



FIRST STRATEGY · CASE STUDY

Mid-Size Manufacturing

A real engagement, anonymized to industry label. The full case study: the story, the deliverables we produced, and the plays that ran it.

The story

Day One Proposal

Day One Audit

Playbook and Delivery Proposal

Charter

The plays

The founding story was true, and it was the wrong pitch

The company exists because of a refusal. When the category's incumbents protected the chemistry their products depend on, the owner went the other way and engineered it out, building a material that is stronger than the commodity standard and clean enough that the usual warnings do not apply. For years that material earned its living quietly, sold as an input into other companies' products, its story told by whoever put their brand on the finished good. The owner wanted the story back. The President, who carries the consumer line, came to us on a referral with the plan: take the material to consumers under the company's own name, and launch it where consumers buy.

The ask sounded like marketing. Messaging, packaging, costing, a storefront, advertising. We asked for a diagnostic stretch before proposing anything, working sessions with leadership and our own research alongside, because the operation that runs in the documents is never the operation that runs, and a launch that has not happened yet hides its problems in the plan instead of on the floor.

The first sessions told us what the company had. The plant was the easy part of the diagnostic: the material is real, the line runs, and plant leadership could explain exactly why the product outlasts the commodity standard it would compete against. The rest of the stretch told us what the company did not have. The consumer plan crossed eight steps from the line to a reorder, and five of them did not exist. No retail costing. No listings. No way to be found. No pitch that wins the moment a buyer compares it to a commodity product at a fraction of the price. And the team that would build all of it was [a handful of] people, carrying the launch alongside the jobs they already had, where a launch like this is conventionally carried by a team of ten.

Then there was the pitch itself. The company planned to lead with the mission, the cleaner material, the founding refusal, because the mission is why the company exists. We did not argue with it in the room. We took it as the first assumption to test, because a founding conviction is the most expensive kind: nobody inside can see past it, and every dollar of the launch was about to follow it.

The shelf had already written the pitch

The research half of the diagnostic let the category argue with the company's instincts. There was no floor to shadow, so we shadowed the shelf: [several thousand] customer reviews across the top sellers in the category, mined for what buyers actually complain about, praise, and pay for.

The complaint pile was a portrait of a shelf that fails its people. Around [half] of the negative reviews described the same death: products tearing in a season, hardware pulling out, sun rot. Another [quarter] came from buyers on their second or third replacement, and the remarkable thing was what they did in those reviews: they wrote the cost-per-use arithmetic themselves, unprompted, adding up what three cheap purchases had cost them. The safety worries were there too, the smell on unboxing, the residue,

the questions about kids and pets and soil, but they read as reassurance sought rather than reasons to buy. And the mission language the company planned to lead with appeared mostly in thin praise that almost never named itself as the reason for the purchase.

The category's own customers had written the pitch, and it was the company's pitch inverted. Open with the failure the buyer already knows. Prove the lifespan with arithmetic. Close with the chemistry no one else can claim, the reason to feel good about a choice the toughness already won. The owner accepted the inversion, which could not have been easy, because the evidence was not our opinion. It was the category's one-star reviews. The mission did not leave the pitch. It moved to the end of it, where it lands.

The same evidence defended the price. The product would sit at [several times] the commodity anchor, and the reflex under pressure would have been to discount toward the shelf. The costing desk and the mining agreed there was nothing to apologize for: a service life of 3 to 5 times the standard makes the premium the cheaper product across three years. The flagship unit holds gross margin in the high fifties of percent, margin the same material never sees as an industrial input, and that margin was the entire point of the launch. One size nearly died in that review, the smallest, its shipping economics too thin to defend. It survived as a named recruitment cost, watched quarterly, written into the decision log so nobody later mistakes a deliberate bet for an oversight.

A brand a machine can be governed by

What we built first was not a listing. It was a constitution: the positioning, every claim with its proof attached, the required language, the banned language, and the two rules that do not bend. Strength and safety travel in the same breath, always, because either one alone collapses into a category the company loses. And nothing is claimed that is not proven, because the mining had shown a shelf where every product says the same two words and not one demonstrates them.

The constitution is what made the lean team possible. A catalog's worth of content cannot be produced by [a handful of] people by hand, and a machine cannot be trusted with a brand's claims unless the brand is written down as law. So the machine drafted everything, every title, every bullet, every block of listing content, every campaign line, and the President read every word against the rules before it shipped. The gate earned its keep in the first weeks. Drafts kept reaching for the category's reflexes, the empty superlatives, the mission-first framing the constitution had banned, and every catch went into a log, and the log went back into the rules. The machine got measurably better at sounding like the brand for the same reason the brand existed at all: someone had written down what it would not say.

The storefront went live under that system, and the advertising went live with it, every keyword an experiment against targets set as law. The reads settled into a pattern that held for the rest of the engagement: specific beats generic, the use-case term beats the category term, the account's best term converting at 2.6 percent cost of sale while the generic terms ran hot and got cut. The chemistry-free buyer turned out to be real and rare, a small pool converting at single digits, a niche worth its own

campaigns and a quiet confirmation of the inversion: worth serving directly, never worth betting the shelf on.

Words drift the way models do

The catalog grew to 19 SKUs across three product families, and the constitution grew with it, learning from the market's feedback. Which produced the failure nobody was watching for, because no dashboard measures it. The dashboards watched sales, spend, and reviews, and all of them looked fine. The question nobody had asked was whether the live catalog still obeyed the law.

The President asked it, and a reconciliation review audited every live listing against the current constitution. The catalog had drifted. A strength multiple was live in one listing that traced to no documented proof point, exactly the kind of unproven claim the constitution exists to prevent, shipped anyway somewhere in the catalog's growth. The core lifespan claim appeared phrased three different ways across the family. The newest messaging angles existed on paper and nowhere on the shelf. The vehicle lines were selling on the family's generic pitch instead of their own.

The unproven claim was retired the day it was found, not defended. Each proof point got one canonical phrasing, written into the constitution and reused verbatim. And the audit became a standing cadence, because the lesson deserved to outlive the incident: words drift the way models drift, quietly, in the blind spots of aggregate metrics, and they get caught the same way, by auditing outputs against the current rules instead of trusting that approval at launch means compliance forever.

Success built the next bottleneck

The launch worked, and the working launch created the engagement's second problem. The consumer line's data lived in three places: the product record in [a master spreadsheet], the listings in the storefront's own console, fulfillment in its own system. At launch, with one family, the gap between them was a tolerable chore. At three families and 19 SKUs, every change meant the same information rekeyed by hand into three systems, and the senior judgment that made the operation work, the costing logic, the fulfillment calls, the feel for what sells where and when, sat in a few heads. Inventory cycles ran long because the operation could not risk anything shorter; nobody dared run lean on data they could not trust to be the same in all three places.

The instinct under that load is to hire someone to do the rekeying. The decision was to connect instead, because rekeying scales with the catalog and headcount patches a structural problem. The second project rebuilt the consumer go-to-market as one connected motion, from inside the commercial leadership team. The product record became the single source of truth and the channel now reads from it: channel details publish themselves, and a product description adapts to market conditions in hours instead of weeks. Fulfillment was wired to the same spine. The judgment moved from heads into the systems, costing logic, fulfillment rules, channel patterns, written down and running. An analytics tool

built over the storefront started feeding the brand's own decisions, and it pulled ad spend down while average order value went up.

The operational results came in hard numbers. Fulfillment dropped by two days. Inventory holding fell by 80 percent, and the operation runs effectively just-in-time now, on cycles that would have been unthinkable when the engagement started. The manual rekeying is gone. The operation that could not risk running lean now runs lean by default, because the data is one thing instead of three.

The second product proved the system

The real test of a system is whether it runs twice. The vehicle lines had already shipped on the same machinery in a fraction of the first family's effort. The second category was the harder test, because it required the discipline to choose boringly. Two candidates were louder, a premium outdoor line at an attractive price band and a product with an easy story. The mining picked the quiet one: floor protection for the professional trades, a product whose incumbents force a choice between durability, grip, easy cleaning, and clean chemistry, whose buyers complain about slip and odor in the same breath, and where not one competitor on the shelf proves a claim.

That last fact set the launch plan. The differentiating claim went to an accredited lab before it went anywhere near a listing. In a category that asserts, the brand that demonstrates owns the trust, and the engagement had made that a standing rule: certify first, claim second. The category goes to the trade first, seeded with working professionals whose reviews carry weight with their own, its first 90 days targeted and written down before launch rather than rationalized after.

The engagement continues, and honestly: we still run the go-to-market machinery day to day, by design, because the client's people belong on the plant and the product, not in campaign hygiene. What has transferred is the part that matters more. The constitution is theirs. The decision log is theirs. Every claim, every price, every next-product call is made by their people, on evidence the system surfaces weekly in a twenty-minute read. The President started this engagement asking for a launch. What the company got was the machine that launches, governed by rules their own judgment wrote, and a shelf of candidates waiting for the same treatment.

Day One Proposal

The opening diagnostic

Prepared for the President, on referral.

What this is

A short diagnostic stretch before anything gets built. Real work, not slides.

We work in sessions with you, the owner, and whoever else carries weight on the consumer line. What the material is and what it proves. What the consumer move is meant to earn. What exists today and what does not. Where AI already shows up in the business, and where it could carry work your team does not have the hands for.

Alongside the sessions, we dig on our own. The plant's story, with the people who can explain what makes the material different. The numbers, with whoever owns costing, because a consumer unit lives or dies on its margin math. And the launch itself as you currently imagine it: the product, the shelf it competes on, the price you think it can hold, and the story you intend to tell. This is a launch that has not happened yet, so the diagnostic stress-tests the plan before money goes into branding, tooling, or ad spend.

We come ready to listen and to think on our feet. No prepared deck.

What you walk away with

A playbook, at the end of the stretch. Not a deck. Not a recommendation memo hiding in a PDF. A written read for operators.

The sessions pull signal from your leadership and your plan. The research alongside tests what we hear against the category's own evidence: what its buyers complain about, what they pay for, and what they ignore. That work turns the signal into a sequenced plan you can run.

The playbook answers three questions:

- Where AI fits in this launch, and where it does not.
- The highest-leverage moves we see, sequenced so you can act on them in order. A roadmap, not a list.
- What it would take to run the sequence: with your own team, with another firm, or with us.

The playbook is yours. Run it however makes sense.

What we need from you

- Your leadership in the sessions: the owner, the President who carries the consumer line, and whoever owns the costing.

- Time with someone who can explain what makes the material different, plant included.
- The launch plan as it stands, however rough, including the price you believe the product can hold.

The terms

A flat fee of [flat fee] for the diagnostic and the playbook. Travel and expenses billed at cost, on top.

No retainer. No commitment beyond the diagnostic itself. If we are the right fit for what comes next, we will already have been talking about what that looks like. If we are not, the playbook is still yours to run.

What happens next

After the playbook, you decide. Run it with your own team, hand it to another firm, or build it with us. If the work points to a build we are right for, we will scope it in a separate proposal once the playbook has shown what is worth building.

Day One Audit

The one-line finding

Your launch is not a marketing project. It is a systems project, and the pitch is backwards. The material is ready: stronger than the commodity standard, free of the chemistry the category depends on, and certified where the category makes claims. What does not exist is the machine that sells it, and the team that would normally be that machine does not exist either. The launch needs the output of a ten-person go-to-market operation from a team a fraction of that size, which means AI carries the production work and your people keep the judgment. And the story you planned to lead with, the environmental story, is the close, not the open. The category's buyers pay for tough. They feel good about clean. Build the system, invert the pitch, and the margin a consumer brand holds becomes yours instead of your customers'.

How we looked, and how we measured

A diagnostic stretch at the start of the engagement, not a single day. Working sessions with the owner, the President, and the people who own costing and the plant, working the launch plan as it stood: product, price, shelf, and story. Our own research ran alongside the sessions, and the research is where the measuring happened. A launch that has not happened yet has no floor to shadow, so the floor we shadowed was the category's: [several thousand] customer reviews across the top sellers on the shelf this product will compete on, mined for what buyers actually complain about, what they praise, and what they say they would pay to fix. Where a figure below comes from that mining, this audit says so. Where a figure is the company's own costing, it is the engagement record.

What exists, and what the launch needs

The operating systems of an industrial manufacturer are real and load-bearing. None of them reach the consumer.

| Capability | State at the start | What the launch needs |
|---------------------------|--|--|
| Manufacturing and quality | Running at industrial volume, the company's muscle | Nothing new. The plant is ready |
| Costing | Built for input pricing, by the roll and the order | Retail unit economics: landed cost, fees, shipping, margin per SKU |
| Brand | The material's story, told so far inside other companies' products | A consumer brand with rules tight enough to govern every word |
| Content | Spec sheets for industrial buyers | Listings, images, copy, and campaign content for every SKU, continuously |
| Channel | Direct industrial relationships | A marketplace storefront, its advertising engine, and the discipline to run it |
| Measurement | Production metrics | A weekly read of revenue, ad spend, reviews, acquisition cost, and returns |

The first row is the asset. The other five rows are the project.

Stakeholder map

Each role holds a different piece of the launch. None is wrong. None alone is sufficient.

| Role | What they own | Where their conviction is | Their definition of the challenge |
|------------------|--|---|---|
| Owner | The company and the material's reason for existing | The mission: the chemistry the material removed | Tell the story we were founded on |
| President | The consumer line | The opportunity: margin and a brand of their own | Get launched, lean, without betting the company |
| Finance | Margin | The arithmetic: consumer economics are unfamiliar terrain | Prove the unit math before the ad spend |
| Plant leadership | Capacity and runs | The material: it does what the category's cannot | Do not let consumer volume disrupt industrial commitments |

The gap that matters is between the owner's conviction and the category's behavior. The mission is real and it is why the material exists. Whether it is why a consumer buys is a different question, and the evidence below answers it.

The shelf, traced end to end

We traced the path one consumer unit must travel, from the line to a reorder. Eight steps, five of which did not exist at the start.

1. Material comes off the line. Exists, proven.
2. The material is converted and packed as a consumer unit. Partially exists; consumer packaging does not.
3. The unit gets a price that survives fees, shipping, and returns. Did not exist; industrial costing does not transfer.
4. The unit gets a listing: title, images, copy, claims. Did not exist.
5. The listing gets found: search placement and paid traffic. Did not exist.
6. The buyer chooses it over a commodity unit at [several times] the commodity price. The pitch that wins this moment did not exist.
7. The buyer reviews it. The product's durability makes this step an asset, but only if step 6 told the truth.
8. The buyer comes back, and the brand's take-back program turns the end of the product's life into the next purchase. Unique to this company; unbuilt.

The trace is the project plan in miniature. Steps 3 through 6 are where the work is, and step 6 is where launches in this category die.

The category's friction, quantified

What the review mining found on the shelf this product enters. The commodity standard is cheap, and its buyers are not happy. The share figures are from the mined sample.

| What buyers complain about | Share of negative reviews | What it tells us |
|--|---------------------------------|--|
| Failure within a season: tearing, grommets pulling out, UV rot | [roughly half] | The category's core promise is broken. Durability is the open |
| Replaced the same product repeatedly | [a quarter] | Buyers already do the cost-per-use math when prompted |
| Chemical smell, residue, worry about kids, pets, soil | [a meaningful minority] | Safety matters, as reassurance rather than as the reason to buy |
| Mission language without proof | thin praise, thinner conversion | Nobody in the category proves claims. Certification is differentiation |

The pattern is consistent: the category's buyers are failure-driven first, value-driven second, and safety-reassured third. The environmental story has an audience, but it is a segment, not the shelf.

The pitch inversion

The company planned to lead with the mission: the cleaner material, the chemistry removed, the founding story. The mining says lead with the failure the buyer already knows: the commodity product that did not last the season. Open with tough. Prove it with the lifespan arithmetic. Close with clean, which no competitor can say and the buyer is glad to hear.

This is not a retreat from the mission. It is the order the mission gets heard in. The two claims must always travel together, strength and safety in the same breath, because either one alone collapses into a category the company does not win: tough-but-toxic is the commodity, clean-but-weak is the niche. Together they are a position no one else on the shelf holds.

The price, defended

The product will sit at [several times] the commodity price, and the instinct will be to apologize for that. The costing says do not. The material's service life runs 3 to 5 times the commodity standard. A buyer replacing a commodity unit every season pays more across three years than a buyer who buys this once. That arithmetic, cost per use rather than price per unit, is the entire defense of the premium, and the mined reviews show buyers already reaching for it on their own. The large-format flagship at a retail price near 100 dollars holds gross margin in the high fifties of percent. The same material sold as an input earns [a fraction of that, held in the client record]. The premium is not a vanity. It is the point of the launch.

The constraint that designs the system

A launch like this is conventionally carried by a team of ten across brand, content, marketplace operations, advertising, and analytics. This one will be carried by [a handful], alongside their existing jobs. That constraint is not a weakness to manage. It is the design input that decides what gets built: a go-to-market system where AI does the producing and people do the deciding.

Concretely: the brand lives in one governing document precise enough that a machine drafting copy cannot wander from it. Listings, campaign copy, and research are AI-drafted and human-approved. The market's feedback, reviews and search terms and ad results, is mined by AI and read by people weekly. The team's scarce hours go to judgment: what to claim, what to price, what to launch next.

Where AI fits, and where it does not

- **Fits: content production at catalog scale.** Every SKU needs a listing; every listing needs to obey the brand rules. AI drafts, a human gate approves.

- **Fits: market evidence.** Review mining, search-term analysis, and ad performance reads that would take a team weeks, run continuously.
 - **Fits: the advertising loop.** Keywords as experiments, results as evidence, spend reallocated weekly.
 - **Does not fit: the claims themselves.** What the brand promises is a human decision, made once, governed in writing. The certification path runs through an accredited lab, not a language model.
 - **Does not fit: the price.** Set by costing and conviction, defended by arithmetic, never delegated.
-

Risks and constraints we observed

- The premium dies without proof. Every claim must trace to a number, a test, or a certification. The category asserts; this brand must demonstrate.
 - The mission, led with, shrinks the market to a segment. Led from behind, it converts the whole shelf. The order is the risk.
 - The mini end of the catalog may not carry its own shipping economics. Small units earn their place as entry points or they do not earn it at all; the math decides.
 - Consumer volume must not disrupt industrial commitments. The plant's existing book is the company's floor.
 - A two-audience company risks speaking both languages to both audiences. The industrial buyer and the consumer must never see each other's pitch.
-

The signal we leave with

The first move is not a campaign. It is the governing document and the system it governs: brand rules tight enough for AI to produce against, a storefront built under them, and an advertising loop that treats every keyword as an experiment. Before committing to the build, three assumptions need cheap tests: that performance language outsells mission language in this category, that the premium holds when cost per use is shown, and that the catalog's entry sizes can carry their economics. Those tests are where the work goes next. The plan, sized by impact, is the Playbook and Delivery Proposal.

Playbook and Delivery Proposal

The playbook and the delivery proposal are one document because they are one act. The playbook says where AI fits and sizes the moves in order. The delivery proposal scopes the build for the moves you choose to start with. The first earns the second. Nothing past the first move is committed until the first move proves the approach in your business.

Part One: The Playbook

A written read for operators, not a deck. It answers three questions: where AI fits in this launch and where it does not, the highest-leverage moves in sequence, and what it takes to run them.

Where AI fits, and where it does not

It fits the gap between the team you have and the team this launch needs. A consumer launch is carried by content production, market evidence, and an advertising loop, and all three are work AI does well under human governance. It does not fit the claims, the price, or the brand's rules, which are decided once by people and then enforced on everything the machine produces. The evidence is in the Day One Audit. The short version: the material is ready, the machine that sells it does not exist, and the team that would conventionally be that machine does not either. Build the machine.

How to read the roadmap

The first two moves we diagnosed in the opening stretch, and we can size them against the category's evidence and your own costing. The third is half diagnosed: the expansion pattern is proven by design, but each new product earns its own diagnosis before it ships. The last two we saw the shape of and did not diagnose, and we say so rather than dress them up. Honesty about what is proven and what is a candidate is the difference between a roadmap and a sales sheet.

Each move is read across six dimensions: time, accuracy and quality, cost and recovered revenue, growth, employee experience, and risk. The first move earns the right to the next.

The roadmap at a glance

| # | Move | Status | Leverage | Containment | Why it sits here |
|---|--|--|----------|---------------------------------|--|
| 1 | The brand constitution and the governed storefront | Diagnosed, sized | Highest | One channel, one product family | Nothing sells until this exists. Start here. |
| 2 | The advertising evidence loop | Diagnosed, sized | High | One spend line | Needs live listings to feed it. Runs from launch week. |
| 3 | Catalog expansion on the proven system | Pattern proven, each product diagnosed in turn | High | One product family at a time | The system pays for itself the second time it is used. |
| 4 | Off-channel demand to feed marketplace rank | Candidate, not yet diagnosed | Medium | One ad budget | Only after the storefront converts reliably. |
| 5 | New channels | Candidate, not yet diagnosed | Medium | Capped experiments | Channels eat lean teams. Experiments only, until proven. |

Move 1: The brand constitution and the governed storefront (start here)

Write the brand down as law, then build the storefront under it. One governing document carries the positioning, the claims and the proof behind each, the language that is used, and the language that is banned. Strength and safety travel together in every sentence; the mission closes rather than opens; every claim traces to a test or a certification. That document is what makes AI safe to use at scale: the machine drafts every listing, every bullet, every campaign line, and a human approves against the rules. The first product family goes live under it.

- **Time:** the difference between a launch and a stall. Listing production for a full family in days per SKU rather than weeks, because drafting is machine work and only judgment waits on a person. The team runs the launch alongside their jobs, which was the constraint that designed the system.
- **Accuracy and quality:** every public word obeys one document. No claim ships without its proof attached. In a category where nobody proves anything, the listing that cites its certification reads like the only adult on the shelf.
- **Cost and recovered revenue:** carries the launch without the conventional team of ten or the agency bill that substitutes for it. The margin case is the audit's: a flagship unit holding gross margin in the high fifties of percent, on material that earns [a fraction of that] as an industrial input.
- **Growth:** the storefront is the beachhead for everything after it. Each later product inherits the constitution, the content engine, and the channel.

- **Employee experience:** the team's hours go to deciding, not producing. The President runs a launch instead of drowning in one.
 - **Risk:** lowest on the board, and contained. One channel, one family, every word human-approved before it ships. Reversible at the cost of the listings themselves.
-

Move 2: The advertising evidence loop

From launch week, every keyword is an experiment and spend follows evidence. Campaigns are read weekly against hard efficiency targets: an advertising cost of sale at or below 25 percent for the standard catalog, 20 percent for the vehicle lines, no cap on the brand's own name. Discovery campaigns surface the search terms real buyers use; winners graduate to dedicated campaigns; losers are cut without sentiment. A weekly scorecard, twenty minutes, reads revenue, spend, reviews, acquisition cost, and returns.

- **Time:** the analysis that would take a marketplace agency a reporting cycle runs continuously. The team reads conclusions, not spreadsheets.
 - **Accuracy and quality:** spend decisions made on the category's actual search behavior rather than on intuition. The mining already shows the pattern to expect: specific terms beat generic ones, use-case terms beat category terms, and the differentiation terms reach a buyer the commodity cannot.
 - **Cost and recovered revenue:** the loop is self-funding by design. Cutting the losing half of an exploratory spend pays for the discipline that found it.
 - **Growth:** search-term evidence is product strategy. What buyers type is what they want; the loop feeds move 3 its shortlist.
 - **Employee experience:** removes the one job a lean team predictably drops, the weekly grind of campaign hygiene, and leaves the decision: scale, hold, or kill.
 - **Risk:** low and self-limiting. Spend is capped by the targets; a bad week costs a week, never a quarter.
-

Move 3: Catalog expansion on the proven system

The same machinery, run again. The constitution, the content engine, and the advertising loop were built for one product family; the second family ships at a fraction of the first one's effort, and the second category proves the system. Each candidate product earns its slot the same way: the category's reviews are mined for the friction the material can fix, the unit economics are proven at the costing desk, and the differentiating claim is certified before it is made. Candidates the mining has already surfaced run from adjacent vehicle lines to a floor-protection line for the professional trades, where the mined gap is sharp: no incumbent combines durability, grip, easy cleaning, and clean chemistry at a workable price, and none of them proves a claim.

- **Time:** each new family inherits the system. Days to draft a family's listings, not the weeks the first one took to stand up.
 - **Accuracy and quality:** new claims go through the same gate: certified first, claimed second.
 - **Cost and recovered revenue:** the second category targets gross margins [near half of retail] on the main sizes, with the niche sizes accepted leaner where they serve a trade the brand wants. The arithmetic is per-SKU and it is done before launch, not after.
 - **Growth:** this is the move that turns a product into a brand. One family is a listing; a catalog with a shared spine is a company.
 - **Employee experience:** the second launch is the proof the team can feel: the machine they ran once runs again, faster, with the judgment in the same hands.
 - **Risk:** medium, and gated. Each family is its own contained bet, sized by its own mining and its own margins. A family that fails the math does not ship.
-

The later moves: named, not yet diagnosed

- **Move 4: Off-channel demand to feed marketplace rank.** Paid traffic from outside the marketplace, pointed at the storefront, run not for direct return but for the rank and velocity the marketplace rewards, with a tracked share of spend recovered through attribution. The number to beat is the blended acquisition cost. Not diagnosed; it earns a cheap test only after the storefront converts reliably.
- **Move 5: New channels.** A second marketplace and a social commerce channel, as capped experiments. The mining and the category both say the same thing: channels eat lean teams without returning profit until proven. Ninety percent of the effort stays on the channel that works. The number to beat is incremental profit per hour of team attention.

Each of these is a contained bet with its own measurable result. None is committed now. They earn their turn only after the moves ahead of them prove out.

What it takes to run the moves

The discipline matters more than the technology.

- **Write the rules before the words.** The constitution precedes the content. A machine can only be governed by what is written down.
- **Prove before claiming.** Certify the differentiator at an accredited lab before the listing says it. The category asserts; the brand demonstrates.
- **Keep a human on every public word.** AI drafts, a person approves, while trust is earned and after.
- **Spend on evidence.** Every keyword is an experiment with a kill threshold. Sentiment is not a campaign strategy.

- **Expand by arithmetic.** A product ships when its mining and its margins say so, not when enthusiasm does.

The plays that run each canon come from our reusable plays library. The ones selected for this engagement are instantiated in the Charter. ## Who runs it

This can run with your own team, with another firm, or with us. It needs a few clear accountabilities: someone who owns the consumer line and clears the way, someone who holds the brand's rules and approves what ships, a builder who runs the content and advertising machinery, the costing desk that proves every unit's math, and the plant, which already does its part. You have the judgment seats. The machinery and the marketplace craft are the pieces you would bring in.

The recommended first move and the 90-day frame

Start with the constitution and the storefront. Nothing else in this launch can exist until the brand is written down as law and the first family is live under it. The first 90 days: the cheap tests that settle the pitch order and the entry-size economics, the constitution authored and agreed, the first family's listings drafted by machine and approved by people, and the advertising loop live from launch week with its targets set. Prove the system on one family and the operation gains not just a launched product but the machine that launches the next one.

Part Two: The Delivery Proposal

The proposal to build the playbook's first moves: the brand constitution, the governed storefront, and the advertising evidence loop, with catalog expansion to follow on the proven system. Scoped only after the playbook showed what is worth building.

What we understand

A differentiated material with industrial proof and no consumer machine. A team a fraction of the conventional size, carrying a launch alongside their jobs. A pitch that needed inverting before money followed it: tough opens, clean closes, and the two never separate. A premium price defended by cost-per-use arithmetic the category's own buyers already reach for.

What we will build

A go-to-market system, not a campaign. The brand constitution that governs every public word. The storefront and its first product family, every listing AI-drafted and human-approved under the rules. The advertising loop that treats spend as experiments and reads results weekly. The measurement spine: one

scorecard, twenty minutes, every week. Then the same machinery run again on the next family, and the next.

How we will work

Four phases, mapped to the WISER canons. Each phase is independently valuable, priced on its own, and earns the next. The engagement can stop at any phase boundary with value already in hand.

Phase 1: Interrogate

Cheap tests before any production build. Settle the pitch order with the category's own evidence. Prove the unit economics SKU by SKU, including whether the entry sizes carry their shipping.

- Mine the category's reviews at scale for the friction the material fixes and the language buyers use.
- Test mission-led against performance-led messaging where a test is cheap.
- Price the catalog at the costing desk: landed cost, fees, shipping, margin per SKU.
- End of phase: a validated pitch order, a priced catalog, and the wrong instincts ruled out for the cost of a few weeks.

Phase 2: Solve

Author the constitution and build the storefront under it. First family live.

- The brand constitution: positioning, claims with their proof, required language, banned language, the rule that strength and safety never separate.
- The content engine: AI drafts every listing against the constitution; the President approves every word before it ships.
- The storefront live with the first family, the advertising loop running from launch week.
- End of phase: the first family selling under governed content, with baseline metrics to expand against.

Phase 3: Expand

The same machinery, run on the rest of the catalog.

- Adjacent families drafted, approved, and launched on the proven system.
- The second category selected by mining, certified before claimed, priced before built, launched to the trade it serves first.
- The advertising loop scaled across the catalog, winners graduated, losers cut.
- End of phase: the full catalog live, ad spend inside its targets, the second category proving the system transfers.

Phase 4: Refine

Govern the system as it grows.

- A standing reconciliation review: every live listing audited against the constitution on a cadence, because words drift the way models do.
 - The weekly scorecard as the operating rhythm: revenue, spend, reviews, acquisition cost, returns.
 - Content autonomy grown tier by tier as the gate's catch rate proves what the machine can be trusted to draft.
 - End of phase: a governed brand machine the team operates, with the judgment where it started, in human hands.
-

What we need from you

- The President as the approving gate on every public word, and the owner's mandate behind the inverted pitch.
 - The costing desk's time to prove the unit math, SKU by SKU.
 - The plant's specifications and test data, because every claim must trace to proof.
 - A weekly twenty minutes, kept.
-

Infrastructure

You provide the marketplace seller account, the product, and the certifications as they are earned. We provide the build, the AI machinery, the brand and marketplace craft, and the implementation.

Who is working on this

A senior practitioner who leads the engagement and owns the system design with your team, and the brand and marketplace build. Your people fill the judgment seats: the line, the costing, the claims. Small team, close to the work.

Investment

Phased. Each phase is priced on its own so the engagement can stop at any phase boundary with value already delivered. The fee basis and amounts are held in the private client record. We did not fabricate figures for this anonymized record.

Charter

What a Charter is

Not a project plan. Not a requirements document that executes once and collects dust. A Charter is the memory that survives the chaos. Its value is the decision log: when someone asks a year later why the pitch leads with toughness instead of the mission, or why the second category beat two louder candidates, the answer is here, with the alternatives that were weighed and the evidence that settled it. The Architect keeps it current, same-day.

Metadata

| Field | Value |
|----------------|---|
| Project | Consumer line launch: a governed go-to-market system. A second project followed: the connected go-to-market. |
| Client | The mid-size manufacturer (anonymized) |
| Charter Keeper | The Architect |
| Dates | Held in the private client record; relative markers used here |
| Current canon | Refine. The launch system and the connected go-to-market are live and governed; the second category is at pre-launch. |
| Version | Running state |

Positions

The work was held together by clear accountabilities, not an org chart.

| Position | Who held it | Tension owned |
|--------------|--|--|
| Sponsor | The owner | Authority. Owned the why and cleared the way. |
| Guide | First Strategy senior practitioner | Translation. Carried the method and kept the Charter honest. |
| Architect | First Strategy | Curiosity and stewardship. System design and Charter Keeper. |
| Sage | Plant leadership | Context. The material's truth and what the line can actually run. |
| Scout | Trade voices recruited for the second category | Empathy. Validated whether the trade would actually adopt. |
| Builder | First Strategy | Execution. The content engine, the storefront, the advertising loop. |
| Finance lead | Client finance | Safety. Proved every unit's math and watched the spend. |
| Brand gate | The President | Integrity. Approved every public word before it shipped. |

On a small team one person can hold several Positions. As the system proved reliable, a Position could be augmented by an AI agent inside documented constraints, with the human shifting from doing to directing and reviewing.

Objectives and constraints

The build specification: what the project set out to do and the lines it would not cross.

Scope

In scope: the consumer go-to-market system, from brand rules to live storefront to advertising loop, and the catalog launched through it. The second project widened the scope to the consumer line's operating spine: product data, storefront publication, fulfillment, and measurement. Out of scope throughout: the industrial business, the plant, and the material itself. The launch borrows the plant's proof; it does not touch its operation.

Objective and success criteria

Launch the consumer line under governed AI content, hold the premium price, and prove the economics SKU by SKU.

| Measure | Baseline | Target | Result |
|------------------------|----------------------|---|---|
| Consumer storefront | None | First family live under the constitution | 19 SKUs live across three product families |
| Advertising efficiency | No consumer spend | Cost of sale at or below 25 percent standard, 20 percent on the vehicle lines | Top terms below 15 percent, the best at 2.6; the working band 15 to 25 |
| Unit margin | Input economics | Gross margin near half of retail on consumer units | High fifties of percent on the flagship; most second-category sizes near half |
| Price position | The commodity anchor | Hold the premium on cost-per-use arithmetic | Held. No discounting toward the commodity |

Constraints

- Every public claim traces to a test, a certification, or the company's own data. No claim ships unproven.
- Strength and safety never separate. The hard line of the constitution.
- The industrial buyer and the consumer never see each other's pitch.
- The team stays lean. The system absorbs scale; headcount does not.
- Consumer volume never disrupts the plant's industrial commitments.

Architecture and human-in-the-loop design

One governing document, the brand constitution, sits at the top: positioning, the claims and the proof behind each, required language, banned language, and the pitch order. Below it, an AI drafting layer produces the volume work: listings, bullets, enhanced content, campaign copy, search-term reads, review mining. Below that, one human gate: the President approves every public word against the constitution before it ships. The market's feedback, reviews, search terms, and the weekly scorecard, is mined by machine and read by people, and what it teaches flows back into the constitution as versioned updates, never as quiet edits.

During the launch the gate held everything: every listing, every bullet, every campaign line. Misses were logged with the pattern that caused them and the banned-language list grew from the log. The gate's grip loosened only later, under the tiers in the Hierarchy of Agency, never before the evidence supported it. A human stays accountable for every public claim.

Current state at the start

Carried from the Day One Audit. A material stronger than the commodity standard and free of the chemistry the category depends on, certified where the category merely claims. No consumer brand, no storefront, no advertising practice, no retail costing. A launch team a fraction of the conventional size, carrying the line alongside their jobs. A pitch instinct, mission first, that the category's evidence said to invert.

Decision log

The decisions that shaped the build, each with the alternatives weighed and the evidence that settled it. This is the part of the Charter that answers "why did we do it this way."

| When | Decision | Alternatives rejected | Rationale | Evidence |
|-------------------|---|--|--|---|
| Interrogate, wk 1 | Invert the pitch: tough opens, clean closes | Lead with the mission and the founding story | The category buys on failure pain; the mission converts a segment, not the shelf | Review mining: failure complaints dominate the negative reviews; mission praise is thin and rarely converts |
| Interrogate, wk 2 | Strength and safety never separate, in any sentence | Lead either alone | Alone, each collapses into a category the brand loses: tough-but-toxic is the commodity, clean-but-weak is the niche | Positioning read of the mined shelf |
| Interrogate, wk 2 | Hold the premium and sell cost per use | Price toward the commodity | The 3-to-5x service life makes the premium cheaper across three years, and buyers already reach for that math | The costing desk's arithmetic; replacement complaints in the mining |
| Interrogate, wk 3 | Keep the entry size despite thin shipping economics | Cut the small units | They recruit first-time buyers into the brand; watched as a metric, not defended as a darling | SKU-level costing; flagged for quarterly review |
| Solve, wk 1 | One constitution governs every word; AI drafts, the President approves | Ad-hoc copy per listing; an agency retainer | The lean team is the design input; consistency at catalog scale is impossible by hand and unaffordable by agency | The diagnostic's constraint, costed |
| Solve, wk 2 | Separate the consumer presence from the industrial one | One storefront for both audiences | Two audiences, two languages; each pitch poisons the other | Brand analysis; the industrial buyer's vocabulary in consumer mock-ups read as catalog copy |
| Launch | Advertising targets set as law: 25 percent cost of sale standard, 20 on the vehicle lines, no cap on the brand's own name | Spend for rank at any cost | Self-limiting spend; evidence over sentiment | Scorecard design |
| Expand | Graduate discovery-campaign winners to dedicated campaigns | Leave everything in automatic discovery | Discovery finds terms the manual plan missed; | Search-term reads: discovery surfaced |

| When | Decision | Alternatives rejected | Rationale | Evidence |
|-------------|--|---|--|---|
| | | | winners deserve their own budget and control | converting terms no one had guessed |
| Expand | Choose the floor-protection line as the second category | A premium outdoor bedding line at a higher price band; a grill cover line | The sharpest mined gap, the same material, the simplest manufacturing; the other two wait their turn | Category mining; plant complexity assessment |
| Expand | Certify the differentiator at an accredited lab before claiming it | Claim it the way the category does | Nobody on the shelf proves anything; proof is the position | The lab certification, in hand before the first listing was drafted |
| Refine | Stand up a standing reconciliation review of every live listing | Trust the launch-time listings | The catalog had drifted from the constitution as it multiplied | The reconciliation found an unproven claim live and the core claim phrased three ways |
| Project two | Rebuild the consumer go-to-market as one connected motion | Hire data entry; live with the spreadsheets | Rekeying scales with the catalog; headcount patches a structural problem | Every new SKU multiplied hand updates across three systems |
| Project two | One source of truth that publishes to the channel | Keep editing in the storefront console | Edits made at the channel drift from the record; the reconciliation lesson, operationalized | Channel details now publish themselves |
| Project two | Build the analytics tool over the storefront | Buy a generic dashboard | The decisions it feeds are the brand's own: spend, price, assortment | Ad spend down, average order value up |
| Project two | Codify senior judgment into the systems | Keep it in heads | Just-in-time inventory needs rules a system can run; heads do not scale and do not stay | Inventory holding fell 80 percent; fulfillment dropped two days |

The decision and experiment record

The supporting narrative behind the log. The project ran the full WISER method. Witness had already found where AI fit; the project picked up at Interrogate and ran through Refine.

Interrogate

The opening diagnostic was spent letting the category argue with the company's instincts. The review mining did most of the talking. The complaint pile was a portrait of a shelf that fails its buyers: products that tear in a season, hardware that pulls out, buyers on their third replacement writing the cost-per-use arithmetic themselves in one-star reviews. The mission language the company planned to lead with appeared in the praise, but thinly, and almost never as the stated reason to buy. The pitch inverted on that evidence, and the owner, whose company exists because of the mission, accepted the inversion because the evidence was the category's own customers. The mission did not leave the pitch. It moved to the close, where it lands as the reason to feel good about the choice the toughness already won.

The same stretch settled the price. The instinct under pressure would have been to apologize for the premium. The costing desk and the mining agreed it needed no apology: a service life of 3 to 5 times the commodity standard makes the premium the cheaper product across three years. The entry size nearly died in the same review, its shipping economics too thin to defend on margin alone. It survived as a recruitment cost, watched quarterly, named in the log so nobody later mistakes it for an oversight.

Solve

The constitution was authored before any content existed: the positioning, the claims with their proof attached, the required language, the banned language, and the two rules that do not bend. Strength and safety in the same breath, always. Nothing claimed that is not proven. Then the machine was turned on under it. The AI drafted the first family's listings, titles, bullets, and enhanced content, and the President read every word against the rules before anything shipped. The gate earned its keep early: drafts kept reaching for the easy vocabulary, the category's empty superlatives and the mission-first framing the constitution had banned, and each catch went into the log and tightened the banned list. The machine got measurably better at sounding like the brand because the brand was written down.

The first family went live with the advertising loop running from launch week. The early reads taught the pattern the rest of the engagement would confirm: specific beats generic, the use-case term beats the category term, and the chemistry-free buyer, a smaller pool, converts with rare efficiency when reached directly.

Expand

The vehicle lines shipped next on the same machinery, in a fraction of the first family's effort, which was the system proving itself. Then came the decision that tested the discipline: what the second category would be. Two candidates were louder, a premium outdoor bedding line at an attractive price band and a grill cover line with an easy story. The mining picked the quiet one: floor protection for the professional trades, where no incumbent combines durability, grip, easy cleaning, and clean chemistry, where the trade's reviews complain about odor and slip in the same breath, and where not one competitor proves a claim. Same material, simplest manufacturing, sharpest gap. The differentiating claim went to an accredited lab before it went into a listing, because in a category that asserts, the brand that

demonstrates owns the shelf's trust. The launch plan for the category goes to the trade first, seeded with working professionals whose reviews carry weight, with targets set for its first 90 days.

Refine

Live and selling was not the same as governed. The constitution had been updated as the engagement learned, and the question nobody had asked was whether the live catalog still obeyed it. A reconciliation review audited every live listing against the current rules. It found drift. The details are in the drift record below; the lesson is the rule the engagement keeps: words drift the way models drift, and they get audited the same way.

The second project: the connected go-to-market

The launch succeeded, and the success built the next bottleneck. The catalog's growth multiplied a cost that had been tolerable at launch: the consumer line's data lived in three places, the product record in [a master spreadsheet], the listings in the storefront's own console, and fulfillment in its own system, and every change meant rekeying the same information by hand into each. Senior judgment, the product and costing logic, the fulfillment calls, the read on what sells where and when, sat in a few heads. Inventory cycles ran long because the operation could not risk anything shorter.

The instinct under that load is to hire. The decision, logged above, was to connect instead: rekeying scales with the catalog, and headcount patches a structural problem. The second project rebuilt the consumer go-to-market as one connected motion, run inside the commercial leadership team. The product record became the single source of truth, and the channel now reads from it: channel details publish themselves, and a product description adapts to market conditions in hours instead of weeks. Fulfillment was connected to the same spine. The senior judgment was codified into the systems, the costing logic, the fulfillment rules, the channel patterns, so the operation no longer depends on the heads that built it. And an analytics tool was built over the storefront to feed the brand's own decisions: it pulled ad spend down and average order value up.

The operational results: fulfillment dropped by two days. Inventory holding fell by 80 percent, and the operation now runs effectively just-in-time, on cycles it could not have risked when judgment lived in heads and data lived in three places. The manual rekeying is gone.

Hierarchy of Agency

Three tiers of human oversight on what the machine produces, by risk. The President's approval is the gate; the tier governs how much of the gate's attention each content type gets.

| Tier | Oversight | Applies to |
|-----------------|--|--|
| 1: Light review | A pass for claim integrity and banned language; structural reuse of approved copy is trusted | Size and color variants of approved listings; campaign copy reusing approved claims verbatim |
| 2: Full review | Every word read against the constitution | New listings within an existing family; search-term reads turned into campaign changes |
| 3: Human-led | The machine assists; a person authors and decides | New categories, new claims, pricing, anything touching certification |

A content type moves to a lighter tier only on evidence: a full review cycle in which the gate's catch rate on that type stays near zero. The variants of approved copy earned Tier 1 that way. Nothing earns it by decree. If a tier drifts, it falls back to heavier review. A human is accountable for every public claim, at every tier.

Risk register

| Risk | Mitigation | Status |
|---|--|---|
| The premium dies without proof | Every claim traces to a test or certification; the lab result in hand before the claim ships | Held; the premium has not been discounted |
| The mission shrinks the market | Pitch order enforced by the constitution: tough opens, clean closes | Held; conversion led by performance terms |
| AI copy wanders off-brand | The constitution, the gate on every word, the banned-language log | Active control; catch rate falling by tier |
| Entry-size economics | Watched quarterly as a recruitment cost, not defended as a darling | Open; named in the decision log |
| Channel sprawl eats the lean team | Effort capped at the proven channel; experiments bounded | Held; no new channel has earned core status |
| Listing drift from the constitution | Standing reconciliation review on a cadence | Realized once; caught; cadence now standing |
| Manual rekeying scales with the catalog | The second project: one connected motion, a single source publishing to the channel | Realized as the catalog grew; rebuilt; the rekeying is gone |
| Senior judgment in a few heads | Codified into the systems: costing logic, fulfillment rules, channel patterns | Resolved in the second project; the resilience risk retired |

Drift and incident record

After the constitution was updated with what the engagement had learned, a reconciliation review audited all 19 live SKUs against the current rules. The catalog had drifted while everyone watched the dashboards: a strength multiple was live in one listing that traced to no documented proof point; the core lifespan claim appeared phrased three different ways across the family; the newest messaging angles existed in the constitution but not in the listings; and the vehicle lines were selling on the family's generic pitch instead of their own. Nothing was failing in the metrics. The words had simply wandered.

The response:

| Action | Detail |
|------------------------------|---|
| Retire the unproven claim | The undocumented strength multiple removed until the plant's data proves it or it dies |
| Standardize the proof points | One canonical phrasing per claim, written into the constitution, reused verbatim |
| Integrate the updates | The constitution's newest angles pushed across the live catalog |
| Differentiate the lines | The vehicle lines repositioned on their own use cases |
| Standing cadence | Reconciliation review on a recurring cadence, every live listing against the current constitution |

The lesson logged: a brand drifts the way a model drifts, quietly and in the aggregate's blind spots. Audit the words like outputs, on a cadence, against the current rules.

Evolution history

How the oversight posture changed over time, and why.

| When | Change | Trigger |
|----------------------------|---|---|
| Launch | The gate reads every word of every listing | Trust not yet earned |
| After the first family | Variants of approved copy moved to light review | A full cycle with the catch rate near zero on variants |
| After the reconciliation | Standing listing audit on a cadence; one canonical phrasing per claim | The drift the reconciliation caught |
| Second category pre-launch | New-category content held at Tier 3, human-led | New claims, new trade, no track record yet |
| Project two | The operating spine connected; inventory run just-in-time; senior judgment codified | The catalog's growth made hand updates across three systems untenable |

Current status and what transfers

The system is live and operating. The storefront runs at 19 SKUs across three product families. Advertising runs inside its targets, with the top terms below 15 percent cost of sale and the best at 2.6. The premium has held. The consumer go-to-market runs as one connected motion: channel details publish themselves, fulfillment runs two days faster, inventory holds 80 percent less and runs effectively just-in-time, and the senior judgment that made it work lives in the systems, not in a few heads. The second category is at pre-launch: certified, priced, its trade-first seeding plan set, its first-90-day targets written down. The weekly scorecard, twenty minutes, is the operating rhythm.

This engagement continues by design. We run the go-to-market machinery day to day; the client runs the plant, the product, and the judgment seats. What has already transferred is the decision discipline: the constitution is theirs, the decision log is theirs, and every claim, price, and product call is made by their people on evidence the system surfaces. If the day comes to bring the machinery in-house, the system is documented to its bones and the Charter is the manual.

Outcomes

- A consumer line live from a standing start: 19 SKUs across three product families, under one governed brand.
- Advertising inside its targets: top terms below 15 percent cost of sale, the best at 2.6 percent, the working band 15 to 25.
- The premium held. Flagship gross margin in the high fifties of percent, on material that earned input margins before.
- The go-to-market rebuilt as one connected motion: the manual rekeying gone, product descriptions adapting in hours instead of weeks, channel details publishing themselves.

- Fulfillment down two days. Inventory holding down 80 percent, effectively just-in-time.
- The analytics tool pulling ad spend down and average order value up.
- Senior judgment codified: product and costing logic, fulfillment rules, and channel patterns live in the systems, not in a few heads.
- The second category certified before it was claimed, selected by evidence over enthusiasm.
- The team that runs it is still [a handful]. The capacity came from the system, not from headcount or an agency stack.

Plays

The WISER plays this engagement ran, instantiated with the client's specifics. This is the index and what each produced. The high-value plays are held as standalone documents; the rest were applied inline in this Charter. | Canon | Play | What it produced | Source | |-----|-----|-----|-----| | Witness | Friction Mapping | The category friction map, mined from the shelf's own reviews | Standalone play | | Witness | Documenting Current State | The exists-versus-needed read of the launch | Inline in the Day One Audit | | Interrogate | Assumption Auditing | The register of instincts tested, the eco-first pitch above all | Standalone play | | Interrogate | Experiment Selection, Logging | The messaging, pricing, and keyword experiments and what each settled | Standalone play | | Solve | Human-in-the-Loop Design | The constitution-and-gate content system | Standalone play | | Solve | Quality Objective Setting, Value Validation | The launch criteria and the SKU-level economics | Inline above | | Expand | Expansion Sequencing, Context Fit | The family-by-family rollout and the second-category selection | Inline above | | Refine | Drift Monitoring, Incident Response | The reconciliation review and the drift fix | Standalone play | | Refine | Hierarchy of Agency Design, Graduation | The content tiers and the evidence that moves a type between them | Inline above |

The first launch is built, and it will not be the last. The operation now has a catalog of candidates for the same treatment: the bedding line, the grill covers, the channels not yet earned. The difference is that the machine that launches them exists, and the people who govern it know how it runs.

The plays

The WISER plays this engagement ran, instantiated with the client's specifics, ordered by canon.

WITNESS PLAY

Category Friction Map

Witness play, instantiated for the mid-size manufacturing engagement. Purpose: locate and quantify where the work breaks. This engagement was a market entry, so the floor we mapped was the category's: [several thousand] customer reviews across the top sellers on the shelf the product would enter, mined for what buyers complain about, what they praise, and what they say they would pay to fix.

The friction, at a glance

Failure within a season

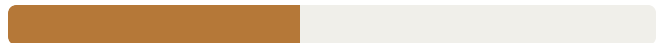
tearing, hardware
pulling out, UV rot



[~half] of negative reviews

Repeat replacement

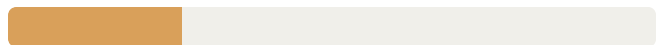
buyers on their third
unit doing the math



[~a quarter]

Chemical smell and safety worry

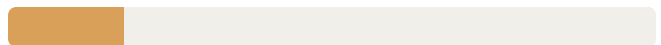
odor, residue, kids,
pets, soil



[a meaningful minority]

Claims without proof

every listing says heavy
duty, none demonstrates
it



[the remainder]

The friction table

| Friction point | Share of negative reviews | What buyers say | What it tells us |
|---------------------------------|---------------------------|--|---|
| Failure within a season | [roughly half] | Tearing, hardware pulling out, UV rot, "lasted one winter" | The category's core promise is broken. Durability is the open |
| Repeat replacement | [a quarter] | "Third one in two years," cost-per-use arithmetic written by the buyers themselves | The premium's defense already lives in the customer's own words |
| Chemical smell and safety worry | [a meaningful minority] | Odor on unboxing, residue, worry about kids, pets, and soil | Safety converts as reassurance, not as the reason to buy |
| Claims without proof | [the remainder] | Skepticism of identical "heavy duty" claims across the shelf | Certification is open ground. Nobody on the shelf proves anything |

The same mining ran later on the second category's shelf, where it found the gap that picked the product: no incumbent combines durability, grip, easy cleaning, and clean chemistry, and the trade's reviews complain about odor and slip in the same breath.

The root-cause read

The shelf is a commodity trap: every product claims the same two words, competes on price, and fails the same way. The friction is not that buyers cannot find a cheap product. It is that the cheap product costs more across three years and the shelf gives them no alternative that proves it is one. A material that is genuinely stronger and genuinely cleaner does not need to invent a pitch. The category's own one-star reviews already wrote it: open with the failure they know, prove the lifespan, close with the chemistry no one else can claim.

INTERROGATE PLAY

Assumption Register

Interrogate play, instantiated for the mid-size manufacturing engagement. Purpose: surface the assumptions the launch was about to spend money on, including the company's own founding instinct, and test each one cheaply before it got expensive.

The verdicts, at a glance

| | | | | | |
|----------|---|-----------|--|----------|--|
| REFRAMED | <p>The mission is the pitch</p> <p>The mission closes; toughness opens. The order inverted on the category's own evidence.</p> | KILLED | <p>The buyer is the eco buyer</p> <p>Three buyer archetypes share durability and safety values. Ideology is a segment, not the shelf.</p> | REFRAMED | <p>The premium is the obstacle</p> <p>The premium holds when cost per use is shown. The obstacle was the missing arithmetic, not the price.</p> |
| KILLED | <p>More channels, more revenue</p> <p>Channels eat lean teams. One channel proven deep beats three run shallow.</p> | CONFIRMED | <p>Proof differentiates</p> <p>Nobody on the shelf proves a claim. Certification before claiming became a standing rule.</p> | | |

The register

| # | Assumption | Source | Cheap test | Verdict |
|---|--|--|---|---|
| 1 | The environmental mission is the selling point | The founding story; the owner's conviction | Mine the category's reviews for stated purchase reasons; test mission-led against performance-led framing | Reframed. The mission converts as the close, not the open. Tough first, clean as the reason to feel good about it |
| 2 | The consumer buyer is the eco buyer | Inherited from the mission | Profile the actual buyers behind the category's reviews | Killed. Three archetypes emerged, a practical homeowner, an outdoor and agricultural professional, a contractor, sharing durability and safety values, not ideology |
| 3 | The premium price is the obstacle | Fear of the commodity anchor | The costing desk's cost-per-use arithmetic against the mined replacement complaints | Reframed. Buyers already write the arithmetic in their reviews. The obstacle was that no listing had ever shown it |
| 4 | More channels mean more revenue | Launch enthusiasm | Size the team attention each channel demands against its proven return | Killed for now. Effort capped at the proven channel; others are bounded experiments until they earn core status |
| 5 | Third-party proof will differentiate | Our read of the shelf | Survey the shelf for any certified claim | Confirmed. None found. Certify before claiming became a standing rule of the constitution |

Why assumption 1 was tested at all

It was the company's reason for existing, which is exactly why it had to be tested. A founding conviction is the most expensive kind of assumption: nobody inside the company can see past it, and every dollar of the launch was about to be spent on it. The test did not disprove the mission. It found the order the mission gets heard in, which saved the launch from leading with the one story the shelf's buyers do not buy on.

INTERROGATE PLAY

Experiment Log

Interrogate play, instantiated for the mid-size manufacturing engagement. Includes the experiment selection for each round. Purpose: record each experiment, its result, and what it changed. The launch experiments ran during the opening diagnostic; the advertising experiments ran continuously from launch week, because in this engagement the ad account is a standing laboratory.

The launch experiments, at a glance

① **Mission-led vs performance-led pitch**

REFRAMED

Performance opens, mission closes. The pitch order inverted.

② **Cost-per-use price defense**

CONFIRMED

The premium holds when the three-year arithmetic is shown.

③ **Entry size on its own margin**

KILLED

Shipping eats the small unit. Reframed as a recruitment cost, watched quarterly.

④ **Specific terms vs generic terms**

CONFIRMED

Size-specific and use-case terms convert; generic category terms burn spend.

⑤ **The chemistry-free buyer, reached directly**

PROMISING

A small pool converting at single-digit cost of sale. A niche to serve, not a shelf to bet on.

⑥ **Discovery campaigns as instruments**

CONFIRMED

Automatic discovery surfaced converting terms nobody guessed. Winners graduate to dedicated campaigns.

The pre-launch experiments

| # | Experiment | Result | What it changed |
|---|---|--|---|
| 1 | Mission-led against performance-led framing, tested against the mined reviews and where a live test was cheap | Performance language carries the open; mission language lands at the close | The pitch order, and with it the constitution's first rule of sequence |
| 2 | The cost-per-use defense of the premium | The arithmetic holds: a 3-to-5x service life makes the premium cheaper across three years, and buyers already reach for the math | The price held without apology; the arithmetic written into the listings |
| 3 | The entry size's unit economics | Shipping costs eat the small unit's margin | Kept deliberately as a recruitment cost, named in the decision log, watched quarterly |

The advertising experiments

The standing laboratory. Targets set as law, 25 percent cost of sale standard and 20 on the vehicle lines, every keyword an experiment against them.

| # | Experiment | Result | What it changed |
|---|---|---|--|
| 4 | Specific terms against generic category terms | Size-specific and use-case terms convert well inside targets, the account's best at 2.6 percent cost of sale; generic terms run hot | Spend reallocated to specificity; generic terms cut or negative-matched |
| 5 | The chemistry-free search term | Converts at a single-digit cost of sale on modest volume | A niche served with its own terms, never mistaken for the core pitch |
| 6 | Automatic discovery campaigns | Surfaced converting search terms the manual plan had not guessed | A standing graduation path: discovery finds, winners get dedicated campaigns and budgets |

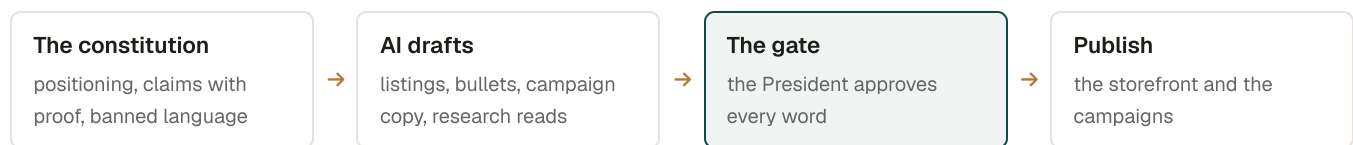
What the log settled

Each experiment narrowed the path by testing rather than asserting. The pre-launch round saved the pitch from the founding instinct and the price from the discount reflex. The advertising round turned the ad account into the engagement's permanent evidence source: what buyers type is what they want, and the search-term record now feeds product selection as much as it feeds spend.

Human-in-the-Loop Design

Solve play, instantiated for the mid-size manufacturing engagement. Purpose: define who reviews AI output, how, and what gets logged. In this engagement the AI's output is public language, every listing and campaign line the brand ships, so the loop guards the brand the way a billing reviewer guards an invoice.

The flow



↻ Every catch is logged with its pattern. The banned-language list grows from the log, and the market's feedback flows back into the constitution as versioned updates.

The design

| Element | Design |
|---------------------------|--|
| Who reviews | The President, as the single brand gate. One person, one standard, no committee |
| What they check | Claim integrity first: nothing ships that does not trace to a test, a certification, or the company's data. Then sequence: strength and safety together, tough opens, clean closes. Then language: nothing from the banned list, the canonical phrasing for each proof point |
| What the AI drafts | Every listing, title, bullet, and block of enhanced content; campaign copy; review mining and search-term reads |
| What the AI never decides | The claims themselves, the price, what launches next, anything touching certification |
| What gets logged | Every catch, with the pattern that caused it: the reach for empty superlatives, the mission-first drift, the unproven flourish |
| Where the log goes | Into the constitution. Each pattern becomes a rule, so the machine stops making the mistake category-wide instead of listing by listing |

How the grip loosens

The gate read every word at launch. Content types earn lighter review only on evidence, under the tiers in the Charter's Hierarchy of Agency: a full cycle in which the gate's catch rate on that type stays near zero. Size and color variants of approved copy earned it. New claims never will, by design.

Why this matters

A lean team cannot produce a catalog's content by hand, and an ungoverned machine cannot be trusted with a brand's claims. The constitution makes the machine governable; the gate makes it accountable. The combination is what let [a handful of] people run a launch that conventionally takes ten, without a single unproven claim shipping. The one that nearly did is in the drift record, which is why the reconciliation review now exists.

REFINE PLAY

Drift Monitoring

Refine play, instantiated for the mid-size manufacturing engagement. Includes the incident response for the drift this engagement caught. Purpose: watch for drift, including drift that looks stable in aggregate but is failing underneath. In this engagement the thing that drifted was not a model's accuracy. It was the brand's own words.

The incident

- The catalog goes live under the constitution**

Every listing gate-approved at launch. The dashboards watch sales, spend, and reviews.
- The words wander, quietly**

The catalog multiplies, the constitution evolves, and the live listings stay where they were approved. Nothing in the metrics flags it.
- Caught by reconciliation, not by the dashboards**

An audit of all 19 live SKUs against the current constitution finds an unproven claim live, the core claim phrased three ways, and the newest angles missing.
- The fix, applied catalog-wide**

The unproven claim retired. One canonical phrasing per proof point. The vehicle lines repositioned on their own use cases.
- A standing cadence, with the lesson logged**

Words drift the way models drift. Audit the words like outputs, on a cadence, against the current rules.

After the constitution was updated with what the engagement had learned, a reconciliation review audited every live listing against the current rules. The catalog had drifted while everyone watched the dashboards. A strength multiple was live in one listing that traced to no documented proof point. The core lifespan claim appeared phrased three different ways across the family. The constitution's newest messaging angles existed on paper and nowhere on the shelf. The vehicle lines were selling on the family's generic pitch instead of their own use cases. Nothing was failing in the metrics. The words had simply wandered from the law.

It was caught by reconciliation, not by the dashboards, because no sales metric measures whether a listing still obeys the brand.

The fix

| Action | Detail |
|------------------------------|---|
| Retire the unproven claim | Removed until the plant's data proves it or it dies. The constitution's first rule has no exceptions clause |
| Standardize the proof points | One canonical phrasing per claim, written into the constitution, reused verbatim everywhere |
| Integrate the updates | The newest messaging angles pushed across the live catalog |
| Differentiate the lines | The vehicle lines repositioned on their own use cases instead of the family's generic pitch |

Standing monitoring

- A reconciliation review on a recurring cadence: every live listing against the current constitution.
- The weekly scorecard: revenue, spend, reviews, acquisition cost, returns, read in twenty minutes.
- Advertising drift watched against the standing targets; any term running past them is cut or repriced at the weekly read.
- Any new constitution version triggers a reconciliation pass before the version is considered shipped.

Incident response pattern

Catch by reconciliation, not by the sales dashboards. Contain by retiring the claim, not defending it. Fix the rule, not just the listing, so the class of drift dies with the instance. Document the cause, detection, and fix in the Charter so the next person does not relearn it.