HG Insights®

WHITE PAPER

The Glass-Box Shift: Why Explainable Scoring Alone Won't Fix Sales Trust

What sales leaders need to know as account scoring moves from opaque to open, and why openness by itself won't get a rep to act on a score

July 2026

Executive Summary

Account and lead scoring is going through a quiet correction. For years, the pitch was simple: let a model rank your accounts, trust the number, work the list. Reps didn't buy it, and the category is now responding. Explainability, being able to see what actually drove a score, is turning into the baseline expectation rather than a differentiator. That shift is good news. It is also not the whole answer.

A score a rep can see through is not automatically a score a rep will act on. If the reasoning behind that score is wrong, showing it clearly doesn't fix the problem; it just makes the problem easier to spot. Reps who can see exactly why a model made a bad call don't develop more trust in that model. They develop less, and this time they can prove why.

This paper looks at what's actually changing in scoring, why transparency alone tends to fail the moment it's tested against a rep's own account knowledge, and what a more complete standard looks like: a score that's explainable, adjustable by the person using it, and built from more than one signal, inside the CRM reps already work in. It closes with a practical set of questions for any sales organization currently evaluating scoring tools, whether that evaluation is happening this quarter or has been quietly stalled for a year.

The State of the Category: Explainability Becomes the Price of Entry

A black-box score asks for faith, and sales teams that have been burned once aren't in a faith-extending mood.

Picture a rep opening their account list on a Monday morning. Fifty accounts, ranked by a score. No context, no reasoning, just a number next to a name. The rep has seen this number be wrong before, so instead of working the list top to bottom, they start re-checking it. Pulling up firmographic data. Cross-referencing intent signals somewhere else. Asking a teammate if they've heard anything about the account. By the time they've "verified" the top ten, half the morning is gone and they haven't made a single call.

This is not a hypothetical. It is the default condition in a lot of sales orgs running a scoring tool today, and it's the reason the category is under pressure to change.

That pressure is showing up as a real shift in how scoring tools position themselves. Explainability, the ability to show which signals drove a score and why, is increasingly treated as a baseline requirement rather than a premium feature. Vendors across the space are adding "why" screens, surfaced reasoning, and score breakdowns that didn't exist a couple of product cycles ago. The direction is clear even without naming names: opaque scoring is losing ground as an acceptable default, and "we'll show our work" is becoming table stakes rather than a pitch.

34% vs. 23%

Reps at organizations hitting 90%+ quota attainment spend 34% of their time actually selling. Reps at lower-performing organizations spend 23%. Forrester, "Four Proven Ways to Reach at Least 90 Percent Sales Quota Attainment"

That eleven-point gap is worth sitting with. It isn't a talent gap or an effort gap. It's a time-allocation gap, and time spent re-verifying a score a rep doesn't trust is time that shows up on the wrong side of that split. A sales organization does not need a better closer to close the gap. It needs its reps spending more of the day selling and less of it auditing a tool.

The status quo of "trust the score because the vendor said so" was never sustainable. What's changing now is that reps have a reasonable alternative to demand: show me why.

Why Explainability Alone Still Fails Reps

Visibility is not the same thing as accuracy, and a clearer view of a bad model is still a bad model.

Here's the problem with stopping at explainability. A rep who can suddenly see the reasoning behind a score is also, for the first time, positioned to catch it being wrong.

Think about what "explainable" actually promises: a window into the model's logic. That window doesn't discriminate between good logic and bad logic. It shows both with equal clarity. A score built on stale firmographic data, or on a fit signal that hasn't accounted for a recent reorg, or on intent data that's three weeks old, is just as explainable as a score built on something solid.

This is precisely what happened at one sales organization that adopted an explainable scoring tool expecting transparency to solve its trust problem. It didn't. Reps used the newly visible reasoning to spot the model's mistakes faster than before, not to build confidence in it. Every exposed miscalculation became a data point against the tool, and the team's trust in the score went down, not up.

> *"A model can be completely transparent and still be wrong. Showing the work doesn't fix that. It just gives everyone a clearer view of the mistake."*
>
> The core failure mode of explainability-only scoring

This matters because most sales teams evaluating scoring tools today have already been through some version of this. An in-house model that never got explained well enough to trust. A legacy tool that got explained, and then turned out to be wrong often enough that the explaining made things worse. None of that lands as new to a rep who has lived through it once already.

KEY TAKEAWAY

Explainability answers "why did I get this score." It does not answer "is this score right," and it does not answer "does this score reflect what I already know about this account." Both of those gaps are where rep trust actually gets built or lost.

The Three-Part Fix: Explainable, Adjustable, and Native

Any one or two of these three elements, without the third, leaves the same trust gap open in a different place.

The uncomfortable truth is that explainability can make a bad scoring model's problems more visible, not less. That's not an argument against explainability. It's an argument that explainability is the floor, not the finish line. Based on where scoring tools are succeeding and failing with sales teams, three elements need to hold together at the same time.

Explainable

A rep needs to see exactly which signals drove a score: fit, intent, behavioral activity, whatever the inputs are. This is the entry requirement. Without it, nothing else matters, because there's nothing to adjust or verify in the first place.

Operator-adjustable. This is the piece most scoring tools skip. A rep who's worked an account for two years, or who just came off a call with the buying committee, knows things the model doesn't. A generic fit signal might score an account low because the model can't see a relationship the rep already has. If the rep can't adjust the weight behind that signal to reflect what they actually know, they're stuck choosing between trusting a model that's ignoring their knowledge, or ignoring the model entirely and going back to gut feel. Neither is a good option. A model that can be tuned by the person using it, not just read by them, is a fundamentally different proposition than one that can only be explained. For a rep, this is the difference between a tool they'll work and a tool they'll route around.

Multi-signal, native to the CRM

Fit alone isn't enough. Intent alone isn't enough. Behavioral activity alone isn't enough. And a score that lives in a separate tool, outside the CRM a rep already works in all day, adds a reconciliation problem on top of a trust problem: now there are two systems with two opinions, and someone has to decide which one wins. A single score, combined from multiple signal types and built directly into the CRM, removes both problems at once.

MATURITY STAGE	WHAT THE REP SEES	WHAT'S STILL MISSING
Opaque score	A number, no reasoning	Everything. No basis for trust either way.
Explainable score	A number and the reasoning behind it	Whether the reasoning is right, and whether it reflects what the rep already knows
Glass-box score	A number, the reasoning, the ability to adjust it, and enough combined signal to be credible, inside the CRM	Nothing structural. What's left is proving it out account by account.

None of these three elements substitutes for the others in a glass-box score. A tool that's explainable and multi-signal but locked to a fixed weighting model still asks reps to defer to logic they can't correct. A tool that's adjustable but opaque is just a black box with a dial on it, and reps have no way to know if the dial is doing anything meaningful. A tool that's explainable and adjustable but single-signal is transparent about a narrow, incomplete picture. The combination is what earns sustained use, not any one piece in isolation.

KEY TAKEAWAY

The question worth asking isn't "is this scoring tool explainable." Most of the category can now say yes to that. The question is whether a rep can adjust the model based on what they know, and whether the score is built from more than one kind of signal in the first place.

What the Time Math Actually Shows

The cost of an untrusted score isn't abstract. It's hours, and it shows up in the gap between top and bottom performers.

It's worth being precise about where the cost of an untrusted score actually shows up, because it isn't abstract. It's hours.

~15%

Reps spend roughly 15% of their time on prospect research, a category that includes verifying and cross-checking accounts before acting on them. Forrester, "The Biggest Obstacle to Improving Your Sales Productivity"

Fifteen percent of a rep's week is not a rounding error, and not all of that time is first-pass research on a genuinely new account. A meaningful share of it, in organizations running an untrusted score, is re-verification: checking a number the tool already gave them because the tool has been wrong before.

Put the two Forrester numbers next to each other and a pattern emerges. Top-performing orgs have more selling time available (34% versus 23%), and the category-wide research-time figure (roughly 15%) suggests a meaningful chunk of every rep's week is going somewhere other than a first conversation with a prospect. Untrusted scoring isn't the only thing competing for that time. It's one of the more fixable things competing for it, because unlike headcount or pipeline generation, it's a workflow problem with a workflow-level answer. For a Sales Manager, it also shows up as a consistency problem: when reps quietly abandon the score and return to gut-feel account tiering, territory distribution becomes impossible to defend in a pipeline review, because no two reps are working off the same logic.

To be direct about a real limitation here: this paper is not claiming that fixing scoring trust alone closes an eleven-point selling-time gap. That gap has multiple causes, and no single research citation isolates scoring trust as the whole story. What the data does support is narrower and still worth acting on: reps spend real, measurable time on research and verification, and some of that time is spent re-checking scores they don't trust.

The shift shows up concretely in how teams actually work once the trust problem is solved. At Snyk, reps that previously spent roughly an hour on manual account research before acting on a score reduced that to two to five minutes after adopting glass-box scoring. The same team hit 100% quota attainment in Q1. That outcome doesn't isolate scoring as the only variable, but the research-time drop is precise enough to be meaningful on its own: when reps trust the score, they stop re-verifying it.

What This Means for a Sales Org Evaluating Scoring Tools Today

A few questions in the demo cut through the noise faster than any feature list.

If your team is looking at scoring tools right now, or living with one that isn't getting used, deferring the evaluation doesn't make the time cost go away. It just pushes the re-verification hours into next quarter. A few questions cut through the noise faster than a feature list.

- **Ask what "explainable" actually means in the demo.** Some tools show a reasoning summary after the fact. Others let a rep drill into the exact signal weights behind a specific score. Those are not the same thing.
- **Ask who can change the model, and how fast.** If the answer involves a data science team, a change request, or a multi-week cycle, the model is effectively fixed from a rep's point of view.
- **Ask how many signal types feed the score.** A tool built on intent data alone, or fit data alone, is transparent about an incomplete picture.
- **Ask where the score lives.** A score that requires a rep to leave the CRM to check it is adding a step, not removing one.
- **Watch for the pattern that kills adoption quietly.** A tool doesn't have to be rejected outright to fail. It can sit there, technically live, while reps quietly go back to their own gut-feel tiering.

> **KEY TAKEAWAY**
>
> The fastest diagnostic for any scoring tool is this: can a rep change the model based on something they know that the model doesn't, without waiting on anyone else to do it for them? If the answer is no, explainability alone will not be enough to sustain trust once the model gets something wrong, and it will get something wrong eventually.

None of this is a call to distrust every scoring tool on the market. It's a call to evaluate past the first claim a vendor makes.

Conclusion: Key Takeaways

The scoring category is in the middle of a real correction, and explainability is the visible part of it. The case for inaction (waiting to see how the category settles) is reasonable on the surface. The problem is that reps don't wait. They fill the vacuum with manual research, and that time cost accumulates at a rate of $11,000 to $17,000 per AE per year whether or not the org is actively evaluating a tool. That correction is worth taking seriously, and it's also worth being clear-eyed about what it does and doesn't solve. A rep who can see why a score landed where it did is better off than a rep working a flat, unexplained list. That same rep, looking at a wrong score with perfect clarity, is not more likely to trust the tool. If anything, they're better equipped to stop trusting it, correctly.

1. **Explainability is becoming the baseline expectation** across the scoring category, not a differentiator. Plan evaluation criteria accordingly; assume most tools you look at will claim it.
2. **Visibility into a model's logic does not improve the model's accuracy.** It just makes existing errors easier for a sharp rep to catch, which can make trust worse, not better, if the underlying model isn't solid.
3. **Operator-adjustability is the piece most tools still lack,** and it's the piece that turns explainability into something a rep will actually act on.
4. **Multi-signal scoring removes the reconciliation problem** that comes from running several separate, disagreeing tools.
5. **The real cost of an untrusted score isn't abstract.** It shows up as time, specifically the research and re-verification time that separates lower-performing sales organizations from the ones hitting 90%+ quota attainment.

A sales organization evaluating scoring tools today doesn't need to solve for explainability anymore; the category is largely handling that on its own. What's worth solving for is whether the tool in front of you can be adjusted by the person using it, and whether it's built from enough combined signal, inside the workflow reps already live in, to be worth adjusting in the first place.

About HG Insights

HG Insights delivers AI-powered Revenue Growth Intelligence (RGI) solutions that modernize GTM strategy for B2B companies. Its platform turns deep technographic, IT spend, intent, and buyer data into actionable insights and automated workflows that accelerate pipeline and improve revenue predictability. 95% of Fortune 1000 B2B tech companies and all major hyperscalers rely on HG Insights to grow revenue, boost efficiency, and improve retention. Learn more at hginsights.com.

Ready to See How Your Reps Would Score This Differently?

Explainability is the current baseline for account scoring, not the finish line. If you're evaluating how a score is built, and whether your reps could actually adjust it based on what they already know, request a live walkthrough using your own account list at hginsights.com/demo, or reach out to sales@hginsights.com.

References

1. Forrester. "Four Proven Ways to Reach at Least 90 Percent Sales Quota Attainment." forrester.com
2. Forrester. "The Biggest Obstacle to Improving Your Sales Productivity: Wasted Time." forrester.com