

THE BUYER'S GUIDE FOR L&amp;D AND SALES ENABLEMENT

# AI Sales Training: The Buyer's Diagnostic Guide

A field-tested toolkit for evaluating AI sales training platforms – built from the three patterns that quietly kill deployments, the ten questions every vendor must answer, and the day-30, 60, and 90 checkpoints that separate a real program from an expensive dashboard.

**10**

Diagnostic questions every vendor must answer before you sign.

**3**

Failure modes that quietly destroy AI training investments.

**90**

Days to know whether the pilot is on track or in trouble.

**INSIDE THIS GUIDE**

- 01 The 10-Question Buyer's Diagnostic
- 02 The Three Failure Mode Detection Framework
- 03 The 90-Day Success Benchmark Sheet

## — THE 10-QUESTION BUYER'S DIAGNOSTIC

# 10 Questions Every AI Sales Training Vendor Must Answer

The critical questions to ask before committing to an AI sales training solution – grouped by what's actually at stake: **Compliance & Security**, **Realism & Adoption**, and **Outcomes & Measurement**.

## Compliance & Security

---

- 1. Model Scoring.** Does the platform score against our approved messaging, not a generic model rubric?
- 2. Non-Compliant Claims.** What happens when a rep makes a deliberately non-compliant claim? How is it flagged and addressed?
- 3. Audit Logs.** Can you produce session-level audit logs on demand, detailing rep interactions and system evaluations?
- 4. Content Usage.** Is our proprietary content used to train, fine-tune, or update your public-facing AI models?

## Realism & Adoption

---

- 5. Voluntary Return Rate.** What is your typical voluntary return rate after week two of deployment, and how do you achieve it?
- 6. International Change Management.** What does your change management approach look like for diverse international markets and cultural contexts?
- 7. Manager Anxiety.** How do you handle the “manager replacement anxiety” problem, ensuring managers feel empowered, not threatened?

## Outcomes & Measurement

---

- 8. Rep Readiness.** How do you define and measure “rep readiness” in actionable terms, beyond simple completion metrics?
- 9. Field Execution Metrics.** What specific field execution metrics improve as a direct result of your training, and on what timeline?
- 10. Quota Attainment Link.** How do you instrument and demonstrate the direct link between training engagement and quota attainment?

## — THE THREE FAILURE MODE DETECTION FRAMEWORK

# The Three Patterns That Quietly Kill AI Sales Training Deployments

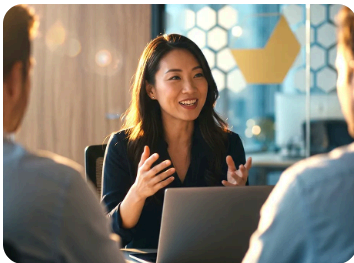
The cost is rarely visible in month one. It shows up in month three, when adoption flattens. In month six, when your VP of Sales asks for proof of readiness. Or in year two, when an audit pulls up an off-label claim no one caught in training. This is the framework for catching all three before they cost you.

## 01

### FAILURE MODE

#### THE VISIBLE SIGNAL

**Clean scores. No flags. Reps moving through certification without friction.**



## Compliance Theater

The platform scores fluent language as compliant language.

A rep can deliver an off-label claim with the cadence and confidence of an approved message, and the scoring engine will pass it. Most AI sales training tools score against a generic rubric of conversational quality – tone, pacing, confidence, keyword presence. They are not scoring against your approved messaging, your authorized claims, or your specific compliance requirements.

The result: a rep who fluently misstates a suitability requirement, an off-label benefit, or a clinical claim does not get flagged in training. They get flagged in audit, **after** they have said it to a real customer.

#### — THE DIAGNOSTIC TEST

Take one of your real, sensitive scenarios. Write a deliberately non-compliant response – off-label, unapproved claim, missing fair balance, whatever applies. Have the vendor run it through their platform live, using your content. Watch what happens.



#### THE INTERVENTION

If the system passes the non-compliant response, the evaluation engine is broken for your use case. Do not deploy until **adversarial test cases are part of the sign-off process**. Require this as a contractual condition, not a verbal assurance.



# 02

## FAILURE MODE

### THE VISIBLE SIGNAL

**High completion.  
Strong month-one  
engagement.  
Voluntary use  
collapses by week  
three.**



## The Adoption Cliff

Mandated completion is a forcing function. It is not a measure of value.

Month one looks great. Completion rates are high. Engagement metrics are strong. The pilot report writes itself. Then month three arrives. Reps do the minimum required to stay compliant. Voluntary engagement drops to near zero. The data still looks fine in aggregate – everyone completed the mandated sessions – but the platform is dead from a behavior-change standpoint.

Reps will do what is required. The question is whether they come back when they do not have to. If the platform is not realistic enough to feel useful, they will not. The simulations feel like work to satisfy a tracker, not practice to get better. You have an expensive tool that meets compliance requirements and produces no measurable change in field performance.

### — THE DIAGNOSTIC TEST

Measure voluntary return rate in **week two** of your pilot. Not week one – novelty is doing the work. Not week six – mandates are doing the work. Week two, when novelty has worn off and mandates have not yet kicked in. That is when realism is doing the work, or it is not.



### THE INTERVENTION

Set a **voluntary return rate floor** before you sign. If a vendor cannot tell you their typical voluntary return rate after week two across comparable deployments, they do not measure it – which means they do not optimize for it – which means you are buying a tool that will not survive month three.

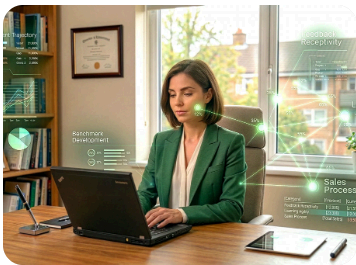


# 03

## FAILURE MODE

### THE VISIBLE SIGNAL

**Sophisticated dashboards. Numbers trend up. Your sales leader still doesn't trust the data.**



## The Dashboard Mirage

You have completion data. You cannot tell your VP of Sales which reps are ready on Monday.

You have login data, time-on-platform data, and certification pass rates. The dashboards look sophisticated. The numbers trend up. Everyone agrees the program is working. But nothing on the dashboard answers the question your sales leader actually needs answered.

Completion is an activity metric. Readiness is a performance metric. Most platforms instrument activity because activity is easy to measure. Performance requires defining what good looks like, scoring against that standard, and producing a verdict that someone can act on. You end up with a defensible budget item and an indefensible business case.

### — THE DIAGNOSTIC TEST

Ask your vendor one specific question: *"Can this data tell me which rep is ready to be in front of a customer on Monday?"* Watch them either show you the screen that answers it – or pivot to talking about dashboards.



### THE INTERVENTION

Define **"certified" in performance terms before deployment**, not after. A certified rep has demonstrated the specific behaviors that predict field success – not completed the modules, not passed the quiz. Get the definition in writing, agreed by sales leadership, before you sign.

### How to Use This Framework

#### FOR NEW PLATFORM EVALUATIONS

Run all three diagnostic tests in the same vendor meeting. Vendors who answer all three confidently are rare. Vendors who deflect on any of them have given you your answer.

#### FOR PLATFORMS YOU ALREADY OWN

Run the same three tests against your existing tool. You may find it passes one or two, but not all three – useful information that tells you where to push for change, and where to start looking for alternatives.

— THE 90-DAY SUCCESS BENCHMARK SHEET

# Three Checkpoints. Three Decisions. One Renewal Call.

Most pilots are evaluated at the end. By then, the conclusions are baked in. The teams who get value from AI sales training evaluate at **three checkpoints – day 30, day 60, day 90**. Each produces a specific decision. Each, if missed, points to a specific intervention.

<p>CHECKPOINT 01 <b>Day 30</b> Realism Is Working</p>	<p>CHECKPOINT 02 <b>Day 60</b> The Signal Becomes Actionable</p>	<p>CHECKPOINT 03 <b>Day 90</b> Field Execution Lifts</p>
---	--	--

CHECKPOINT 01 **Day 30** Realism Is Working Voluntary return > 60%

<p><b>● WHAT GOOD LOOKS LIKE</b></p> <ul style="list-style-type: none"> <li>■ Voluntary return rate above 60% in weeks two and three.</li> <li>■ Every rep has completed at least one full compliance scenario.</li> <li>■ Auto-fail logging is active and producing readable session-level records.</li> <li>■ Manager dashboards show rep-level engagement, not just cohort averages.</li> </ul>	<p><b>● IF YOU ARE BEHIND</b></p> <ul style="list-style-type: none"> <li>■ The realism is not landing. The simulations feel like a tracker.</li> <li>■ Ask five reps directly: <i>"What would make you come back?"</i></li> <li>■ Content answers (more scenarios, buyers, product context) are fixable.</li> <li>■ Experience answers (feels fake, avatar flat, scoring unearned) – pilot in trouble.</li> </ul>
--	---

**MEASURE**

Voluntary return rate · weekly	Session-level scoring spread	Compliance flag rate
--------------------------------	------------------------------	----------------------



CHECKPOINT 02

# Day 60 The Signal Becomes Actionable

Coaching ratio inverts

● WHAT GOOD LOOKS LIKE

- Skill-gap analytics identify the bottom decile by specific behavior, not by completion.
- Manager time has shifted from reviewing recordings to running development conversations.
- Sales leaders start asking for the data unprompted.

● IF YOU ARE BEHIND

- The data is descriptive, not prescriptive. Managers can't act on it.
- Usually a definition problem (what "good" looks like) or a feedback-loop problem.
- If the vendor can't translate data into coaching practice, you have a tool – not a partner.

MEASURE

- Manager review hours vs. coaching hours
- Rep-level skill gaps identified
- Platform-data 1:1s per manager / week

CHECKPOINT 03

# Day 90 Field Execution Lifts

CF0-defensible lift

● WHAT GOOD LOOKS LIKE

- Pilot cohort shows measurable lift vs. control: call quality, message adherence, ramp velocity, time-to-certification.
- The lift is large enough to defend to a CFO – not just visible in a chart.
- You can name three specific behaviors that changed at the rep level, with examples.

● IF YOU ARE BEHIND

- Wrong metric measured. Recheck the day-zero success criteria – the lift may be real on a different one.
- Cohort too narrow. 25 reps in one region is noisy. Signal may be real but unprovable.
- Platform may not produce lift on this role. The answer no one wants – the answer that saves a year-two budget fight.

MEASURE

- Pilot vs. control on field execution metrics
- Time from start to readiness
- Field manager qualitative assessment

## THE DAY-90 CLOSING DECISION

### Expand

The data supports rollout. Move from pilot cohort to full deployment with a named lift metric.

### Extend

The data is promising but incomplete. Add reps, regions, or time – and re-run the checkpoint.

### End

The data doesn't support investment. The teams who get the most from AI training are the ones willing to make this call.

---

**ABOUT QUANTIFIED**

# The AI sales coaching platform built for the rigor of regulated industries.

Quantified personalizes readiness for every rep, scales manager coaching across the field, strengthens compliance before reps reach a client, and gives commercial leaders the visibility they need to improve performance.

From new-hire onboarding to product-launch certification and continuous coaching, Quantified helps high-stakes commercial teams prepare for the conversations that drive success.

**TALK TO SALES****512-240-2522****sales@quantified.ai****LEARN MORE****quantified.ai**

Request a diagnostic walkthrough with our team.

**SCAN**

Visit  
quantified.ai