · CS 9%	ICONIQ, State of GTM 2026 — rounded
Win rate by pipeline source	CS-sourced 52% · Sales 38% · Partner 35% · Marketing 23%	ICONIQ, State of GTM 2026
Sales cycle by ACV	<$10K 6wk · $10–50K 12wk · $50–100K 17wk · $100K+ 24wk	ICONIQ, State of GTM 2026
Odds of reaching scale	~50% reach $1M ARR · ~10% reach $10M · ~2% reach $25M	ChartMogul, n=6,525 (2025)
Early-adopter durability	<20% still using after 2 years	Garbugli, LANDR case (Lean B2B) — single-company, directional

USE THE 40% RULE CAREFULLY

It is **necessary but not sufficient.** Tristan Kromer documents his own startup scoring above 40% while lacking real fit — respondents were scoring the promise of a solution, not the product's performance.

Three specific traps: **survivorship** (surveys reach active users; the churned are structurally excluded, which inflates the score), **segment masking** (the aggregate hides the segment that is killing you), and **the B2B respondent problem** (a 60% user score can coexist with 90% logo churn when the economic buyer sees no value).

WHAT A GOOD ANSWER SOUNDS LIKE

"Ops leads at UK payments companies between 50 and 400 staff, who have already hired an analyst to do this manually and are about to hire a second. They find us when the second hire gets blocked in budget. If they have not made the first hire, we lose to nothing — so we disqualify them."

Note what is in there: a segment, a trigger, an alternative, a disqualification rule. Four things. That is a 3. "Mid-market fintech" is a 1.

If Layer 0 is your constraint

This is founder work. It cannot be outsourced, it takes roughly a quarter, and it is the highest-return quarter available to a stalled B2B SaaS company.

1. **Interview 12 buyers from the last 90 days**

   Not a survey. Reconstruct the actual timeline — first thought, passive looking, active looking, decision, purchase. Ask where they stalled and what unstalled them. Expect most of what you hear to be unusable; the part that matters is the effort they spent overcoming their own resistance. Include three who evaluated and chose not to buy.

2. **Write the disqualification list before the ICP**

   It is easier and far more honest. Who do you refuse to sell to, and why? A team that cannot name five disqualifiers does not have an ICP, it has a preference. Circulate it to sales and hold a deal to it within the fortnight.

3. **Run the competitive-alternatives exercise**

   List what buyers would truly do instead — usually a spreadsheet, an intern, an incumbent suite module, or nothing. Then check your last two quarters of loss reasons against that list. If "no decision" is your largest single loss reason, your messaging is aimed at the wrong opponent.

4. **Choose a market frame that centres your strengths**

   Dunford's process: competitive alternatives → unique attributes → the value they enable → who cares disproportionately → the category that makes that value obvious. Do it with sales and CS in the room, not marketing alone.

5. **Rewrite for all four forces**

   Homepage, primary deck, first three onboarding emails. Add the anxiety and habit material you found missing on page 15. This is usually a two-week job and it typically moves win rate before it moves traffic.

DO NOT, THIS QUARTER

- Increase paid spend
- Hire a demand-gen lead
- Start a content programme
- Rebuild the website
- Run A/B tests

Each of these scales whatever fit you currently have. If fit is the problem, they scale the problem.

EVIDENCE YOU ARE DONE

- Written ICP with five disqualifiers, used to kill a real deal
- Loss reasons categorised for two quarters
- Every rep gives the same "why you, why now"
- Anxiety and habit copy live in three places
- Segmented PMF survey with a named best-fit segment

WHAT USUALLY MOVES FIRST

Win rate	4–8 wks
No-decision losses	1 qtr
Sales cycle	1–2 qtr
Churn	2–4 qtr

Churn moves last because you are still working through the bad-fit cohort you already sold.

1

Acquisition.

Can you reach the right people, repeatably, at a cost your average contract value can actually fund? Most acquisition problems are not volume problems. They are concentration problems, arithmetic problems, or a founder's network being mistaken for a channel.

What breaks at Layer 1

Three failures account for most of it, and none of them is solved by increasing the budget.

1 **Channel spray — six channels at 10% effort each**

Balfour's observation is that high-growth companies get 70%+ of growth from a single channel. Channels obey a power law, not a portfolio logic. Running six at a fraction of the effort needed for any one to work is the most common way a Series A marketing budget is spent producing nothing. The tell: you cannot name the channel that produced most of last quarter's pipeline.

2 **The founder's network mistaken for a repeatable channel**

The Series A pattern. Early pipeline came from the founder's relationships, warm intros and reputation. That is real revenue and it is not a channel — it does not scale with spend and it does not transfer to a hire. The tell: strip founder-sourced deals out of last year's pipeline and see what is left. If it is under half, you have not yet found a channel.

3 **Channel–model arithmetic that has never been done**

CAC has to be compatible with ARPU. A channel that delivers a £4,600 CAC is excellent for a £30K ACV enterprise product and fatal for a £4K one. Most companies know blended CAC and have never costed a single channel separately, which means they cannot tell a good channel from a subsidised one.

THE CONCENTRATION TEST

What share of new revenue came from your single largest channel last quarter?

Over 70%	`Channel fit`
40–70%	`Emerging`
Under 40%	`Spray`

Diversification is what you do *after* a channel works, to reduce risk. Doing it before is not risk management, it is an inability to choose.

THE FOUNDER-STRIP TEST

Take last year's closed-won list. Mark every deal where the founder made the first contact, was personally known to the buyer, or closed it themselves. Remove them.

What is left is your actual acquisition capability. Most Series A companies are shocked by this number, and it is an hour of analysis that changes the plan more often than anything else on this page. **If the founder went dark for 90 days, what happens to pipeline?** That is the same question with the answer made vivid.

Engine, turbo boost, lubricant, or *fuel?*

Lenny Rachitsky and Dan Hockenmaier's Racecar framework, originated at Reforge. Classify every growth activity you run into one of four buckets. The classification is the diagnosis.

COMPOUNDS

Engine

The self-sustaining loop. Four archetypes: content/SEO, paid, sales, virality. Output feeds back in as input.

Test: does last month's output make next month's input cheaper or larger?

ONE-OFF

Turbo boost

PR, launches, partnerships, a platform moment, an award. Real effects. They do not compound.

Test: if you stopped, would the effect decay to zero within two quarters?

EFFICIENCY

Lubricant

CRO, friction removal, better copy, faster forms. Makes the engine more efficient. Is not an engine.

Test: does it create demand, or convert demand you already had?

CONSUMED

Fuel

Capital, and the activity the engine burns. Paid spend with negative unit economics is fuel wearing an engine costume.

Test: does the activity survive a funding freeze?

THE TWO PATHOLOGIES

No engine, all lubricants

Enormous activity, no compounding. Every quarter starts from zero. The team is busy, the dashboards are full, and the growth rate is flat because nothing that happened last quarter makes this quarter easier. This is the most common state of a Series A marketing function.

Paid mistaken for an engine

Paid *is* a legitimate engine — but only when the revenue it generates funds the next round of spend at an acceptable payback. If payback runs beyond 24 months, you are not running an engine, you are converting investor capital into customers at a loss and calling it growth.

A USEFUL EXERCISE FOR YOUR NEXT LEADERSHIP MEETING

List every growth activity currently consuming someone's time. Put each in one bucket. Then total the **share of team hours going to things that compound**. In most companies scoring below 15 on this layer, the answer is under 20% — and nobody had noticed, because each individual activity was defensible.

Layer 1 — diagnostics

Score each 0–3. ◆ marks a high-signal question.

01 ◆ What share of new revenue came from your single largest channel last quarter? 0 1 2 3

Under 40% suggests no channel fit yet. A 3 requires the number to be known without looking it up.

02 What is your CAC *by channel*, fully loaded — and which channels have you never costed separately? 0 1 2 3

A 3 means salaries, tools and agency fees are allocated, not just media spend.

03 ◆ How much of last year's pipeline was founder-sourced, and what remains without it? 0 1 2 3

A 3 means the analysis exists and the non-founder number is growing.

04 Does your product have the attributes the channel you are betting on actually requires? 0 1 2 3

Time to value, breadth of value proposition, and a model that funds the spend.

05 ◆ Does your ACV support the acquisition motion you are running? 0 1 2 3

A £4K ACV product with enterprise AEs, or a £30K ACV product through a self-serve funnel, both score 0.

06 Which channels have you actually killed in the last year, and on what evidence? 0 1 2 3

A team that has never killed a channel is not running channels, it is accumulating them.

07 ◆ When a channel starts working, what mechanism reinvests its output back into its input? 0 1 2 3

If nothing does, it is not an engine — growth stops the moment spend stops.

08 What is your CAC payback in months, and how has it moved over eight quarters? 0 1 2 3

A 3 requires the trend, not the point. Median across B2B SaaS is 18 months and rising.

LAYER 1 SCORE

— / 24

0–11 Broken · 12–17 Fragile · 18–21 Solid · 22–24 Compounding

A WARNING ABOUT THIS LAYER

Layer 1 is where founders most often *want* the constraint to be, because it is the layer with the most obvious tactics and the easiest budget conversation. Be suspicious of a result that puts your constraint here when Layer 0 scored close. If Layer 0 and Layer 1 are within three points of each other, work Layer 0 first — the acquisition numbers you are measuring were produced by whatever fit you currently have.

Layer 1 — evidence

Use these to locate yourself, not to set a target. Note the sample each comes from — a benchmark drawn from VC-backed high-growth companies is not a benchmark for a bootstrapped one.

EFFICIENCY

METRIC	MEDIAN	SOURCE
CAC payback	18 mo	Benchmarkit 2025 (n=148); up 12.5% since 2022
New-customer CAC ratio	$2.00	Benchmarkit 2025 — per $1 new ARR; fourth quartile $2.82
Blended CAC ratio	$1.81	Benchmarkit 2025 (n=43)
Expansion CAC ratio	$1.00	Benchmarkit 2025 — up from $0.69 in 2022
S&M as % revenue	37%	Benchmarkit 2025 — VC-backed 45–47%, PE-backed 33%
Marketing as % ARR	8%	SaaS Capital 2026 (n=1,000+) — mostly bootstrapped
Sales as % ARR	15%	SaaS Capital 2026

The 37% vs 23% gap between Benchmarkit and SaaS Capital is population, not performance: VC-backed high-growth versus a largely bootstrapped private panel. Pick the one you resemble.

FUNNEL & CHANNEL

METRIC	VALUE	SOURCE
Visitor → lead	1.4% median	First Page Sage 2025 — top quartile 2.1–2.3%
Visitor → lead by audience	Small 2.3% · Ent 0.7%	First Page Sage 2025
Lead → MQL, best sources	Referrals 56% · Exec events 54% · SEO 41%	First Page Sage 2025
MQL → SQL	27%	ICONIQ 2026 — sales-accepted definition
SQL → closed won	28%	ICONIQ 2026
Pipeline coverage	3.6×	ICONIQ 2025 — PLG 3.2×, sales-led 3.5×, hybrid 3.7×
Ramped AEs at quota	62%	ICONIQ 2026 — declining trend since 2023

CAC BY CHANNEL — DIRECTIONAL

Email $510 · Webinars $603 · Thought-leadership SEO $647 · Paid search $802 · LinkedIn Ads $982 · Content $1,254 · Trade shows $1,390 · Outbound SDR $1,980 · ABM $4,664.

First Page Sage, ~120 B2B firms, data to Nov 2024, reported in USD. The publisher is an SEO agency, so organic channels are almost certainly flattered. Use the *ordering*, not the values.

TWO RULES WITH NO DATA BEHIND THEM

"LTV:CAC should be 3:1." A 2000s venture heuristic. No 2024–2026 report publishes a median LTV:CAC with a disclosed sample. Useful as a sanity check, worthless as a benchmark.

"You need 3× pipeline coverage." Also undocumented. ICONIQ's measured 3.2–3.7× by motion is the citable substitute — and note their counterintuitive finding that teams with higher AI adoption need *less* coverage, not more.

If Layer 1 is your constraint

The instinct is to add. The work is almost always to subtract, then concentrate what is left onto one bet.

1. **Cost every channel properly, once**

   Fully loaded: media, tools, agency fees, and the salary share of everyone who touches it. Most teams discover that their "cheapest" channel is expensive once a person's time is counted, and that a channel they were about to cut is their best. This is a two-day job and it changes the decision more often than not.

2. **Do the channel–model arithmetic**

   CAC by channel divided by monthly ARPA gives payback by channel. Anything past 24 months is on notice. Anything past 36 is being subsidised by your balance sheet and should stop this quarter, regardless of how good the volume looks.

3. **Kill everything except two**

   One channel that is closest to working, and one deliberate experiment. Nothing else. This will feel reckless and it is the entire intervention — a channel needs concentrated effort to reach the point where you can judge it, and you have been denying every channel that concentration.

4. **Give the chosen channel a real test window**

   On First Page Sage's data — an SEO agency's own campaigns, so read it as a best case — content and SEO break even at roughly 7–10 months, and the first 4–6 months of any channel run materially above its steady-state CAC. If you cannot commit two quarters, do not start; choose a channel whose feedback loop matches your patience.

5. **Build the reinvestment mechanism explicitly**

   Write down the sentence: "When this channel produces X, that output becomes input for the next cycle by doing Y." If you cannot complete the sentence, you have a campaign, not an engine. Campaigns are fine — but budget them as one-offs, not as growth.

IF PIPELINE IS FOUNDER-DEPENDENT

Do not hire an AE to fix it. Document why the last ten deals were won, in the founder's words, before anyone else is expected to reproduce them. Undocumented founder-led sales is, in my experience, the most common reason a first sales hire fails inside a year.

IF CAC IS RISING

Check the model before the bidding. Rising CAC in a channel your ACV cannot fund is not an optimisation problem — it is the arithmetic arriving. The fixes are: move upmarket, raise price, or change channel. Better creative is not on the list.

EVIDENCE YOU ARE DONE

- One channel above 60% of new pipeline, and rising
- Payback by channel, tracked monthly
- A written reinvestment mechanism
- Non-founder pipeline growing quarter on quarter

2

Activation.

The gap between someone signing up and someone experiencing the thing you actually sell. It is the cheapest layer to improve, the least often owned by a named person, and the one where a two-point gain compounds through every layer beneath it.

What breaks at Layer 2

Activation fails quietly. Nobody complains, nobody churns loudly, and the metric that would reveal it usually does not exist.

FAILURE ONE

No defined activation event

If you cannot name the specific action that means "this person has now experienced the value," you cannot measure activation, cannot improve it, and cannot tell a healthy cohort from a doomed one until renewal.

This is the most common state, and it is why cross-company activation benchmarks are close to meaningless.

FAILURE TWO

Single-user value in a team product

The product genuinely requires three people to be useful, and onboarding is designed for one. The first user reaches a dead end that looks like disinterest and is actually a structural mismatch.

Elena Verna's escalator: individual → team → enterprise. Each rung needs motivation, ability and — her addition to Fogg's model — organisational permission.

FAILURE THREE

Nobody owns the handover

Sales owns up to signature. Customer success owns from month two. The three weeks in between — where activation actually happens — belong to nobody, and it shows in the cohort data.

Winning by Design treats onboarding retention as a first-class conversion metric on the right side of the Bowtie. Almost nobody measures it.

THE NUMBER THAT JUSTIFIES THE WORK

Users who reach a real activation milestone retain at least twice as well as those who do not.

An association, not a proven cause — see the causal test on page 27. It is still the strongest signal available at this stage.

Lenny's Newsletter, 500+ products (2022). Amplitude's 2,600-company benchmark found 69% of products with strong early activation also showed strong three-month retention.

This is why activation sits above retention in the stack and not beside it. A retention problem that originates in activation cannot be fixed by customer success, by better QBRs, or by a renewal playbook. Those act on people who never reached the value in the first place — which is why they feel so hard.

If your Layer 3 score is low and your Layer 2 score is lower, **your churn problem is an activation problem wearing a retention costume**, and the intervention belongs eleven months earlier in the customer's life than you think.

Defining the activation event

Most companies pick a milestone that is easy to instrument rather than one that predicts anything. There is a method for doing it properly, and it takes a day.

1. **Split your cohorts by outcome, not by behaviour**
   Take customers from 12–18 months ago. Split into those still active and those churned. You are looking for what the survivors did that the leavers did not — not what everyone does.

2. **Look at week one only**
   Compare the two groups on actions taken in the first seven days. The candidate events are the ones with the widest gap between groups, not the ones with the highest overall usage. Logging in is high-usage and predicts nothing.

3. **Test whether it is causal or correlated**
   Engaged users do everything more. The question is whether pushing an ambivalent user to the milestone changes their trajectory. If you cannot test it, at least check that the milestone requires deliberate effort rather than being a by-product of interest.

4. **For team products, require more than one person**
   Verna's point: your activation metric should embed multiple roles and company-wide distribution, not a single user's success. "Three users from two departments completed X within 14 days" is a real B2B activation event. "Completed onboarding" is not.

5. **Write it down and give it an owner**
   One sentence, with numbers and a time bound, on a wall. Then name the person whose job it is. An activation metric without an owner reverts to a dashboard nobody opens within a quarter.

WEAK ACTIVATION EVENTS

- Completed signup
- Verified email
- Logged in twice
- Finished the product tour
- Invited a teammate

These measure compliance with your onboarding, not the arrival of value. They all go up when you add a nag screen and predict nothing.

STRONG ACTIVATION EVENTS

- Connected a production data source
- Ran the workflow on real data, twice, in week one
- Three users across two teams took the core action within 14 days
- Received the first output they then shared internally
- Replaced the previous tool for one real task

Each requires effort, uses real data, and implies the product has entered the customer's actual workflow.

> **The onboarding-completion trap.** Median onboarding checklist completion is around 10% — the average is 19%, and the gap between them tells you the distribution is heavily skewed (Userpilot, n=188, 2024). Teams respond by optimising checklist completion. That is optimising the proxy. Completion is only worth anything if completers retain better than non-completers, which is a question almost nobody checks and which frequently comes back negative.

Layer 2 — diagnostics

Score each 0–3. ◆ marks a high-signal question.

01 ◆ What specific event means a user has experienced value, and what share of signups reach it? 0 1 2 3

A 3 requires the event to be written down, instrumented, and validated against retention.

02 ◆ How long is it from contract signature to first measurable impact — and who is accountable for that number? 0 1 2 3

The Bowtie makes time-to-first-impact a first-class metric. If nobody owns it, score 1 at most.

03 Does a single user get value alone, or does your product need a team before it is useful at all? 0 1 2 3

A 3 means onboarding is designed for whichever of the two is true, deliberately.

04 For the next rung of expansion, does the user have motivation, ability *and* permission? 0 1 2 3

Most B2B product-led stalls are permission failures — procurement, IT or security, not interest.

05 ◆ Where in onboarding do users most commonly stop, and have you watched a recording of it happening? 0 1 2 3

A 3 requires someone to have actually watched. Session data alone is a 2.

06 What does your best-performing cohort do in week one that your worst does not? 0 1 2 3

A 3 means this comparison has been run on real cohorts, split by 12-month outcome.

07 Does your activation metric include multiple roles and company-wide distribution? 0 1 2 3

For a team product, a single-user activation metric will systematically overstate health.

08 ◆ Who owns the period between signature and month two, by name? 0 1 2 3

"Sales hands to CS" without a named owner of the gap scores 0.

LAYER 2 SCORE

— / 24

0–11 Broken · 12–17 Fragile · 18–21 Solid · 22–24 Compounding

WHY THIS LAYER IS USUALLY UNDERSCORED

Activation sits in the organisational gap between product, sales and customer success, so it is nobody's quarterly objective. Teams therefore score it from impression rather than data, and the impression is generous because the failures are invisible — a user who never reached value simply stops appearing, and produces no complaint, no ticket and no exit interview. **If you score above 15 here without opening an analytics tool, score it again.**

Layer 2 — evidence

Read the definitional note before using any of this. Activation and trial-conversion benchmarks are the least comparable numbers in SaaS.

FREE-TO-PAID, SELF-SERVE

MODEL	GOOD	GREAT
Freemium, self-serve	3–5%	6–8%
Freemium + sales assist	5–7%	10–15%
Free trial	8–12%	15–25%
Reverse trial	4–6%	8–12%

Lenny's Newsletter / Kyle Poyar, 1,000+ products (2023); reverse-trial band from ChartMogul's conversion report (2026, n=200). Developer-focused products run at roughly half the rate of non-developer products.

TRIAL DESIGN

DESIGN	SIGNUP RATE	TRIAL → PAID
Opt-in, no card	8.5%	18.2%
Opt-out, card required	2.5%	48.8%
Freemium	13.3%	2.6%

First Page Sage, 86 SaaS companies, Q1 2022–Q3 2025. End to end, opt-in produces ~1.55% visitor-to-paid against opt-out's ~1.22% — opt-in wins on volume, opt-out on efficiency. ChartMogul's 2026 panel puts card-required conversion nearer 30% rather than 49%; both find the same direction, and the gap is sample composition.

ACTIVATION & ONBOARDING

METRIC	VALUE	SOURCE
Activation rate, SaaS median	30%	Lenny's, 500+ products (2022); average 36%
Activation rate, B2B median	37%	Userpilot, n=62 (2024)
Activation, PLG vs sales-led	34.6% vs 41.6%	Userpilot, n=547 (2024) — runs against the PLG narrative
Onboarding checklist completion	10.1% median	Userpilot, n=188 (2024); average 19.2%
Time to value	~1.5 days	Userpilot, n=547 (2024) — the only TTV figure with a disclosed sample
Retention lift from activation	≥2×	Lenny's (2022)
Trial → paid with PQL scoring	~25%	ProductLed, 600+ cos (2025) vs 9% baseline
Companies running a PQL motion	~24%	ProductLed (2025)
Enterprise trial/POC → paid	50%	ICONIQ 2026 — a late-funnel stage, not a signup
Demo → closed won	36%	ICONIQ 2026

THE ONE INTERVENTION WITH REAL EVIDENCE

OpenView's 2023 product benchmark (1,000+ respondents) found free-trial conversion **doubles**, and freemium conversion **quadruples**, when sales reaches out to more than half of signups. Not a nurture sequence. A person.

The PQL correlation is weaker evidence than it looks: companies mature enough to run PQL scoring are probably better at conversion generally. Treat 3× as an upper bound.

NUMBERS CIRCULATING THAT YOU SHOULD NOT CITE

A cluster of viral time-to-value statistics — "every 10-minute delay costs 8% of trial conversion," "feature overwhelm cuts conversion 45%," a claimed 10,000-company study — has no locatable primary source, methodology or research footprint.

Similarly, a widely-shared trial-conversion percentile table attributed to ChartMogul does not appear in any ChartMogul publication. Time to value is the weakest-evidenced metric in this entire document.

If Layer 2 is your constraint

The good news about activation: it is the cheapest layer to move, it moves fastest, and the gains compound into every layer beneath it.

1. **Define and instrument the activation event — week one**
   Use the cohort method on page 27. Do not skip to interventions. Every hour spent on onboarding before the event is defined is spent optimising something you cannot measure.

2. **Watch ten session recordings, end to end**
   Not a summary. Ten actual recordings of new users, watched by the founder and whoever owns product. Almost nobody does it, and almost everyone who does finds one obvious, embarrassing blocker in the first three.

3. **Name an owner for signature-to-month-two**
   One person, whose objective it is. This can be a CS lead, a product manager or a founder. It cannot be "the team." The gap between sales and CS is where activation problems live precisely because it belongs to nobody.

4. **Add human contact to the first week**
   The best-evidenced intervention available. Reach out to more than half your signups with a person, not a sequence. For sales-led products, this means the AE stays involved through first value rather than handing over at signature.

5. **Remove one step, do not add a tooltip**
   The reflex is to explain the product better. The better move is usually to require less before value: pre-populate, provide sample data, defer the configuration, shorten the form. Adding guidance treats the symptom; removing steps treats the cause.

IF YOU SELL TO TEAMS

Redefine activation to require multiple people and cross-team spread. Then design the first week around getting the second and third user in — not around making the first user proficient. Most B2B onboarding does the opposite.

IF YOU SELL TO ENTERPRISE

Your constraint is usually *permission*, not motivation. Security review, data residency, procurement and IT approval are the real activation blockers. Instrument those as funnel stages with owners and cycle times, because that is what they are.

EVIDENCE YOU ARE DONE

- An activation event with a number and a time bound
- Validated against 12-month retention
- A named owner of signature-to-month-two
- Activation rate reviewed monthly, by cohort

3

Retention.

Does the value hold once the novelty, the onboarding attention and the launch enthusiasm are gone? Retention is the layer that decides whether everything above it was an investment or an expense — and it is the layer most often hidden behind a single flattering number.

What breaks at Layer 3

Retention problems are rarely retention problems. They are usually Layer 0 or Layer 2 problems arriving on a delay of eleven months.

THE PRESENTING SYMPTOM	WHAT IT USUALLY IS	WHERE THE FIX LIVES
Churn spikes at month 11–12	Annual renewal exposing customers who never activated	Layer 2
A specific segment churns at 3× the rest	You sold to people outside the ICP because the quarter needed it	Layer 0
Champion leaves, account leaves	No multi-threading and no re-selling motion	Layer 3
Usage decays quietly from month four	The promised value never became measured impact	Layer 3
Payments fail and nobody notices	Involuntary churn, which is roughly a third of all SaaS churn	Layer 3, and it is cheap
NDR looks fine, cash does not	Expansion is masking gross churn	Layer 3, measurement

THE SINGLE MOST USEFUL RETENTION QUESTION

Of the customers who churned last year, how many were bad-fit at the point of sale?

Then the follow-up that matters more: **who decided to sell to them anyway, and what incentive drove that decision?**

In most companies the honest answer is a quarter-end commission structure that pays the same for a customer who renews and one who does not. That is not a customer success problem. It is a compensation design problem, and it will keep manufacturing churn indefinitely.

THE CHEAPEST WIN IN THIS DOCUMENT

Involuntary churn

Roughly **20–40% of subscription churn is involuntary** — failed cards, expired cards, bank declines. In Recurly's SaaS network data it is 1.06 points of a 3.22-point churn rate: about a third.

Baseline retry logic recovers around 53% of failed payments. Optimised, network-informed retries recover around 71%. And **90% of all recoveries happen within ten days** of the failure — so the entire opportunity is in a very short window.

Recurly Research (2026). If you sell month-to-month or self-serve and have never looked at dunning, this is a week of work for a measurable point of retention.

The number that hides behind your *good number.*

If someone quotes you a single retention figure and it is NDR, expansion is covering for churn. Gross revenue retention is the honest number.

GRR

Gross revenue retention

Starting revenue minus churn and contraction. **Expansion excluded.** It cannot exceed 100%.

This is the question "do customers stay?" It cannot be gamed by a large upsell into one account, and it is the number an experienced investor asks for first.

NDR / NRR

Net revenue retention

The same, plus expansion. Can exceed 100%, and often does for reasons that have nothing to do with satisfaction.

A 110% NDR built on 85% GRR plus aggressive seat expansion in three accounts is a fundamentally different business from 110% built on 96% GRR — and looks identical on a slide.

THE GAP

What the spread tells you

NDR minus GRR is your expansion contribution. Track both, always together.

A widening gap with flat GRR means you are getting better at extracting more from the customers who stay while doing nothing about the ones who leave. That works until the base stops growing, and then it stops working suddenly.

CONTRACT LENGTH — WHERE THE REAL EFFECT IS

Month-to-month	NRR 100.0% · GRR 89.5%
Annual	NRR 101.0% · GRR 90.0%
Multi-year	NRR 105.5% · GRR 95.0%

SaaS Capital, 1,500+ private B2B companies (2023).

The widely repeated claim that annual contracts dramatically reduce churn is **not supported by this data**. The annual-versus-monthly GRR gap is half a point. The real step change is at **multi-year**: +5.5 points of gross retention.

Multi-year agreements have gone from 14% of SaaS contracts in 2022 to 40% in 2025 (Maxio, 2025). If you are pushing customers from monthly to annual expecting a retention transformation, you are doing the right thing for cash flow and the wrong thing for the reason you think.

On the "smile curve." In ChartMogul's primary cohort data the best B2B cohorts — above roughly $100/month ARPA — do not dip: they hold at or above 100% of month-zero revenue at both month three and month twelve, the two points the data reports. The shape is flat-to-rising, not a smile. Treat it as a qualitative story, not a benchmarked shape you should expect in your own data.

Layer 3 — diagnostics

Score each 0–3. ◆ marks a high-signal question.

01 ◆ What is your GRR, separately from your NDR? 0 1 2 3

If someone can only quote NDR, expansion is masking churn. Score 1 at most.

02 Does your retention curve flatten, and at what level — for which segment? 0 1 2 3

A 3 requires cohort curves by segment, not a blended monthly churn number.

03 ◆ Of the customers who churned last year, how many were bad-fit at the point of sale? 0 1 2 3

And what incentive caused you to sell to them anyway?

04 What is the gap between logo churn and revenue churn, and in which direction? 0 1 2 3

Revenue churn above logo churn means you are losing your larger customers. That is the urgent case.

05 ◆ What share of your churn is involuntary, and what is your dunning recovery rate? 0 1 2 3

Around a third of SaaS churn is failed payments. A 3 requires knowing your own number.

06 When a champion leaves an account, what happens? 0 1 2 3

A 3 requires a named re-selling motion and multi-threading as standard, not as heroics.

07 Can you show that the value promised at sale became measured impact for your top ten accounts? 0 1 2 3

The Bowtie distinction between Value and Impact. Renewal depends on the second.

08 ◆ If you stopped all acquisition spend tomorrow, would revenue grow, flatten or shrink? 0 1 2 3

The cleanest single test of whether the base is an asset or a leak.

LAYER 3 SCORE

— / 24

0–11 Broken · 12–17 Fragile · 18–21 Solid · 22–24 Compounding

BEFORE YOU ACT ON A LOW SCORE HERE

Check Layer 0 and Layer 2 first. A retention score below 12 alongside a fit or activation score below 12 means **the churn you are seeing was created eleven months ago and cannot be fixed downstream.** Customer success can slow the bleeding on the cohort you already have; only the upstream fix stops manufacturing the next one. Work upstream, and accept that your retention number will keep deteriorating for two to four quarters while the bad cohort works through.

Layer 3 — evidence

The most sample-sensitive numbers in SaaS. Find the row that matches your ACV before comparing yourself to anything.

RETENTION BY ACV — THE DEFENSIBLE SPINE

ANNUAL CONTRACT VALUE	NRR	GRR
Under $12K	93%	90%
$12–25K	96%	90%
$25–50K	97%	93%
$50–100K	98%	93%
$100–250K	101%	93%
Over $250K	105%	93%

SaaS Capital, 1,500+ private B2B companies (2023). Use this table rather than any single headline number: NRR rises monotonically with ACV, while GRR plateaus around 93% above $25K. Median GRR in this panel is 91%; Benchmarkit's 2025 median, on a different and more VC-weighted panel, is 88%. Both are right for their samples.

CHURN

METRIC	VALUE	SOURCE
SaaS churn, total	3.22%	Recurly network data (2026), monthly
Voluntary / involuntary	2.16% / 1.06%	Recurly (2026)
Involuntary as share of churn	20–40%	Recurly (2026)
Dunning recovery, base → optimised	53% → 71%	Recurly (2026)

WHY PUBLISHED NRR MEDIANS DISAGREE

SOURCE	MEDIAN NRR	SAMPLE
ChartMogul 2025	82%	~2,700 B2B, billing data, long tail of low-ACV
Benchmarkit 2025	101%	~1,000 B2B, self-reported, VC/PE-skewed
SaaS Capital 2023	102%	1,000+ private B2B, lender panel
ICONIQ 2025	110–120%	~100 growth-stage enterprise, survivorship-selected

A 38-point spread that is almost entirely sample composition. Any "SaaS NRR benchmark" quoted without an ACV band and a sample description is not usable.

RETENTION DRIVES GROWTH

FINDING	EFFECT ON GROWTH	SOURCE
High vs low NRR	43.6% vs 13.1% a year	ChartMogul 2023
GRR above 85%	1.5–2.5× faster	ChartMogul 2023
NRR 90–100% → 100–110%	~+5 points	SaaS Capital 2025
Usage-based pricing	~3 pts lower GRR	Benchmarkit 2024

CUSTOMER SUCCESS COVERAGE

The famous "$2–5M ARR per CSM" benchmark traces to a **2019** Gainsight survey and has never been refreshed. Treat any CS ratio benchmark as directional.

If Layer 3 is your constraint

Sequence matters here more than anywhere. Do the cheap mechanical work first, because it buys you the quarter you need for the structural work.

1. **Split voluntary from involuntary churn — week one**

   If a third of your churn is failed payments, a fortnight on dunning, card-updater services and retry logic recovers more retention than a year of QBRs. Target the ten-day window: 90% of all recoveries happen there. This is the only item on this list that is purely mechanical, and it is why it goes first.

2. **Separate GRR from NDR in every report, permanently**

   Two numbers, always together, in the board pack and the weekly. This single reporting change surfaces problems that a blended NDR has been hiding, often for years.

3. **Categorise last year's churn by root cause**

   Four buckets: bad fit at sale, never activated, value delivered but champion left, value genuinely declined. The distribution tells you which layer owns the fix. In the engagements I have run, the first two buckets usually dominate — which means retention is not the constraint, it is the symptom.

4. **Instrument impact, not satisfaction**

   For your top ten accounts, can you produce evidence that the value promised at sale became a measured result? Winning by Design's distinction between Value and Impact is the whole game: renewal depends on impact, and impact has to be visible to the buyer, not just to you.

5. **Fix the compensation before the process**

   If new-business commission pays identically for a customer who renews and one who churns in month seven, no process will hold. Clawbacks, or a renewal-linked component, change behaviour faster than any enablement programme.

MULTI-THREAD, DELIBERATELY

"Champion leaves, account leaves" is not bad luck, it is single-threading. Make a named second and third contact a stage requirement in your CRM, and treat a single-threaded account as a renewal risk regardless of how happy the champion is.

CONSIDER MULTI-YEAR, NOT ANNUAL

The retention step-change is at multi-year (+5.5 pts GRR), not at annual (+0.5 pts). If you are going to spend negotiating leverage on contract length, spend it on the one that actually moves the number.

EVIDENCE YOU ARE DONE

- GRR and NDR reported separately, monthly
- Churn categorised by root cause
- Involuntary churn measured and actively worked
- Multi-threading enforced in the CRM
- Impact evidence for the top ten accounts

4

Monetisation.

Are you charging for the thing that actually creates the value, and is anyone accountable for that decision? Pricing is the most neglected growth lever in B2B SaaS — not because founders think it is unimportant, but because it belongs to nobody and changing it feels dangerous.

What breaks at Layer 4

Pricing failures are quiet, cumulative and almost entirely reversible — which is what makes them the most under-worked lever in most SaaS companies.

THE LEVER STUDY

3.32% per 1% acquisition

6.71% per 1% retention

12.7% per 1% monetisation

A 1% improvement in monetisation produced roughly **four times** the bottom-line impact of a 1% improvement in acquisition, and about twice that of retention.

Read the caveats. Price Intelligently, N=512 SaaS companies. The study is roughly a decade old, its full methodology was never published, and it was produced by a pricing consultancy with a commercial interest in the conclusion. It is still the most-cited monetisation statistic in SaaS and the direction is almost certainly right. Cite it with the vintage attached, and do not treat the precise multiples as current.

FAILURE ONE

No value metric

You charge per seat for a product whose value scales with transactions, or per workflow for one whose value scales with the number of people who see the output. The price and the value drift apart, and every renewal becomes an argument.

FAILURE TWO

Discounting as policy

A wide discount distribution means the list price is not believed — by your buyers or by your own reps. The variance is the signal, not the average.

FAILURE THREE

Nobody owns pricing

It was set by the founder before the first ten customers and has not been revisited since, because it sits between product, sales and finance and touching it feels like a risk with no obvious owner.

The frequency finding. Across 500 leading SaaS and AI companies, PricingSaaS tracked more than 1,800 pricing and packaging changes in 2025 — roughly **3.6 changes per company per year**. Of those, 43% were feature and packaging changes, 40% were price or structure, and 17% were usage limits. If your pricing has not changed in two years, you are not being cautious relative to your market; you are standing still while it moves.

Finding your value metric

The unit you charge by should be the unit that grows as the customer gets more value. Get this right and expansion becomes automatic; get it wrong and every upsell is a negotiation.

A VALUE METRIC HAS THREE PROPERTIES

1. **It tracks value received**
   As the number goes up, the customer is demonstrably getting more from you.
2. **The customer can predict it**
   If they cannot forecast their bill, procurement will cap it and your expansion ceiling is set by their risk tolerance rather than by their success.
3. **It is hard to game**
   If a customer can get the same value while keeping the metric flat, they will, and your revenue decouples from your usage.

THE TEST

Take your ten happiest customers. Plot what they pay against the value they would say they receive. If the line is flat or noisy, your value metric is wrong — you are systematically underpricing your best customers and overpricing your worst.

THEN ASK THE HARDER QUESTION

Which of your customers would happily pay 30% more, and how do you know? If the answer is a feeling rather than evidence — a willingness-to-pay study, a win/loss pattern, an unusually fast close — then you have a pricing hypothesis, not a pricing strategy.

PRICING MODEL SHIFTS WORTH KNOWING ABOUT

FINDING	NUMBER	SOURCE
Hybrid subscription-plus-usage models post the highest median growth rate	21%	Maxio / Benchmarkit Pricing Trends 2025
SaaS companies now charging for AI-powered features	44%	Maxio 2025
Private SaaS companies already monetising AI	67%	KeyBanc / KBCM 2025
AI credit-based pricing among 500 tracked companies	35 → 79 in one year	PricingSaaS 2025, +126% YoY
Usage-based pricing vs traditional, gross retention	~3 pts lower GRR	Benchmarkit 2024
Multi-year contracts as a share of SaaS agreements	14% (2022) → 40% (2025)	Maxio 2025

The usage-based caveat. Usage-based models are frequently sold as strictly better. The primary data is more nuanced: Benchmarkit finds usage-based pricing associated with roughly three points *lower* gross retention. The plausible reading is that usage models churn harder at the bottom while expanding harder at the top — excellent for NDR, worse for GRR, and a genuinely different risk profile rather than a free upgrade. Claims that usage-based pricing delivers "18–23% higher NRR" circulate widely and trace to a single unverifiable study of ~100 companies.

Layer 4 — diagnostics

Score each 0–3. ◆ marks a high-signal question.

01 ◆ **What is your value metric, and does it actually correlate with the value customers receive?** 0 1 2 3

A 3 requires having plotted it, not asserted it.

02 ◆ **What is your average discount — and what is the *distribution*?** 0 1 2 3

A wide spread means the list price is not believed. The variance matters more than the mean.

03 **Who owns pricing, by name, and when did they last change something?** 0 1 2 3

The market median is roughly 3.6 pricing or packaging changes a year. Two years of stasis scores 0.

04 **Does your pricing match your GTM motion?** 0 1 2 3

Opaque 'contact us' pricing on a self-serve funnel, or published pricing on a six-month enterprise cycle, both score low.

05 ◆ **Which customers would pay 30% more, and what evidence do you have?** 0 1 2 3

A 3 requires research — willingness-to-pay work, win/loss patterns, or a tested price point.

06 **When did you last run a willingness-to-pay study or a structured price test?** 0 1 2 3

'We looked at competitors' pricing pages' is a 1.

07 **Do your packaging tiers map to distinct segments, or did they emerge by accretion?** 0 1 2 3

A 3 means each tier exists because a named segment needs it, and you can say which.

08 ◆ **Can a customer predict their bill next quarter within 10%?** 0 1 2 3

If not, procurement will cap your expansion regardless of how well the product performs.

LAYER 4 SCORE

— / 24

0–11 Broken · 12–17 Fragile · 18–21 Solid · 22–24 Compounding

THE MOST COMMONLY LOWEST-SCORING LAYER

In practice this layer scores lower than any other, and companies still do not work on it — because pricing changes feel irreversible in a way that a campaign does not. They are not. **Grandfathering exists.** You can change price for new customers only, test a new package alongside the old one, or run a structured price increase on a single segment. The perceived risk is much larger than the real one, and that gap is where the opportunity sits.

Layer 4 — evidence, and what is *not* evidence

Pricing is the worst-benchmarked area in SaaS. Being explicit about where the data runs out is more useful than filling the gap with numbers that look authoritative.

REASONABLY WELL EVIDENCED

METRIC	VALUE	SOURCE
Median free-to-paid, self-serve	**8% in six months**	ChartMogul n=200 B2B (2026)
Spread, top vs bottom quintile	**~10×**	ChartMogul (2026)
Card-required trials convert at	**~30%**	ChartMogul (2026); >5× no-card
Model mix	**Trial 57% · Freemium 26% · Reverse trial 7%**	ChartMogul (2026): rest are demo / paid trial
Trial length	**14 days 62% · 7 days 14% · 30 days 14%**	ChartMogul (2026): remainder other lengths
Pricing changes per company per year	**~3.6**	
Change mix	**Packaging 43% · Price 40% · Limits 17%**	
Hybrid pricing, median growth	**21%**	Maxio / Benchmarkit (2025)

WHERE THE EVIDENCE RUNS OUT

No credible primary source publishes any of the following for B2B SaaS. If you see a confident number for one of them, it has been invented or extrapolated:

- Average discount rate from a buyer-side dataset
- The causal effect of discounting on churn
- Adoption rate of willingness-to-pay research
- Decomposition of expansion into upsell, cross-sell and price increase
- Upsell conversion rates by account
- Referral or word-of-mouth as a share of new ARR

This matters practically: it means you cannot benchmark your discounting against the market, so the useful comparison is **against yourself** — your own discount distribution this quarter versus last, by rep and by segment.

A NOTE ON TIER COUNTS

You will see "the average SaaS company has 3.2 public tiers" and "standard discount is 23% for three-year commitments." Both come from a single self-published study of about 100 companies with no methodology appendix. Directionally plausible; not something to plan around.

If Layer 4 is your constraint

Pricing work is unusually fast. Most of the value comes from two weeks of research and one deliberate change — the rest is nerve.

1. **Assign an owner this week**
   One named person, with authority to change price for new customers. Not a committee. The absence of an owner — not the absence of analysis — is why pricing has not moved.

2. **Pull your discount distribution**
   By rep, by segment, by quarter. Plot it. A tight distribution around a modest discount means your price is believed. A wide one means it is fiction, and that fiction is costing you more than any campaign is producing. This takes an afternoon.

3. **Test the value metric against your happiest customers**
   Ten accounts. Plot what they pay against the value they describe. Where the relationship breaks down, you have found either a mispriced segment or the wrong unit of charge. Both are fixable within a quarter.

4. **Run willingness-to-pay research on one segment**
   Van Westendorp questions on 30–50 buyers in your best-fit segment, or a structured analysis of which deals closed fastest at which price. Either beats competitive benchmarking, which tells you what other people guessed.

5. **Change one thing, for new customers only**
   A price, a tier boundary, or a packaging line. Grandfather existing customers. Measure win rate, sales cycle length and discount depth for a quarter. This de-risks the change almost entirely and produces the evidence for the next one.

IF EVERYONE NEGOTIATES YOU DOWN

That is a positioning symptom, not a pricing one. Buyers negotiate hardest when they cannot tell what makes you different from the alternative. Check your Layer 0 score before you rebuild the price list.

IF EXPANSION IS MANUAL

Every upsell requiring a negotiation means your value metric does not grow with the customer. Fixing the metric converts expansion from a sales activity into an accounting one — which is the whole mechanism behind high-NDR businesses.

EVIDENCE YOU ARE DONE

- A named pricing owner
- Discount distribution reviewed quarterly
- A value metric that grows with the customer
- One structured price change shipped and measured

5

Expansion.

Do your existing customers become an engine, or an endpoint? Expansion is the layer that decides whether growth compounds or has to be bought again every quarter — and by $100M ARR it is where most of your net new revenue will come from, whether or not you have built for it.

What breaks at Layer 5

Expansion is the most consistently benchmarked metric in SaaS and the least consistently owned function. The shape of the curve is not in dispute; whether anyone in your company is responsible for it usually is.

EXPANSION AS A SHARE OF NEW ARR, BY STAGE — FIVE SOURCES, ONE SHAPE

ARR BAND	EXPANSION SHARE OF NEW ARR	SOURCE
$1–5M	12–14%	Emergence Capital (600+ cos); OpenView
$20–50M	35–38%	OpenView, via Ordway's compilation
All companies, median	40%	Benchmarkit 2025 — up from 25% in 2022
$50–100M	58%	Benchmarkit 2025
$100–200M	50%	ICONIQ (96 B2B companies)
$200M+	67%	ICONIQ; corroborated by Benchmarkit at >$100M

The trajectory is the point, not the individual rows: expansion goes from roughly an eighth of new revenue at $1–5M ARR to two thirds at $200M+. The $50–100M and $100–200M rows come from different panels and invert — a sample artefact, not a real dip. A company that has not built an expansion motion by $20M ARR is planning to acquire its way through a stage where almost nobody successfully does.

THE COST ARGUMENT

$2.00
S&M cost per $1 of new-customer ARR

$1.00
S&M cost per $1 of expansion ARR

Expansion revenue costs about half what new revenue costs. Note that expansion CAC has risen sharply — from $0.69 in 2022 — so this gap is not fixed, and worth re-checking against your own numbers rather than assuming.

Benchmarkit, 2025.

THE WIN-RATE ARGUMENT

Customer-success-sourced pipeline	52%
Sales-sourced	38%
Partner-sourced	35%
Marketing-sourced	23%

Pipeline that originates in your existing customer base wins at more than twice the rate of marketing-sourced pipeline — and at most companies it is the smallest and least resourced source, at 4–9% of total.

ICONIQ, State of GTM 2026.

Four mechanisms. One owner *each.*

Winning by Design identifies four distinct expansion mechanisms. Most companies run one and a half of them, and cannot say who is accountable for any.

01

Upsell

More of the same — seats, volume, capacity. The mechanism most companies have, because it happens whether or not anyone drives it.

Owner test: if this is nobody's objective, you are collecting passive expansion and calling it a motion. Who has a number?

02

Cross-sell

A different product or module. Requires a second thing worth buying and a reason for this customer to buy it now.

Owner test: can you name the trigger event that makes a customer ready for product two? If not, cross-sell is happening by accident.

03

Renewal & extension

Longer terms, broader scope, more of the organisation. The multi-year data on page 33 lives here: +5.5 points of gross retention.

Owner test: is renewal a date in a calendar, or a motion that starts 90 days out with a defined play?

04

Re-selling

Re-earning the account when your champion leaves. **The most-missed mechanism in B2B SaaS** and the one that quietly determines your logo retention.

Owner test: when a champion's email bounces, does a defined play trigger — or does someone notice at renewal?

THE LOOP TEST

Draw your growth loop. Input, output, and the mechanism that turns the output back into input.

If you cannot complete that diagram, you do not have a loop — you have a funnel, and funnels do not compound. This is the core of Balfour, Winters, Kwok and Chen's argument in *Growth Loops are the New Funnels*: a funnel models a linear input to output, which means growth stops the moment the input stops.

Be honest about whether your loop is a *product* loop or a *sales* loop. Many B2B SaaS companies have no product-native loop, and their engine is revenue funding more reps funding more revenue. That is a real engine — capital-constrained rather than compounding, and it should be planned as such.

Layer 5 — diagnostics

Score each 0–3. ◆ marks a high-signal question.

01 ◆ What share of net new ARR comes from expansion versus new logos, and how has it moved over eight quarters? 0 1 2 3

Compare against the stage table on page 44 — roughly 12–14% at $1–5M, 35–38% at $20–50M.

02 Which of the four expansion mechanisms do you actually run, and who owns each by name? 0 1 2 3

A 3 requires four names, or an explicit decision not to run one.

03 What is the trigger that tells you an account is ready to expand? 0 1 2 3

A 3 means it is instrumented and fires automatically, not noticed in a QBR.

04 ◆ When a champion leaves, does a defined re-selling play trigger? 0 1 2 3

The most-missed mechanism. If the answer is 'we find out at renewal', score 0.

05 ◆ Draw your growth loop — input, output, and the reinvestment mechanism. 0 1 2 3

If you cannot draw one, what is your engine? Score honestly.

06 Which loop or channel is closest to saturation, and what is the S-curve after it? 0 1 2 3

A 3 requires work on the successor to have already started, 12–18 months ahead.

07 What proportion of new business comes from referrals or existing-customer introductions? 0 1 2 3

There is no credible published benchmark for this. Measure your own number and trend it.

08 ◆ Is expansion in anyone's compensation plan? 0 1 2 3

If expansion belongs to everyone it belongs to nobody, and it will stay passive.

LAYER 5 SCORE

— / 24

0–11 Broken · 12–17 Fragile · 18–21 Solid · 22–24 Compounding

A STAGE CAVEAT

Below roughly $5M ARR, a low Layer 5 score is expected and not necessarily your constraint — expansion is 12–14% of new ARR at that stage even in healthy companies. What matters at that size is not *having* an expansion engine but having **decided who will own it and when you will start**. Above $20M ARR, a low score here is a real constraint and it will get more expensive every quarter you leave it.

Layer 5 — evidence

Expansion sits inside the broader question of what growth rate is actually normal at your stage — which is worth confronting directly, because the folklore is badly out of date.

MEDIAN GROWTH BY ARR BAND (2025)

ARR BAND	ALL B2B SAAS	AI-NATIVE
Under $1M	75%	100%
$1–5M	40%	110%
$5–20M	30%	90%
$20–50M	35%	60%
Over $50M	15%	40%

2025 SaaS Benchmarks Report via Growth Unhinged, n=800+. The $20–50M band sits above $5–20M in both columns; that is what the data says, not a transcription error.

THE T2D3 REALITY CHECK

"Triple, triple, double, double, double" describes roughly the **90th percentile**, not a plan of record. The median $1–5M ARR company grows 40%, not 200%. Only top-quartile companies below $1M ARR (around 300% YoY) approach the pattern.

AI-native companies at $1–5M, with a 110% median, are the only cohort above $1M ARR where a "double" is an ordinary outcome rather than an exceptional one.

EFFICIENCY AT YOUR STAGE

METRIC	BENCHMARK	SOURCE
Rule of 40, top quartile	47	2025 SaaS Benchmarks, n=800+
Rule of 40 by motion	PLG 34 · Sales-led 20	Benchmarkit 2024
Burn multiple guidance	<1× great · 1–2× good · >2× concerning	Benchmarkit 2024/25
ARR per FTE, $20–50M	$350K, +42% YoY	Growth Unhinged 2025
Growth endurance	~65%, down from ~80%	Benchmarkit 2025
Median growth, all private B2B	22% (2025), from 25%	SaaS Capital, n=1,000+
Time to $100M ARR	7.5 years; AI companies 5.7	Bessemer Cloud 100 (2025)
2024 plan vs actual	Actual came in 9 pts below plan	Benchmarkit / Maxio 2025

READ THAT LAST ROW AGAIN

In Benchmarkit's panel, actual growth landed **nine points below plan** — 26% against a planned 35%. SaaS Capital's larger, more bootstrapped panel puts 2024 at 25%. If your board plan assumes you will be the exception, the plan needs a stated reason why — and "we will execute better" is not one.

If Layer 5 is your constraint

Expansion is the only layer where the fix is almost entirely organisational. The mechanisms are well understood; what is missing is ownership and a trigger.

1. **Measure the split, then trend it**
   Expansion ARR as a share of net new ARR, by quarter, for eight quarters. Put it next to the stage table on page 44. This single chart tells you whether you are on the normal trajectory or acquiring your way through a stage where nobody wins that way.

2. **Assign each of the four mechanisms a named owner**
   Upsell, cross-sell, renewal and extension, re-selling. Four names, or an explicit and recorded decision not to run one. This meeting takes an hour and is usually the intervention.

3. **Build the re-selling play first**
   It is the most-missed and the cheapest. Trigger: a champion's departure detected through email bounce, LinkedIn, or a usage drop from a named account. Play: a defined re-introduction sequence to the successor within fourteen days. Most companies discover the departure at renewal, which is most of a year too late.

4. **Instrument the expansion trigger**
   What usage, headcount or workflow signal indicates readiness? Make it fire into a queue rather than waiting for a quarterly review. Passive expansion is the default state, and it caps you at whatever your customers happen to ask for.

5. **Put expansion in a compensation plan**
   Anyone's. A CS lead, an account manager, an AE. The specific design matters far less than the fact that a number exists and a person carries it.

IF YOUR LOOP IS A SALES LOOP

That is legitimate — revenue funds reps funds revenue. Plan it as capital-constrained rather than compounding: model the hiring, the ramp and the payback explicitly, and stop describing it internally as though it were a product loop. The two fail in different ways and need different early warnings.

IF YOUR BEST LOOP IS SATURATING

Verna and Balfour both make the S-curve point: the successor loop has to be started well before the current one flattens — a year or more, in practice. If you are noticing saturation now, you are already late — which is an argument for starting the second curve while the first still looks healthy.

EVIDENCE YOU ARE DONE

- Expansion share tracked quarterly, against stage
- Four mechanisms, four named owners
- A re-selling play that triggers automatically
- An instrumented expansion signal
- Expansion in someone's comp

III

Acting on the diagnosis.

A score is only worth the decision it changes. This part turns 48 answers into one constraint, one quarter of work, and a dozen numbers you will actually look at again.

The scorecard

Short forms of all 48 diagnostics. Score 0–3 using the anchors on page 11; ◆ marks the twenty-four high-signal questions. Give this spread to three colleagues and have them fill it in blind before comparing.

0 Fit

TOTAL ___ / 24

1	ICP written down, with disqualifiers	◆	0 1 2 3
2	Closed-won vs closed-lost pattern analysed		0 1 2 3
3	Real competitive alternative known	◆	0 1 2 3
4	No-decision losses measured	◆	0 1 2 3
5	Market category chosen deliberately		0 1 2 3
6	Anxiety and habit addressed in the funnel	◆	0 1 2 3
7	Segmented PMF survey run		0 1 2 3
8	Model-market arithmetic done honestly		0 1 2 3

1 Acquisition

TOTAL ___ / 24

1	Largest channel's share of new revenue known	◆	0 1 2 3
2	Fully loaded CAC by channel		0 1 2 3
3	Founder-sourced pipeline stripped and measured	◆	0 1 2 3
4	Product has the attributes the channel needs		0 1 2 3
5	ACV supports the acquisition motion	◆	0 1 2 3
6	A channel has been killed on evidence		0 1 2 3
7	A written reinvestment mechanism exists	◆	0 1 2 3
8	CAC payback trend over eight quarters		0 1 2 3

2 Activation

TOTAL ___ / 24

1	Activation event defined and instrumented	◆	0 1 2 3
2	Time from signature to first impact, owned	◆	0 1 2 3
3	Single-user vs team value understood		0 1 2 3
4	Motivation, ability and permission checked		0 1 2 3
5	Drop-off point observed in recordings	◆	0 1 2 3
6	Best vs worst cohort, week one compared		0 1 2 3
7	Activation metric spans multiple roles		0 1 2 3
8	Signature-to-month-two has a named owner	◆	0 1 2 3

3 Retention

TOTAL ___ / 24

1	GRR reported separately from NDR	◆	0 1 2 3
2	Cohort curves by segment, flattening level known		0 1 2 3
3	Churn categorised as bad-fit at sale	◆	0 1 2 3
4	Logo churn vs revenue churn gap known		0 1 2 3
5	Involuntary churn and dunning recovery measured	◆	0 1 2 3
6	Re-selling play for champion departures		0 1 2 3
7	Impact evidence for top ten accounts		0 1 2 3
8	Growth without acquisition spend, understood	◆	0 1 2 3

4 Monetisation

TOTAL ___ / 24

1	Value metric correlates with value received	◆	0 1 2 3
2	Discount distribution, not just the average	◆	0 1 2 3
3	Named pricing owner, recent change		0 1 2 3
4	Pricing matches the GTM motion		0 1 2 3
5	Evidence of who would pay 30% more	◆	0 1 2 3
6	Willingness-to-pay study or price test run		0 1 2 3
7	Tiers map to named segments		0 1 2 3
8	Customers can predict next quarter's bill	◆	0 1 2 3

5 Expansion

TOTAL ___ / 24

1	Expansion share of net new ARR, trended	◆	0 1 2 3
2	Four mechanisms, four named owners		0 1 2 3
3	Expansion trigger instrumented		0 1 2 3
4	Re-selling play fires automatically	◆	0 1 2 3
5	Growth loop drawn: input, output, reinvestment	◆	0 1 2 3
6	Successor S-curve already started		0 1 2 3
7	Referral share measured		0 1 2 3
8	Expansion sits in a compensation plan	◆	0 1 2 3

0 FIT	1 ACQUISITION	2 ACTIVATION	3 RETENTION	4 MONETISATION	5 EXPANSION	CONSTRAINT LAYER

Each out of 24. Your constraint is the lowest — if two tie, take the lower-numbered one.

Reading your score

Three things to read, in this order: the lowest layer, the shape of the profile, and the spread between the people who scored it.

COMMON PROFILES AND WHAT THEY MEAN

PROFILE	WHAT IT LOOKS LIKE	THE READ
The leaky bucket	Layers 0–1 strong, 2–3 weak	You are good at getting attention and bad at keeping it. Every pound of acquisition spend is currently leaking. Stop spending more and fix activation.
The invisible good product	Layers 2–5 strong, 0–1 weak	Customers who find you love you and stay. Almost nobody finds you. This is the most fixable profile in the set and usually a positioning problem, not a budget one.
The bought growth	Layer 1 strong, everything else mid	Growth is real and entirely purchased. Check CAC payback and burn multiple honestly; this profile is only safe while capital is cheap.
The plateau	Everything 14–17, nothing below 12	No single constraint, which is genuinely harder. Work Layer 4 first — monetisation is almost always the most neglected, and it is fast.
The false positive	Everything above 20 on first pass	You scored generously. Re-score with the rule that any answer requiring an explanation is a 1, and have someone outside the leadership team score it too.

THE SPREAD IS THE FINDING

If your CEO scored Layer 1 at 19 and your head of sales scored it at 11, the useful output of this exercise is not the average. It is that two senior people have incompatible models of how the company acquires customers.

Any layer where independent scorers differ by three or more points is an **alignment problem before it is a performance problem**, and no plan built on the average will survive contact with either of them.

ONE QUARTER, ONE LAYER

Take the "If this is your constraint" page for your lowest layer. That is your quarter. Not that page plus three items from elsewhere because they seemed easy.

The discipline is the entire intervention. Most teams already know their weak layer — they work on everything else because everything else is more comfortable, better understood, and produces visible activity sooner. **Visible activity is not the goal.**

What to do with a tie. Fix the lower-numbered layer first. Layer 1 problems corrupt Layer 3 measurements — the bad-fit customers you acquired in January are your churn number in October — while the reverse is not true. Working upward means you are always fixing causes rather than symptoms, even when the symptom is louder.

The stage-fit check

Separate from the layer scores. This asks a different question: is any dimension of your company running ahead of the others? That inconsistency — not any single weakness — is the strongest predictor of failure in the Startup Genome data. The first five dimensions are theirs; * channel is my addition.

DIMENSION	AHEAD OF STAGE LOOKS LIKE	BEHIND STAGE LOOKS LIKE	✓
Customer	Buying traffic before the ICP is validated	No systematic acquisition attempt at all	
Product	Building for imagined future users	Manual delivery propping up the product	
Team	Specialists and managers for processes that are not repeatable	Founders doing work that has been repeatable for a year	
Financials	Capital raised ahead of proof, and now must be deployed	Under-investing in a channel that has already proven itself	
Model	Monetising broadly before value is demonstrated	Leaving money on the table from proven value	
Channel	Scaling a channel validated by the founder's network	Refusing to commit to any channel	

SEED — THE PATTERN TO AVOID

- Hiring a head of growth before the founder has proven the pitch
- Mistaking early adopters for a market
- Building for the loudest customer
- Six channels at 10% effort each
- Buying growth without product/market fit

The compounding harm: with no fit, you cannot tell whether stagnation is bad execution or an absent market.

SERIES A — THE TRANSFER PROBLEM

- Founder-led sales that never got documented
- Channel–model mismatch arriving as rising CAC
- Premature specialisation into SDR / AE / CSM
- Positioning inherited from a competitor
- Marketing measured on MQLs sales rejects

The tell: the founder is still on every deal above a threshold, and rep ramp time keeps slipping.

SERIES B — THE COMPOUNDING PROBLEM

- Scaling a channel the founder's network was actually feeding
- No engine — only lubricants and turbo boosts
- Retention debt surfacing as first cohorts renew
- ICP drift — two ICPs, two motions, one budget
- S-curve saturation with no successor started

Gross bookings grow while net growth collapses. This is the most expensive surprise in SaaS.

How to use this page. Tick the dimensions that are ahead. If you tick more than one, the correction is not to accelerate the laggards — it is usually to *slow the leaders* until the rest catch up. That is a much harder conversation and it is the one the evidence supports.

Prioritising the work

Once you know your constraint, you will have more ideas than quarter. Three published frameworks, and the honest case for when each one is wrong.

SEAN ELLIS

ICE

Impact, Confidence, Ease — each rated 1–10, averaged.

Use when: seed stage, weekly triage, no data.

Wrong when: you need comparability. Scores are entirely subjective and not comparable across people or across time. Peep Laja's objection lands: if you could guess the impact, why run the test?

SEAN MCBRIDE, INTERCOM

RICE

(Reach × Impact × Confidence) ÷ Effort, with reach as a real count and effort in person-months.

Use when: you have volume data and multiple teams to compare across.

Wrong when: enterprise B2B, where reach is small and deal value is large. It systematically under-rates strategic bets. Common fix: substitute revenue-at-stake for reach.

PEEP LAJA, CXL

PXL

Binary, evidence-based questions instead of subjective ratings. Is it above the fold? Is it backed by user testing? By session recordings?

Use when: you run a mature experimentation programme.

Wrong when: you have no user testing, heatmaps or qualitative research — most Series A companies score zero on every evidence criterion, which makes PXL unusable rather than demanding.

THE EXPERIMENTATION REALITY CHECK

Win rate, all experiments	~20%
Win rate, revenue-linked tests	~10%
Conclusive rate (reach significance)	35–40%
Below-33% conclusive rate	Red flag

Optimizely programme benchmarks; corroborated by ConversionTeam's audit of 2,288 tests (19.1% statistically significant winners per test).

A win rate that is too high is a red flag, not a success. Above 50% significant wins usually means peeking at results, no significance discipline, insufficient sample size, or only testing changes you already knew would work — which means you are not learning anything.

And the finding that matters most for companies at your stage: **most Series A B2B SaaS companies do not have the traffic to run more than one or two statistically valid web tests a month.** That is itself the diagnosis. It should redirect you from CRO towards qualitative research and larger, bolder changes that do not require statistical proof to justify.

The 90-day rhythm

The same shape regardless of which layer is your constraint. Two weeks of evidence, six weeks of a single intervention, four weeks of measurement, then re-score.

W1–2

Evidence, not opinion

Pull the data the diagnostics revealed you did not have. This is deliberately the least satisfying fortnight: no shipping, no campaigns, no announcements. The temptation to skip it is exactly proportional to how much you need it.

W3–8

One intervention

The single highest-leverage item from your constraint layer's playbook. One. If your team can name three things they are working on this quarter, you have not chosen — you have listed. Write the intervention on one line and put it where the team sees it daily.

W9–12

Measure and hold

Resist adding a second intervention while the first is being measured. Confounded results are worse than no results, because they produce confident wrong conclusions that persist for years.

W13

Re-score, publicly

All 48 diagnostics again, blind, with the same people. The constraint will usually have moved. That is what success looks like — not a layer reaching 24, but the bottleneck relocating.

WHAT WILL GO WRONG

Week 4: someone brings an urgent idea

It will be genuinely good. Write it down, date it, and do not start it. The list of good ideas you deferred is a healthy artefact; a quarter with four half-finished interventions is not.

WHAT WILL GO WRONG

Week 7: the metric has not moved

Most of these interventions have lags. Win rate moves in weeks; churn moves in quarters. Check the leading indicator you defined in week two, not the outcome you want.

WHAT WILL GO WRONG

Week 13: the score barely changed

Then either the intervention was too small, or the constraint was misidentified. Both are useful findings and both are cheaper than a year of undirected effort. Re-score honestly and choose again.

The one-page dashboard

Twelve numbers. Two per layer. Reviewed monthly by the same people who scored the audit. If a number here has no owner, it will stop being true within a quarter.

LAYER	METRIC	WHY THIS ONE	OWNER	CADENCE
0 Fit	Win rate, and share of losses to no decision	No-decision share is the cleanest positioning signal you have		Monthly
	Deals disqualified on ICP criteria	Proves the ICP is real rather than decorative		Monthly
1 Acquisition	Share of new pipeline from the largest channel	Concentration, not diversification, is the health signal		Monthly
	CAC payback, by channel	The channel-model arithmetic, kept live		Quarterly
2 Activation	Activation rate, by cohort	Leading indicator for everything below it		Monthly
	Time from signature to first measured impact	The Bowtie's Δt6; almost nobody tracks it		Monthly
3 Retention	GRR — reported separately, always	The honest retention number		Monthly
	Involuntary churn and dunning recovery rate	A third of churn, and the cheapest to recover		Monthly
4 Monetisation	Discount distribution, by rep and segment	Variance reveals whether your price is believed		Quarterly
	ARPA for new cohorts vs the base	Shows whether pricing work is actually landing		Quarterly
5 Expansion	Expansion as a share of net new ARR	Compare against the stage table on page 44		Quarterly
	Pipeline sourced from existing customers	Wins at 52% versus marketing's 23%		Monthly

DELIBERATELY NOT ON THIS LIST

MQLs. Website sessions. Social followers. Email open rates. Blended CAC on its own. NDR without GRR beside it. Each of these can move substantially in either direction without telling you anything about whether the business is healthier.

THE TEST FOR ADDING A THIRTEENTH

Before a metric joins this page, answer: **what decision will we make differently depending on its value?** If there is no answer, it belongs in a dashboard nobody opens, not on the page the leadership team reviews.

Benchmark library

The published benchmarks used in this document, with source, year and sample. Read the note on page 10 before comparing yourself to any of it.

GROWTH & EFFICIENCY

METRIC	VALUE	SAMPLE / SOURCE
Median ARR growth, all private B2B	22% (2025), from 25% (2024)	SaaS Capital, n=1,000+
Median growth by ARR band	<$1M 75% · $1-5M 40% · $5-20M 30% · $20-50M 35% · >$50M 15%	2025 SaaS Benchmarks / Growth Unhinged, n=800+
Same bands, AI-native	100% · 110% · 90% · 60% · 40%	Growth Unhinged 2025
Actual vs planned growth (2024)	9 points below plan	Benchmarkit / Maxio 2025
Rule of 40, top quartile	47	2025 SaaS Benchmarks, n=800+
Rule of 40 by GTM motion	PLG 34 · sales-led 20	Benchmarkit 2024
ARR per FTE, $20–50M band	$350K, +42% YoY	Growth Unhinged 2025
Burn multiple guidance	<1× great · 1–2× good · >2× concerning	Benchmarkit 2024/25
Time to $100M ARR	7.5 years; AI companies 5.7	Bessemer Cloud 100, 2025
Odds of reaching scale	~50% reach $1M · ~10% $10M · ~2% $25M	ChartMogul 2025, n=6,525

FIT

METRIC	VALUE	SAMPLE / SOURCE
Product/market fit threshold	≥40% “very disappointed”	Sean Ellis, ~100 startups (2009); Slack 51%; Superhuman 22% → 58%
Channel concentration in high-growth companies	>70% from one channel	Balfour (2017)
Early-adopter durability	<20% still using after two years	Garbugli, LANDR case — single company, directional
SQL → closed won, marketing-vetted definition	12%	First Page Sage 2025 — compare ICONIQ’s 28%

ACQUISITION — EFFICIENCY

METRIC	VALUE	SAMPLE / SOURCE
CAC payback, median	**18 months**	Benchmarkit 2025, n=148; up 12.5% since 2022
New-customer CAC ratio	**$2.00 median · $2.82 fourth quartile**	Benchmarkit 2025, n=73
Blended CAC ratio	**$1.81**	Benchmarkit 2025, n=43
Expansion CAC ratio	**$1.00, from $0.69 in 2022**	Benchmarkit 2025
SaaS magic number	**~0.75**	Benchmarkit 2025, n=101
S&M as % of revenue	**37% median · VC-backed 45–47% · PE-backed 33%**	Benchmarkit 2025, n=157
Marketing / sales as % of ARR	**8% / 15%**	SaaS Capital 2026, n=1,000+ (mostly bootstrapped)

ACQUISITION — FUNNEL

METRIC	VALUE	SAMPLE / SOURCE
Visitor → lead	**1.4% median · 2.1–2.3% top quartile**	First Page Sage 2025
Visitor → lead by audience size	**Small 2.3% · SMB 1.4% · mid-market 1.2% · enterprise 0.7%**	First Page Sage 2025
Lead → MQL by source	**Referrals 56% · exec events 54% · SEO 41% · email 38% · PPC 29%**	First Page Sage 2025
Lead → MQL → SQL → closed won	**25% → 27% → 28%**	ICONIQ 2026 (sales-accepted definitions)
Win rate by pipeline source	**CS 52% · sales 38% · partner 35% · marketing 23%**	ICONIQ 2026
Pipeline mix, high-growth companies	**Sales 62% · channel 19% · marketing 15% · CS 4%**	ICONIQ 2026
Sales cycle by ACV	**<$10K 6wk · $10–50K 12wk · $50–100K 17wk · $100K+ 24wk**	ICONIQ 2026
Pipeline coverage	**3.6× (2025), from 3.9×**	ICONIQ 2025
Ramped AEs at quota	**62%**	ICONIQ 2026
Cost per opportunity	**SMB $5.2K · mid/ent $6.3K · strategic $8.0K**	ICONIQ 2026

ACTIVATION — RATES

METRIC	VALUE	SAMPLE / SOURCE
Activation rate, SaaS median	**30% median · 36% average**	Lenny's Newsletter, 500+ products, 2022
Activation rate, B2B	**37% median**	Userpilot 2024, n=62
Activation, PLG vs sales-led	**34.6% vs 41.6%**	Userpilot 2024, n=547
Retention lift from activation	**≥2×**	Lenny's 2022
Strong activation → strong 3-month retention	**69% of products**	Amplitude, 2,600+ companies, 2025
Onboarding checklist completion	**10.1% median · 19.2% average**	Userpilot 2024, n=188
Time to value	**~1.5 days**	Userpilot 2024, n=547 — weakest-evidenced metric here

ACTIVATION — CONVERSION

METRIC	VALUE	SAMPLE / SOURCE
Free trial → paid	**good 8–12% · great 15–25%**	Lenny's / Kyle Poyar, 1,000+ products, 2023
Freemium → paid	**good 3–5% · great 6–8%**	Lenny's 2023
Freemium + sales assist	**good 5–7% · great 10–15%**	Lenny's 2023
Median free-to-paid, six months	**8%; ~10× spread top to bottom quintile**	ChartMogul 2026, n=200
Card-required trial conversion	**~30% — more than 5× no-card**	ChartMogul 2026
Opt-in vs opt-out trial	**18.2% vs 48.8% trial→paid; 8.5% vs 2.5% signup**	First Page Sage, n=86, 2022–2025
Sales outreach to >50% of signups	**Trial conversion 2×, freemium 4×**	OpenView 2023, 1,000+ respondents
Trial → paid with PQL scoring	**~25% vs 9% baseline**	ProductLed 2025, 600+ companies
Enterprise trial/POC → paid	**50%**	ICONIQ 2026 — a late-funnel stage, not a signup
Demo → closed won	**36%**	ICONIQ 2026

RETENTION

METRIC	VALUE	SAMPLE / SOURCE
NRR by ACV	**<$12K 93% · $25–50K 97% · $100–250K 101% · >$250K 105%**	SaaS Capital 2023, n=1,500+
GRR by ACV	**<$12K 90% · ≥$25K plateau at 93%**	SaaS Capital 2023
Median GRR	**88% (2025), from 90%**	Benchmarkit 2025
Contract length effect on GRR	**Monthly 89.5% · annual 90.0% · multi-year 95.0%**	SaaS Capital 2023
Multi-year share of contracts	**14% (2022) → 40% (2025)**	Maxio 2025
SaaS churn, monthly	**3.22% total · 2.16% voluntary · 1.06% involuntary**	Recurly network data 2026
Involuntary share of churn	**20–40%**	Recurly 2026
Dunning recovery	**53% baseline → 71% optimised; 90% within 10 days**	Recurly 2026
NRR >100% vs <60%	**43.6% vs 13.1% annual growth**	ChartMogul 2023
GRR above 85%	**Grows 1.5–2.5× faster**	ChartMogul 2023
NRR 90–100% → 100–110%	**~+5 points of growth**	SaaS Capital 2025
Usage-based pricing effect on GRR	**~3 points lower**	Benchmarkit 2024

MONETISATION & EXPANSION

METRIC	VALUE	SAMPLE / SOURCE
Relative impact of growth levers	**Acquisition 3.32% · retention 6.71% · monetisation 12.7%**	Price Intelligently, N=512 — ~decade old, methodology unpublished
Pricing changes per company per year	**~3.6**	PricingSaaS, 500 companies, 2025
Change mix	**Packaging 43% · price/structure 40% · usage limits 17%**	PricingSaaS 2025
Hybrid subscription + usage pricing	**21% median growth — highest of any model**	Maxio / Benchmarkit 2025
SaaS companies charging for AI features	**44%**	Maxio 2025
Expansion as share of new ARR	**$1–5M 12–14% · $20–50M 35–38% · median 40% · $50–100M 58% · $100–200M 50% · $200M+ 67%**	Emergence, OpenView, Benchmarkit, ICONIQ — five sources; the middle bands come from different panels

EXPERIMENTATION

METRIC	VALUE
Win rate	**~20% all · ~10% revenue-linked**
Significant winners	**19.1% per test**
Conclusive rate	**35–40%; <33% a red flag**
Velocity growth target	**+10–20% YoY minimum**

Optimizely programme benchmarks; ConversionTeam audit of 2,288 tests.

PREMATURE SCALING

METRIC	VALUE
Startups scaling prematurely	**~70%**
Share of failures attributed	**74%**
Advantage of consistency	**~20× by scale stage**
Markers	**2.3× spend · 3× team · ~3× capital**

Startup Genome, *Report Extra: Premature Scaling* (2011).

Numbers you will see that you should *not* cite.

Each of these circulates widely and is either unverifiable, misattributed, or contradicted by its own claimed source.

THE CLAIM	THE PROBLEM
"LTV:CAC should be 3:1"	A 2000s venture heuristic. No 2024–2026 report publishes a median LTV:CAC with a disclosed sample.
"You need 3× pipeline coverage"	No dataset behind it. ICONIQ's measured 3.2–3.9× is the citable substitute.
"Every 10-minute delay in time-to-value costs 8% of trial conversion"	Part of a cluster of viral statistics with no locatable primary study, methodology or research footprint.
A trial-conversion percentile table attributed to ChartMogul	Does not appear in any ChartMogul publication. Attribution appears to be an error propagated between blogs.
"Usage-based pricing delivers 18–23% higher NRR"	Traces to one self-published study of ~100 companies. Primary data shows usage-based models with ~3 points *lower* GRR.
Clean NRR-by-segment tables (118 / 108 / 97)	Published without methodology by AI-content sites. Plausible-looking and fabricated.
"$2–5M ARR per CSM"	Real, but from a 2019 survey, never refreshed, and partly contradicted by more recent stage-based figures.
The cohort "smile curve"	A qualitative story. Primary cohort data shows the best B2B cohorts holding flat or rising, never dipping.
"Annual contracts dramatically reduce churn"	The measured annual-versus-monthly GRR gap is 0.5 points. The real step change is at multi-year (+5.5).
T2D3 as a plan of record	Describes roughly the 90th percentile. The median $1–5M ARR company grows 40%.

Currency note. Benchmark sources publish in US dollars and their figures are reproduced unconverted. Where this document gives illustrative amounts in pounds, they are illustrative. Do not convert a benchmark and present it as a local figure — the underlying samples are predominantly US companies, and conversion adds false precision to an already wide distribution.

> **Why this page exists.** A benchmark you cannot defend is worse than no benchmark, because it produces confident decisions. If you take one habit from this document, take this one: before acting on a number, find out the sample size, the year, who paid for the research, and what definition was used. It takes four minutes and it will save you at least one bad quarter.

Sources

FRAMEWORKS

Brian Balfour, *Four Fits for $100M+ Growth* and *Product–Channel Fit* (2017) · brianbalfour.com

Brian Balfour, Casey Winters, Kevin Kwok & Andrew Chen, *Growth Loops are the New Funnels* · reforge.com

Lenny Rachitsky & Dan Hockenmaier, *The Racecar Growth Framework* · reforge.com, lennysnewsletter.com

Dan Hockenmaier, *Why Growth Models Fail* · danhock.co

Sean Ellis, *The Startup Pyramid* (2009) · seanellis.substack.com

Rahul Vohra / First Round Review, *How Superhuman Built an Engine to Find Product/Market Fit* (2019)

Tristan Kromer, *False Positives and Product/Market Fit* · kromatic.com

April Dunford, *Obviously Awesome* (2019) and *Sales Pitch* (2023)

Bob Moesta & Chris Spiek, the Forces of Progress · jobstobedone.org

Winning by Design, *The Bowtie: A Proposed Standard* · winningbydesign.com

Elena Verna, *To PLG or not PLG* and *My 9 Favorite Growth Frameworks* · elenaverna.com

Amplitude & John Cutler, *The North Star Playbook*

Dave McClure, *Startup Metrics for Pirates* (2007); Gabor Papp & Thomas Petit, RARRA (2017)

Sean McBride, *RICE* · intercom.com · Peep Laja, *PXL* · cxl.com

Startup Genome, *Report Extra: Premature Scaling* (2011)

Etienne Garbugli, *Lean B2B* · leanb2bbook.com

BENCHMARK DATASETS

Benchmarkit / Pavilion, *B2B SaaS Performance Metrics* 2024 and 2025 · benchmarkit.ai

SaaS Capital, *Retention Benchmarks* (2023), *Spending Benchmarks* (2026), *Growth Rate Benchmarks* · saas-capital.com

ICONIQ Growth, *The State of GTM* 2025 and 2026; *State of Software* 2025 · iconiq.com

ChartMogul, *SaaS Retention Report* (2023, 2025), *SaaS Conversion Report* (2026), *SaaS Growth* (2025) · chartmogul.com

Recurly Research, *Churn Rate Benchmarks* and failed-payment recovery (2026) · recurly.com

Maxio, *2025 SaaS Pricing Trends Report* · maxio.com

PricingSaaS via Growth Unhinged, *State of SaaS Pricing Changes* (2024, 2025)

Growth Unhinged, *2025 SaaS Benchmarks Report* · growthunhinged.com

Bessemer Venture Partners, *Cloud 100 Benchmarks* (2025) · bvp.com

KeyBanc / KBCM, *Private SaaS Company Survey* (2025)

ProductLed, *Product-Led Growth Benchmarks* (2025, 600+ companies)

OpenView, *Product Benchmarks Report* (2023)

Lenny's Newsletter / Kyle Poyar, activation and free-to-paid benchmarks (2022, 2023)

Userpilot, activation, onboarding and product-metrics benchmark reports (2024)

Amplitude, *B2B Technology Product Benchmarks* (2025); Mixpanel, *Benchmarks Report* (2024)

First Page Sage, CAC-by-channel, funnel-conversion and free-trial benchmark reports (2025) — agency-published; see the caveat on page 23

Optimizely, experimentation programme metrics; ConversionTeam, audit of 2,288 A/B tests

Emergence Capital and Ordway, expansion-ARR compilations

Send me your *scores.*

If you have worked through the diagnostics and want a second read on them, send me the six numbers and your one-line answer to the lowest-scoring layer. I will come back with what I think your constraint actually is and the first thing I would do about it.

No deck, no proposal, no sequence. If it turns out you want help afterwards, we can talk about that separately — and if the answer is that you should fix it yourself, I will say so, because that is the answer more often than a consultant's incentives suggest.

SEND SCORES TO

partneringrowth.co.uk

Book a 30-minute call — no commitment

OR ON LINKEDIN

Hilal Tasdan

Message "audit" with your six scores

PARTNER IN GROWTH

Fractional CMO and go-to-market work for B2B SaaS companies scaling pipeline in crowded markets. London, working across the UK and Europe. Embedded leadership rather than a retainer — in your standups, your Slack and your CRM.

SELECTED RESULTS

Access PaySuite · TOFU growth	54×
Checkbook.io · paid CAC	-41%
Uptechlab · qualified leads	3.1×
Client retention	94%