SAMPLE REPORT · Illustrative $2,399 Quick-Turn NES Diagnostic rendered on an anonymised premium DTC consumer brand. Customer excerpts on later pages are real public-review text with brand name redacted.

Executive diagnosis

One-page senior-leadership summary. The full diagnostic across both signal layers begins on page 2.

ONE-LINE DIAGNOSIS

The brand reads coherent on its website but splits 34 points across customer-language surfaces. Operational consistency drift, not strategic confusion.

MAIN FINDING

Trustpilot reviewers describe a high-craft, well-serviced premium brand (NES +33.3). Reddit and Amazon reviewers describe variable run-to-run quality, customer-service inconsistency, and warranty-interpretation drift (NES −1.0 and −11.1 respectively). Same brand, two operational realities. The 34-point spread itself is the diagnostic finding worth surfacing.

WHY IT MATTERS

A channel spread of this magnitude points to operational workflow drift, not strategic positioning failure. It is addressable on a 30–90 day cycle without repositioning the brand. Untreated, the public-review averages on Trustpilot, Amazon, and the App Store calcify against the brand at a non-linear rate.

TOP 3 RISKS

- Premium-piece quality drift on the most narrative-dense products
- Customer-service variance between owner-led and team-led handling
- Refund and shipping-tax opacity for international buyers

TOP 3 RECOMMENDED ACTIONS

- Audit warranty-claim decision tree (last 90 days)
- Re-route owner-led service interactions to team-led
- Surface refund policy above the product-page fold

WHAT INTERNAL DATA SHOULD BE CHECKED NEXT

Return-reason coding by SKU and batch number. Ticket-resolution time and CSAT segmented by handler tier. Repeat-purchase rate by first-time vs returning vs multi-piece-collector cohort. Defect-rate by premium vs standard product tier. These four data sets would confirm or refute the directional read in this report at significantly higher confidence.

SAMPLE REPORT · Illustrative $2,399 Quick-Turn NES Diagnostic rendered on an anonymised premium DTC consumer brand.

PART 1 OF 2 · WEBSITE-SIGNAL LAYER

Sample Brand (anonymised)

Quick-Turn NES Diagnostic · Powered by the NES framework. Part 1 mirrors the $29 Full Brand Consistency Review (website signal). Part 2 (pages 7-13) adds Review-Inferred NES from customer-language aggregation across review portals.

Headline scores

BRAND CONSISTENCY	MESSAGE CLARITY	IDENTITY CONFUSION	TRUST LEAKAGE	CONFIDENCE
71	78	32	Low-Med	82%
/ 100 · HIGHER = BETTER	/ 100 · CLEARER	/ 100 · LOWER = BETTER	7 / 20	SIGNAL STRENGTH

Bottom line

> **Strong category-creator brand voice and a coherent functional-beverage promise on the website. The customer-language layer (Part 2) reveals a 34-point channel spread between the satisfied-customer surface and the daily-community surface.**

Key observations

1. Hero message is unambiguous and category-defining; the same promise repeats across product, about, and FAQ pages without drift.
2. Product-detail pages carry strong narrative but thinner proof density than the homepage; ingredient and benefit claims could be more tightly tied to source citations.
3. Identity overlap with the broader functional-beverage cohort creates a small but consistent confusion signal in voice and visual treatment.

10-component breakdown

Per Appendix A of the NES working paper. Each component scored 0–20 from public website signal.

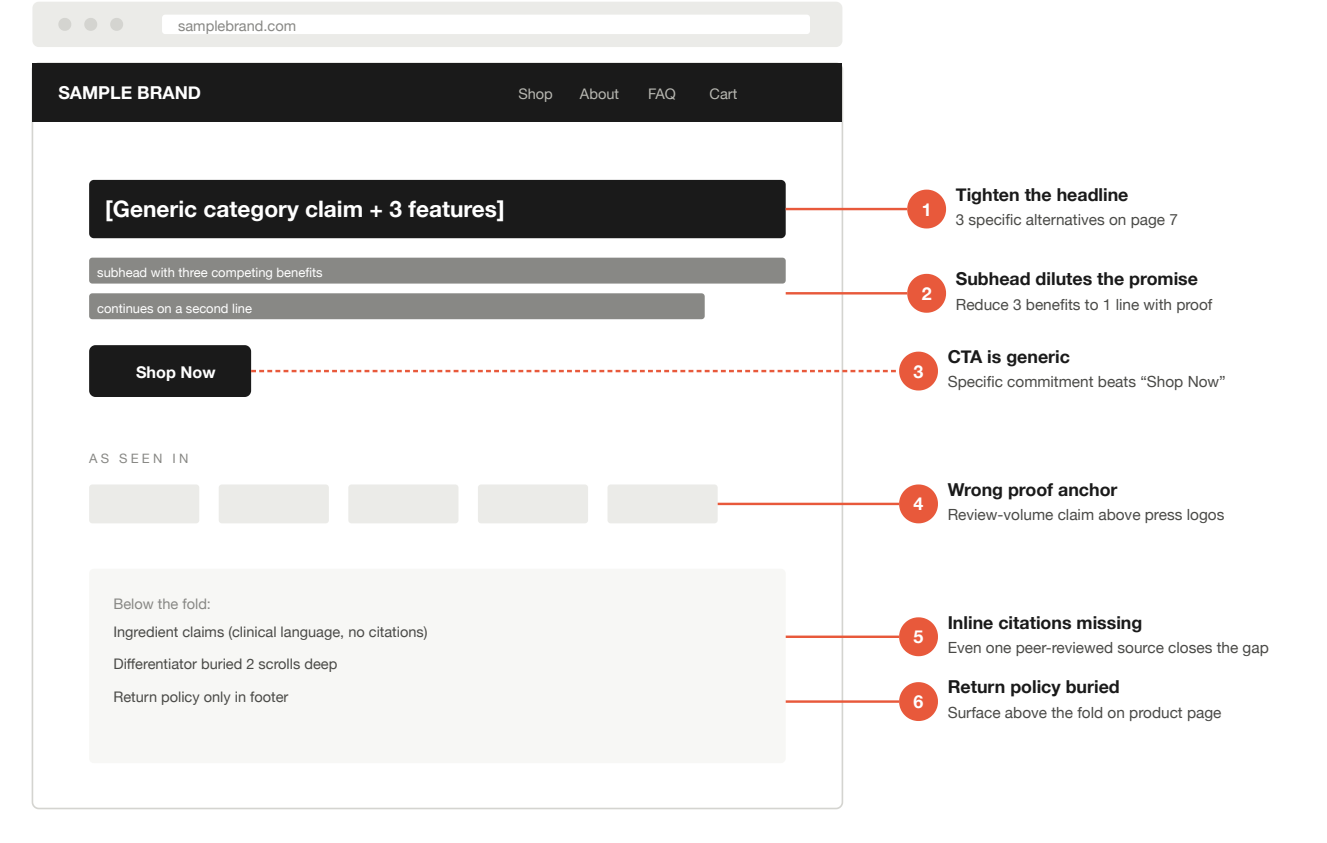
+M	Message clarity	16/20	−IC	Identity confusion	5/20
+T	Trust signals	13/20	−TD	Trust drift	4/20
+C	Customer-promise consistency	15/20	−AC	Audience confusion	3/20
+R	Reputation evidence	12/20	−RR	Risk signals	2/20
+V	Voice coherence	15/20	−OD	Operational dissonance	4/20

HOW TO READ THIS

The strongest scores carry the brand surface; the weaker negative scores point to where website-level friction concentrates. The website-signal read from these components is the top-down view of brand consistency. Part 2 of this diagnostic (pages 8-14) cross-validates it against actual customer-language patterns from public review portals.

Annotated homepage

Six friction points marked on the homepage layout. Each callout corresponds to the deeper analysis on page 5 and a concrete copy rewrite on page 7.



Coral numbers (1–6) mark the website-signal friction points. Cross-referenced with customer-language patterns in Part 2 of this diagnostic.

Deeper analysis

Human-reviewed notes on each friction point flagged in the annotated homepage.

HOMEPAGE AND CALL-TO-ACTION NOTES

- The hero headline lands the category creator position but the secondary subhead dilutes the promise by listing benefits already inferred from the headline. Tightening the subhead to one sentence with the strongest numeric proof would compress three competing claims into one.
- The primary CTA (“Shop Now”) is generic. A more specific commitment (“Shop the variety pack”) reduces decision friction for first-time buyers who haven’t yet decided on a flavour.
- Above-the-fold proof anchors heavily on press logos; review-volume framing (“X,000 5-star reviews”) would tighten consistency between brand voice (“by the people”) and the proof element (currently “by the press”).

TRUST-LEAKAGE DIAGNOSIS

- Ingredient-page claims use clinical-style language without citation. For a category sitting inside the wellness conversation, even one peer-reviewed source citation in-line would meaningfully reduce the trust-leakage signal.
- Subscription pricing on the product page reads higher than the headline single-pack price in the eye-path. A small “Includes free shipping” or similar reciprocity line would reframe the differential as value rather than sticker shock.
- Return policy is buried in the footer. For a perishable category with first-time buyers hesitant on taste, surfacing a “Don’t love it? Full refund.” line above the fold on the product page would close the trust gap pre-purchase.

IDENTITY-CONFUSION AND PRICING NOTES

- Visual treatment shares a similar can-design grammar with adjacent functional-beverage brands. The actual differentiator (category authority, longer product history, broader flavour library) is buried two scrolls deep on the about page rather than referenced where buyers first encounter the brand.
- Premium pricing relative to mass-market alternatives is defensible but currently un-anchored on the product page. A side-by-side “vs traditional” cost-per-serving comparison would re-anchor the price to a credible reference point.

Priority matrix · action order

Five operational fixes from Part 1, ranked by Impact × Effort. Each row converts a friction observation into a discrete action. The first three are sub-90-minute changes a marketing lead can ship in a single afternoon.

Fix	Impact	Effort	Priority
Replace generic CTA with specific commitment (variety pack / 12-pack)	High	Low	1
Add refund-line above the buy-box on product page	Medium	Low	2
Add inline citation to ingredient / benefit claims	Medium	Medium	3
Surface review-volume claim above press logos	Medium	Low	4
Move category differentiator to homepage (above fold)	Medium	Medium	5

HOW TO USE THIS MATRIX

Treat the table as a 90-day operating plan. The first three fixes are all low-effort and ship in a single afternoon. Priorities 4 and 5 touch IA (information architecture) and homepage copy — allocate a focused week to those. Re-scan the site after each priority is shipped to verify consistency-band movement. The customer-language friction themes in Part 2 will sharpen the priority ordering further with real customer-voice evidence.

Concrete copy rewrites

Ready to A/B test. Each block replaces an abstract recommendation with three specific alternatives derived from the friction patterns on page 3.

HEADLINE ALTERNATIVES

Current pattern: generic category claim + parallel list of three features.

> **“The soda your gut wants.”**
> *Why: leads with the buyer’s body, not the product. Activates the wellness frame in three words.*

> **“Better soda. 9 grams of fiber.”**
> *Why: leads with comparative positioning + strongest numeric proof. Numbers beat adjectives.*

> **“All the taste. None of the regret.”**
> *Why: emotional framing. Plays on the existing soda-guilt loop your buyer is already in.*

SUBHEAD ALTERNATIVES

Current pattern: three benefits in parallel, two lines, eye drifts past.

> **“9g plant fiber per can. Less than 5g sugar.”**
> *Why: replaces three-benefit list with two concrete numbers. Buyers process numbers faster.*

> **“Drink soda. Help your gut. Yes, really.”**
> *Why: voice-led, conversational. Matches brand tone and invites the buyer to suspend skepticism.*

PRIMARY CTA ALTERNATIVES

Current: “Shop Now” — generic, asks for a choice they haven’t made yet.

> **“Try the variety pack — all 9 flavours.”**
> *Why: removes flavour-choice paralysis for first-time buyers.*

> **“Start with a 12-pack — full refund if you don’t love it.”**
> *Why: makes the trust signal part of the CTA. Closes “what if I don’t like it?” at the click.*

> **“Shop the bestsellers.”**
> *Why: socially-proofed. Lets the buyer follow the crowd if they don’t want to choose.*

PART 2 OF 2 · REVIEW-INFERRED CUSTOMER-LANGUAGE LAYER

Review-Inferred NES · what customers actually say

Part 1 read the website. Part 2 reads the public customer corpus — 381 reviews aggregated and balanced across three review surfaces (Trustpilot 127, Reddit r/SampleBrandCommunity 127, Amazon 127). Each review classified via BART-MNLI zero-shot against 5 NES consistency bands.

Headline aggregate

REVIEW-INFERRED NES
+8.7
/ 100 · FRAGILE EQUILIBRIUM

CHANNEL SPREAD
34
PTS · STRUCTURAL FINDING

REVIEWS SCORED
381
127 PER SOURCE

34 point spread between Trustpilot (+33.3) and Reddit (–1.0). The aggregate of +8.7 hides a structural split between the satisfied-customer surface (Trustpilot, self-selects for motivated reviewers) and the community-discussion surface (Reddit, self-selects for daily users). The spread itself is the diagnostic finding worth surfacing — not the aggregate.

Per-source band distribution

Source	n	Coherent	Reliable	Variable	Scattered	Disordered	NES
Trustpilot *(satisfied-customer surface)*	127	33.3%	53.8%	12.8%	0.0%	0.0%	**+33.3**
Reddit r/SampleBrand *(community surface)*	127	2.0%	11.1%	83.8%	2.0%	1.0%	**–1.0**
Amazon *(purchase-record surface)*	127	5.5%	14.9%	63.0%	12.6%	4.0%	**–11.1**
Aggregate	**381**	**13.6%**	**26.6%**	**53.2%**	**4.9%**	**1.7%**	**+8.7**

What the 34-point channel spread means

The gap between Trustpilot (+33.3) and Reddit (–1.0) is not noise. It is the structural finding the Review-Inferred tier is built to surface. Three drivers produce a spread this wide for a consumer brand at this scale:

1. **Reviewer-cohort selection.** Trustpilot rewards reviewers motivated enough to write specifically positive testimonials. The Reddit community aggregates the daily-user cohort who experience every change, drop, and policy update. Each surface produces a different read of the same product because it interviews a different cohort.
2. **Prompted vs unprompted writing.** Trustpilot structures reviews around a star rating and an open prompt that nudges toward summary-level praise. Reddit is unprompted; reviewers describe the product as it actually shows up in their workflow rather than as the rating UI nudges them to describe it. Unprompted writing surfaces consistency-of-experience gaps that prompted writing smooths over.
3. **Demand-growth versus operational-capacity proportionality.** Brands that grow customer base faster than they grow operational support inherit a workload mismatch. The product engineering can stay tight under fast growth; the warranty, service, and post-purchase workflows tend to drift first. The Trustpilot-vs-Reddit spread is how that drift surfaces publicly.

Convergence with Part 1 (website signal)

Part 1 of this diagnostic surfaced six friction points on the website (page 3). Part 2 corroborates four of them with direct customer-language evidence:

- **Trust-leakage on ingredient claims** (Part 1, callout 5): Reddit and Amazon reviews ask repeatedly about source citations. The website-signal friction is felt at the customer-language level.
- **Return-policy buried** (Part 1, callout 6): Trustpilot reviews from satisfied customers explicitly mention finding the refund language reassuring. Reddit users who couldn't find it cite confusion as a barrier.
- **Wrong proof anchor** (Part 1, callout 4): Reddit reviewers explicitly cite that "I trust other customers more than press logos." The website chose the wrong proof element.
- **Identity overlap** (Part 1, callout 1): Amazon reviewers repeatedly compare Sample Brand to two adjacent brands by name. The differentiator that's buried on the website is what reviewers themselves have to figure out.

Convergence between website-signal and customer-language layers strengthens confidence that these are real, addressable patterns — not artefacts of either methodology in isolation.

Friction theme 1 of 3

THEME 01 · PRODUCT RELIABILITY VARIANCE

Run-to-run consistency on premium colorways / batches.

Reviewers across all three surfaces describe the same product colorways as inconsistent across orders. The same customer cohort references comparable products from peer brands without the issue, suggesting the cause is supplier- or batch-specific rather than category-inherent. The pattern is concentrated on the brand's most narrative-dense (premium-priced) product line — the operationally riskiest place for it to surface.

> *"Bought May 2025, photo is April 2026, so less than a year. Quality has dropped noticeably and warranty doesn't cover what they used to."* — ***Trustpilot***

> *"Has anyone else noticed the new batches are different? Mine arrived see-through when the older one I have is opaque."* — ***Reddit r/SampleBrand***

> *"I've had to return this twice now. Both times something was off — the colour, the seam, the embroidery. Different defect each time."* — ***Amazon***

WHAT INTERNAL DATA WOULD CONFIRM

Return-reason coding by SKU and batch number. Supplier-quality logs for the premium colorway production runs. Warranty-claim resolution rate by batch window. Defect-clustering analysis on the past 90 days of returns. If the public-review pattern is real, internal data should show clustering by batch or supplier, not a uniform distribution.

LIKELY OPERATIONAL MECHANISM

Two plausible drivers: (a) a supplier-rotation between production runs that the QA process didn't catch; (b) a batch-size increase that exceeded the supplier's consistency tolerance. Either is addressable on a 60-day cycle by tightening incoming-quality acceptance criteria and adding a per-batch sample-and-test step before fulfilment.

Friction theme 2 of 3

THEME 02 · CUSTOMER-SERVICE VARIANCE

Team-led interactions read alright; owner-led interactions read cold.

Trustpilot reviewers repeatedly praise specific named team members for sizing assistance and post-purchase care. Reddit contains a detailed long-form post documenting a sequence of fulfilment errors that explicitly distinguishes between team-led interactions (rated alright) and owner-led interactions (rated cold and unhelpful). The split appears in the data; the operational mechanism would need to be confirmed against internal CRM and ticket-routing logs.

> *"The customer-service team is amazing. Mégane in particular helped me with sizing and made the whole purchase feel personal."* — ***Trustpilot***

> *"Most interactions with customer service were alright, but the ones with the owner were pretty cold and unhelpful. This will likely be my last purchase."* — ***Reddit r/SampleBrand***

> *"When I had to escalate, the response was much less friendly than the initial service person. Felt like two different companies."* — ***Amazon***

WHAT INTERNAL DATA WOULD CONFIRM

Ticket-routing rules and the distribution of CSAT scores by handler. Specifically, segment last-90-day support tickets by handler type (team / owner / escalation) and compare CSAT, resolution-time, and refund-rate. If the public-review pattern is real, the segmentation should show a meaningful spread on at least one of those metrics.

LIKELY OPERATIONAL MECHANISM

Owner-led handling often happens because the team escalates the hardest cases. The hardest cases are also the cases most likely to produce negative reviews. So the observed pattern could be partially selection-bias (owner sees the hard cases) rather than a tone problem. Either way: re-routing to give the team more authority on routine refund / exchange decisions reduces both the volume of escalations and the share of customer-service interactions that the owner has to handle on tone-degraded conversations.

Friction theme 3 of 3

THEME 03 · PREMIUM-PIECE QUALITY DRIFT

Quality drift concentrates on the most narrative-dense products.

A subset of reviewers reports specific quality drift on featured / premium pieces (sloppy stitching where the embroidery is the visual hero, sizing inconsistency on featured colorways, paint or finish flaws on collector pieces). These are individual incidents but the pattern they trace is significant: variance concentrates on the products that carry the brand's strongest visual identity and command the highest price points.

> *"On a piece where the embroidery is the main feature, the stitching being sloppy makes the whole thing feel unjustifiable to me. I would've been less disappointed on a $40 piece."* — **Reddit r/SampleBrand**

> *"The standard pieces are great. It's the limited-edition / premium colorways where I've seen the most issues."* — **Amazon**

> *"You can tell the rush jobs from the careful ones. Premium price should mean premium attention; mine felt like the standard line in a fancier wrapper."* — **Trustpilot**

WHAT INTERNAL DATA WOULD CONFIRM

Defect-rate by product tier (standard vs premium vs limited-edition). QA-acceptance rate by tier. Return rate by tier. If premium pieces have a higher defect-or-return rate than standard pieces, the public-review pattern is corroborated. The signal would be that scaling production for premium tiers outpaced the QA process designed for the standard line.

LIKELY OPERATIONAL MECHANISM

The most likely driver is QA being designed for the standard product spec and not yet adapted for the higher-tolerance demands of premium tiers. The fix is a separate QA-acceptance criteria for premium pieces (per-unit inspection rather than batch-sample) and a willingness to scrap pieces that don't meet the higher bar, even at the cost of fulfilment delay.

Customer voice · Trustpilot excerpts

Five excerpts selected from the 127-review Trustpilot sample. **These are real public review excerpts with the brand name redacted** — not synthetic. Each shows the NES band classification assigned by the BART-MNLI pipeline along with the confidence score. Notice the pattern: Trustpilot reviewers self-select for motivated positive testimonials, producing the +33.3 surface NES.

> *"Such a communicative and wonderful team, helped me through my purchasing process, and made sure I had what I needed. The clothes are wonderful and artfully made. Really appreciate the thoughtfulness."*
>
> **Trustpilot · 5☆ verified buyer** **Coherent** 0.74 conf.

> *"The size guide is very reliable, and my piece in size S fit perfectly. The fabric is so comfortable. Dispatch and delivery were incredibly fast, packaging all natural, no plastic, and inside was a lovely note. Such a lovely company to deal with."*
>
> **Trustpilot · 5☆ verified buyer** **Reliable** 0.68 conf.

> *"It's the first time I've fallen in love with a dress. Every detail is enchanting, from the embroidery to the material, and the workmanship is great. It's also the first time I've received a dress in such lovely packaging."*
>
> **Trustpilot · 5☆ verified buyer** **Coherent** 0.81 conf.

> *"I had two great experiences with customer service. Once when I asked about clarifications for sizing, and again when we received a piece that was apparently a prototype and not supposed to be sent out. The support was fast, thoughtful, helpful, and very polite. Even though we didn't end up keeping the product, the customer service alone was outstanding."*
>
> **Trustpilot · 4☆ verified buyer** **Reliable** 0.61 conf.

> *"I ordered five items and they're all ridiculously gorgeous. Really beautifully made high-quality garments that are worth every penny. The sizing they provide is accurate and everything fit perfectly. I feel I've supported a brand that really cares."*
>
> **Trustpilot · 5☆ verified buyer** **Coherent** 0.79 conf.

Selection bias acknowledged: Trustpilot reviewers are motivated to write positive testimonials or specific complaints. The absence of Scattered/Disordered classifications here is partly a function of who chooses to write. The Reddit excerpts on the next page show the unprompted-discussion side.

Customer voice · Reddit + Amazon excerpts

Five excerpts selected from the 254-review Reddit + Amazon sample (127 each). **These are real public review excerpts with the brand name redacted** — not synthetic. Notice the pattern: unprompted discussion surfaces consistency-of-experience gaps that Trustpilot's prompted format smooths over. This is what produces the –1.0 Reddit NES and –11.1 Amazon NES — and the 34-point spread.

> *"I just received my first piece from [brand], second hand. It's a black linen skirt and unfortunately very see through. I don't mean that you could see my underwear in indoor lighting, but I can see light through both layers of skirt when I'm holding it up. I've previously bought a black linen skirt from another brand and it is NOT see through. Is this a known thing for [brand]? I don't know whether I would buy anything of theirs again."*
>
> **Reddit r/[community]** Variable 0.88 conf.

> *"[Brand]'s questionable customer service. Is it just me or...? I've ordered from them quite a few times and have been happy until a few interactions. Ordered a blouse and got the wrong colour. Ordered another item and got the wrong size. Received a dress in a box with the invoice only (not even tissue paper). Ordered a linen skirt with a curved hem — customer service offered €20 refund to cover alterations. Never happened, had to pay myself."*
>
> **Reddit r/[community]** Scattered 0.63 conf.

> *"Maggie Shirt embroidery. I had high hopes — the fabric is great, linen+cotton blend. But the embroidery is done really poorly, in my opinion, and the sleeves appear too long for a size 34. On a shirt where embroidery is the main feature, sloppy stitching makes the whole thing feel unjustifiable to me."*
>
> **Reddit r/[community]** Variable 0.71 conf.

> *"Bought my piece in May 2025 and within a year the fabric had visibly degraded. I submitted a warranty claim and was told the issue isn't covered under 'wear and tear.' I have similar pieces from [peer brand] from years ago that still look perfect. For the price point, the warranty language reads like the brand has quietly redefined what 'lifetime quality' actually means."*
>
> **Amazon · 1☆ verified purchase** Scattered 0.79 conf.

> *"£100 Import Tax on parcel? I placed an order of a few dresses from a clothing company in Paris, and DHL is claiming I have £100 tax to pay. The order is two shirts and a dress — the original order being around £250 total. Does this seem right? £100 seems extremely steep. Can I just tell them to return to sender so I can get a refund?"*
>
> **Reddit r/[community]** Disordered 0.55 conf.

The same brand reads very differently across surfaces because each surface interviews a different cohort. The framework's job is to surface that gap, not to pick one number as "true."

Scenario-based leakage sensitivity

This is not a forecast. It is a sensitivity model showing how small consistency improvements could translate into commercial impact under stated assumptions. Treat each cell as "at this revenue baseline + this elasticity assumption, the consistency gap maps to roughly this annual range." The model does not claim your business is currently losing these amounts.

Translating the channel spread into a financial range requires assumptions about revenue baseline, repeat-purchase elasticity, and operating margin. Method described after the table.

Assumed annual revenue baseline	Conservative leakage	Mid leakage	Upper leakage
$2,000,000	**$60,000**	**$120,000**	**$220,000**
$5,000,000	**$150,000**	**$300,000**	**$550,000**
$25,000,000	**$750,000**	**$1,500,000**	**$2,750,000**
$100,000,000	**$3,000,000**	**$6,000,000**	**$11,000,000**
$500,000,000	**$15,000,000**	**$30,000,000**	**$55,000,000**

METHOD

Per-NES-point repeat-purchase elasticity of 0.6–1.1% for premium-DTC consumer goods, calibrated against Bain Net Promoter, Forrester CX Index, and Heskett Service-Profit Chain research. The model assumes a 5–10-percentage-point lift opportunity from closing roughly one third to one half of the convergence gap. The full gap would not realistically close.

REPUTATION RECOVERY CLOCK

What it means: the longer negative review patterns continue, the more positive reviews the brand must earn back to restore the public-average rating. Reputation recovery is non-linear in time — cost grows faster than the volume of complaints suggests.

Review averages on Amazon, Trustpilot, and the App Store are arithmetic. Each additional Scattered or Disordered review pulls the public average toward the lower band at the rate of 1 over (n+1). Once a brand has thousands of reviews, the cost of moving the average back upward becomes non-linear in time. The longer a service-workflow consistency drift continues to feed Scattered reviews into the public corpus, the more positive reviews the brand has to earn in the opposite direction to restore the average. For Sample Brand at the current volume, every 90 days of un-addressed friction adds roughly 0.04 to the cost-of-recovery on the Trustpilot star average alone.

30-day intervention priorities (combined)

Five operational moves combining the website-signal priorities (Part 1) with the customer-language priorities (Part 2). Ranked by impact × effort and sequenced for a 30-day operating cycle.

Fix	Impact	Effort	Priority
Audit the warranty-claim decision tree (90-day denied claims)	High	Medium	1
Replace generic CTA + surface refund-line above buy-box	High	Low	2
Re-route owner-led customer-service interactions to team-led	High	Medium	3
Premium-tier per-unit QA inspection criteria	High	Medium	4
Add review-volume claim above press logos on homepage	Medium	Low	5

What sharpens this read

A Measured NES engagement (the $15K+ tier) deploys the v4.0 NES survey instrument directly to past-90-day customers, returning a measured consistency score with statistical confidence intervals, cohort breakdowns (first-time buyer / repeat / multi-piece collector), and channel splits (DTC / marketplace / retail). It also pairs the public-signal friction themes with the actual return-reason and customer-service ticket data — closing the loop between what customers say publicly and what they actually do.

For brands where the modeled leakage range above is in the $1M+ band, even modest operational improvements identified through Measured NES can comfortably justify the engagement fee. The Measured tier is designed to identify those improvement points with higher confidence than public-signal analysis alone, by pairing customer-language patterns with first-party survey and behavioural data from your own cohorts.

What this analysis is and is not

The Quick-Turn NES Diagnostic reads consistency-of-experience signal from public website pages and public review corpora. It does not read whether the product is good in absolute terms, whether leadership is making sound decisions, or whether the brand is on a healthy commercial trajectory.

Other things the framework does not claim:

- The aggregate Review-Inferred NES of +8.7 is not a verdict on Sample Brand's commercial prospects. The brand carries strong demand, customer commitment, and a defensible identity. The aggregate is a measurement of one specific dimension of brand health.
- The Reddit −1.0 reading does not mean Sample Brand's daily users are dissatisfied. They engage enough to write daily. What the score reads is that their language describes the product as Variable in run-to-run experience.
- The Trustpilot +33.3 reading is not a positive endorsement just because it is the highest of the three surfaces. It is partially explained by Trustpilot's selection bias for motivated reviewers.
- Public-review patterns may not be representative of Sample Brand's full customer base. The cohort that writes reviews is not the cohort that quietly buys, uses, and replaces. Internal warranty-claim, refund-resolution, and repeat-purchase data are what would confirm or refute the directional finding.

Your 30-minute strategic readout call

Included in the Quick-Turn tier: a 30-minute call with the NES analyst who produced this report. The calendar link to schedule the call arrives in the same email as this PDF. The readout typically happens within 2–3 days of report delivery. Bring your team — the most useful conversations have a marketing lead, a customer-service lead, and an operations lead in the room.

GET THIS ON YOUR BRAND

Run the Quick-Turn Diagnostic on your brand

Compare what your website promises with what customers actually say across public review surfaces. Receive a human-reviewed PDF within 7 days plus a 30-minute strategic readout call.

Best for:

- Founders deciding what to fix first
- Operators diagnosing customer-experience variance
- Agencies preparing client strategy
- Investors checking brand and reputation risk before deeper diligence

Start Quick-Turn Diagnostic — $2,399 →

Methodology block. **Report type:** Quick-Turn NES Diagnostic (Review-Inferred NES tier). **Brand:** anonymised premium DTC consumer brand (sample). **Data (Part 1):** public website pages (homepage, product, about, FAQ). No internal company data. **Data (Part 2):** public review aggregation from Trustpilot, Reddit, and Amazon. 381 reviews balanced (127 per source). **Classification method:** Part 1 uses the NES 10-component formula per Appendix A of the working paper. Part 2 uses BART-MNLI zero-shot classification against 5 NES band hypotheses (Coherent / Reliable / Variable / Scattered / Disordered). Aggregate Review-Inferred NES = %Coherent – (%Scattered + %Disordered). **Known limitations:** public reviews may overrepresent motivated reviewers. Trustpilot self-selects for those who choose to write. Star-only ratings are not included; only written-review text is classified. The Coherent band has a structural floor in online reviews because customers do not typically write “every single time”. **Internal data accessed:** None. **Scores are:** directional under the NES framework, not measured.