Assessing the AI-Ready People Leader

SIX CRITERIA, THEIR REFERENCE TWINS, AND THE ASSESSMENT UNDERNEATH | TL PARTNERS | ARCHITECTURE BUILT 2021-2023; CRITERIA FROM FIELD PROGRAMS 2025-2026 | THOMAS IGEME

Every sponsor we talk to is suddenly asking the same question about the people seat: who do we need for the future? The searches are kicking off, and the old profile has stopped predicting. The leader who ran annual cycles beautifully, kept the org chart clean, and delivered hiring at pace may be exactly the wrong hire for a company whose thesis assumes an AI-driven margin structure, whatever the title on the card, CHRO with a team of sixty or VP People with a team of eight. Not because they got worse. Because the job added a second half, redesigning how work gets done, and the assessment playbook has not moved with it.

These six are the delta, not the whole scorecard. The function still has to run, employee relations, compensation, compliance, retention under load, and your assessor already tests for that. What follows is what they do not. Each comes with the question that exposes it and the tell that separates real from rehearsed.

ONE: THEY READ WORK, NOT JUST PEOPLE

The AI-era people leader's raw material is the work itself, where it flows, where it stalls, what stopped being human. Ask: pick a function you supported, and walk me through how one deliverable moves through it today, and what changed about that flow in the past year. The tell: real answers name systems, handoffs, and one thing that got automated, with a number attached. Rehearsed answers name programs, partnerships, and relationships. A leader who cannot see the work will manage the org chart while the work moves out from under it.

TWO: THEY TOOK THE FIRST REP IN PUBLIC

Nothing in an enablement program outperforms a leader visibly working differently. Ask: what did you and your team automate out of your own work in the last six months, and can you walk me through the artifact? Not a live demo; a before, an after, and what broke in between. A real answer includes the story of being bad at it for two weeks, in front of people. Rehearsed answers say "my team uses it daily."

THREE: THEY ANCHOR THE PEOPLE PLAN TO THE THESIS

People strategy that cannot name the strategy line it serves is programming, however polished. Ask: take your last company's strategy, line by line. Which line did your people plan accelerate, and what work did you kill, or try to kill? The tell: killed work, named, and full credit for the kill that lost, because "I proposed ending the engagement survey, lost to the ops partner, and here is what it cost" is more diagnostic than a clean win. Anyone can list programs launched. The leaders who move a P&L can tell you what they stopped.

FOUR: THEY REDESIGN ROLES, NOT JUST FILL THEM

Ask: describe a role you redesigned because of AI. What stopped being human, where did the boundary land, and what happened to the person? The tell: a real case carries dates, a boundary that was actually published, and an honest cost. A hypothetical answer describes a framework. Then listen for the span question: as their org's throughput became partly machine-produced, what did they do to spans and layers, and why?

FIVE: THEY CAN HOLD A ROOM THAT HAS REASON NOT TO TRUST THEM

Most of this cohort will operate after a reduction, in rooms where "adopt AI and you will be safe" has already been said and already broken. Ask: walk me through the week after your last reduction. What did you say, what did you stop, and what did regretted attrition in your top decile do over the next two quarters? The tell: an honest capacity story, told plainly, and a

retention number they know cold. The leaders who held scared rooms know that number. The ones who hid behind the comms deck know the talking points.

SIX: THEY SHOW EVIDENCE A BOARD CAN AUDIT

Ask: what number about AI would you defend to a board, and would it survive an audit? The tell: real answers name workflows and system-measured cycle time or cost, and include something that was killed. Answers built on adoption percentages and self-reported hours saved are activity, not evidence, and a people leader who cannot tell the difference will spend your enablement budget on theater.

THE REFERENCE TWIN

These questions are rehearsable, and once this circulates, they will be rehearsed. That is fine, because each has a twin that cannot be coached: the same question, aimed at the people who lived it. Ask the CFO peer what the people plan killed and whether the savings survived. Ask a skip-level what actually changed about their job, and when. Ask the CEO what the people leader told them that they did not want to hear about the AI number. Run the six in the room and the twins in the backchannel, and the gap between the two answers is the finding.

WHAT AN INTERVIEW CANNOT TELL YOU

Everything above, questions and references alike, is still narration, and the leader who narrates well and the one whose organization executes are different people more often than anyone admits. The good assessment firms know this; the best already score your finalist against the value-creation plan, and if yours does, keep them. What none of them can do is read the work itself, because every input they hold is an account of the work by someone with a stake in the account.

The Executive Assessment closes that gap, in two modes, because they are not the same. Pre-close and on incumbents, all three input families run: intrinsics, a standardized instrument, because the trait this era actually prices, the tolerance for being publicly wrong while accountable for being right, is precisely what a polished interview conceals; outcome metrics by function, in that function's own currency, including the working-level distribution of the teams they ran; and work-function metrics, how work actually moved through their organization, read from the systems where work lives, every input invited, never passively collected. In an external search, no former employer opens its systems, and candidate-supplied telemetry is self-report with better formatting, so the instrument runs on what is legitimately available, intrinsics, verified outcomes, and the reference architecture above, and says so. An assessment that overstates its inputs is the same theater it claims to replace. Either mode is scored against your thesis, delivered deal-team-legible, findings labeled by confidence: Strong Signal, Emerging Pattern, Hypothesis to Validate. Run it alongside the assessor you already trust; it reads what interviews cannot. And the assessed leader receives their own findings, because the work-function read doubles as a first-90-days map, which is why strong candidates volunteer for this rather than endure it.

THE QUESTIONS A REAL ONE WILL ASK YOU

In a pool this small, the strongest candidates are assessing you back, and how you run this search is your pitch. Expect: what margin does the thesis assume from AI, by when, and do I own that number or does the CFO? What happened at your last portfolio company when the people leader said the AI number was not real? Is the mandate work redesign, or headcount reduction with better optics? Sponsors who can answer those three close this pool. Sponsors who cannot should fix that before the search opens.

ONE RUN, SANITIZED

Mid-hold at a portfolio software company of several hundred people, the sponsor was preparing an AI reset and had concluded the people leader could not carry it. Strong on the function, thin on the new half of the job, was the board's read, and a search was about to open. We ran the assessment instead of the search.

The interview layer agreed with the board: polished on programs, hesitant on work redesign, one visible AI initiative that on inspection was mostly communications. The instrument disagreed, in both directions. Intrinsics showed the profile this era prices, high experimentation tolerance and a low need to be publicly right, exactly what a defensive interview conceals. The work-function read, run with her participation from the systems where work lives, showed her function was one of only two in the company where a workflow had been rebuilt end to end: intake triaged automatically, cycle time down by roughly a third, one role redesigned and its boundary republished. What was missing was not capability but context: the CEO had never put the AI margin number in front of her, so she ran redesign as a side project and narrated programs, because programs were what the board asked about.

Finding, labeled Strong Signal: the constraint was the operating model above the seat, not the seat. The search did not open. She took her findings as a first-90-days map, the margin number moved into her goals, and two quarters later the second and third workflows were rebuilt, regretted attrition in her top decile held at zero, and the reset was on its number. The search fee was the cheapest thing we saved. The year was the point.

THE COST OF THE MISS

The math is unforgiving in a specific way: a mis-hire in this seat is visible by month nine, takes six months to re-search and six more to ramp, and the enablement program runs as theater the whole way through. Most theses can absorb that once. Run the six questions and the twins at your next search kickoff; they are yours. When the seat is on the line, run the instrument. Fifteen minutes of material, your questions after.