



PROOF BEFORE HIRE

Seeing What Candidates Can Actually Do, Before You Hire Them

A program from The 5-Star Approach

The Interview Illusion

Manufacturers across the country are running into the same wall: candidates who present well in an interview don't always perform well in the role. The interview measures how well someone talks about work. It does **not** measure how someone does work. This gap can be especially expensive for mid-level to senior engineering, logistics, analyst, and leadership roles, among others.

The scale of the underlying hiring challenge is well documented. NIST's Manufacturing Extension Partnership (MEP) reports that U.S. manufacturers will need to fill 4.6 million jobs by 2028, and that employee turnover costs typically run 50 to 150 percent of a departing employee's base salary. Those numbers describe a systemic problem: manufacturers are hiring more, competing harder for the same talent, and paying heavily every time a hire doesn't work out.

The Manufacturing Institute's own benchmarking research points to part of the fix: reducing reliance on subjective judgment in favor of skills tests, standardized scoring, and work samples, methods that let employers evaluate what a candidate can actually do rather than how well they interview.



The Science Behind Seeing Real Work



This isn't a new idea. It's one of the most thoroughly studied practices in employee selection research.

The U.S. Office of Personnel Management (OPM), which governs hiring practices across the federal government, describes work samples as tasks that mirror actual job duties as closely as possible. OPM's own guidance states that these assessments carry a “high degree of content validity” because they directly replicate job duties, and a “high degree of criterion-related validity” because performance on the task correlates strongly with actual job performance. OPM also notes that applicants tend to perceive work samples as fair, since the connection between the test and the job is obvious to them.

Dr. Elaine Pulakos reached the same conclusion in her landmark 2005 report for the SHRM Foundation, *Selection Assessment Methods*. Reviewing decades of selection research, she found that “work sample tests consist of tasks or work activities that mirror the tasks that employees are required to perform on the job,” and that they combine high validity with low adverse impact, a combination few other assessment methods achieve. It remains one of the most cited references in the field precisely because later research has not overturned its core finding.

The Society for Industrial and Organizational Psychology (SIOP), the leading professional body for the science of employee selection, sets the standard for how these assessments should be built and used: grounded in a thorough job analysis, validated against real job performance, monitored for adverse impact, and documented so the decisions behind them can be defended. SIOP's position is direct: without this structure, “the concept of merit is likely to be interpreted inconsistently, leading to favoritism, inefficiency, and subjectivity.”

Put simply: the research is not new, and it is not a small effect. Structured work samples, built around real job tasks and scored consistently, are among the most validated tools in hiring science. What's new is bringing that rigor into a paid, real-world, manufacturing-specific engagement, which is exactly what Proof Before Hire is built to do.

The Proof Before Hire Process

Proof Before Hire translates this research into a five-step process built specifically for manufacturing hiring decisions.

1. Define Success

Every validated assessment starts with a job analysis: a clear, specific picture of what the role actually requires. We sit down with your team, the hiring manager, key stakeholders, and often the people who will work alongside the new hire, to define what success looks like before we design anything.

This mirrors the approach behind Tooling U-SME's manufacturing competency framework, which connects specific roles and tasks to the knowledge, skills, and performance standards required to do them well. In manufacturing, success is rarely just technical competence. It can include how a candidate makes system-wide architecture decisions, creates supply chain resilience, offers cross-functional leadership, or executes root-cause problem solving under pressure. Those expectations have to be defined before they can be tested.

2. Create the Project

Candidates complete a paid micro-consulting engagement: real, valuable work drawn directly from the role itself, not a hypothetical quiz or a generic personality test. This is the step OPM's guidance describes directly: recreate the actual work environment and tasks “to the greatest extent possible.” It's also the step where Pulakos's research applies most directly, since a work sample's validity comes specifically from how closely it mirrors real job tasks.



Seeing Around Corners. A short, real engagement compresses weeks of on-the-job behavior into a few hours you can actually observe. How a candidate approaches an ambiguous problem, prioritizes competing demands, collaborates, communicates progress, and handles a deadline today is a strong, direct indicator of how they'll do the same things on your team tomorrow. You're not guessing about the future. You're watching a compressed, real version of it.

Proof in Practice

Manufacturers who move away from gut-feel hiring and toward structured, validated assessment are already seeing it pay off. In one documented case, a large packaging manufacturer adopted the Wiesen Test of Mechanical Aptitude to screen candidates for machine operator roles. The results were clear: employees who scored above the 50th percentile earned significantly higher performance ratings than those who scored below it (3.9 versus 3.4, on average), and every single employee rated a top performer had passed the assessment. The correlation between test score and job performance was statistically significant. The company adopted the test as its standard screening tool for every future manufacturing hire. We chose this example even though the roles are lower level, because it carried enough volume to demonstrate the statistical findings, which is harder to show with less frequent, specialized hiring.

Indeed, the case makes the core point: structured, validated assessment beats gut feel in manufacturing hiring, and the data proves it. Other industries have since taken the same principle a step further, replacing a test of aptitude with an actual paid sample of the work itself.

At Automattic, the company behind WordPress.com, CEO Matt Mullenweg has built the company's entire hiring process around paid trial “auditions” lasting three to eight weeks. Candidates work on real tasks alongside their future teammates, paid the entire time. Mullenweg has said plainly that being “well-spoken or charming in an interview often had little bearing on how a candidate would perform the job in question,” a finding that echoes exactly what the selection-science research above describes.

PostHog, a software company known for its transparent internal practices, runs a paid one-day “SuperDay” for engineering candidates: a real, scoped task, evaluated by a designated reviewer alongside the hiring team, with a hiring decision made within days. The task is deliberately built to be more work than one person can finish in a day, specifically so the company can observe how candidates prioritize and make trade-offs, the same behaviors a Candidate Score Card is designed to capture.

Gumroad takes this further still, hiring candidates as paid contractors for four to six weeks before making a full offer, giving them real access to the company's tools and live roadmap work. Founder Sahil Lavingia's reasoning is simple: “It's so very hard a priori to know if someone can do the job. So we just pay them to do it.” We note that if some of your candidates are employed at the time of consideration, this multi-week scenario won't work. This is why we shrink the tasks into a “side-hustle” sized chunk of work, still covering multiple days but not to interfere with candidate's current employment.

No one, in manufacturing or otherwise, has yet combined all three pieces into a single process: validated, job-specific assessment; a real, paid work sample instead of a synthetic test; and AI-assisted scoring that keeps evaluation consistent across every candidate. That combination, purpose-built for manufacturing roles, is what Proof Before Hire brings to the table.



3. Score What Matters

We build a Candidate Score Card tied directly to the outcomes defined in Step 1, so every candidate is measured against the same standard. These scorecards may measure demonstrated skills in executive decision-making, strategic trade-offs, or technical leadership in complex manufacturing environments, just to name a few. This directly reflects both SIOP's guidance on structured, standardized scoring and the Manufacturing Institute's own recommendation that companies use “standardized interview questions and scoring systems” to ensure candidates “are evaluated on the same basis.” Pulakos found the same principle holds for any structured evaluation: standardized criteria help evaluators judge quality and effectiveness consistently, rather than relying on impression or gut feel.

4. Build the AI Assistant

We build an AI assistant trained on your Score Card to apply it consistently to every candidate's work, every time. SIOP's published statement on the use of artificial intelligence in selection is clear that AI can support consistent, evidence-based judgment, but should never replace human accountability. That's exactly how Proof Before Hire uses it: the AI assistant helps evaluate the work quickly and consistently against your rubric. Your hiring manager makes the final call.

5. Implementation Guidance & Handoff

We teach your hiring manager how to run the process and make the final decision, and we stay through implementation to make sure it works in practice, not just on paper. NIST MEP's own guidance on manufacturing talent frames hiring as one part of a connected system: recruitment, training, onboarding, and retention all depend on each other. A single well-designed assessment doesn't help if it isn't implemented consistently. This step exists to make sure it is.



What This Whitepaper Does Not Claim

In the interest of giving you an accurate picture, not just a compelling one:

- The 5-Star Approach is not affiliated with, endorsed by, or certified by the Manufacturing Institute, NIST MEP, SME, the SHRM Foundation, SIOP, or OPM. We draw directly on their published research and guidance, and apply it to your context.
- The specific combination used in Proof Before Hire, a paid micro-consulting engagement, a manufacturing-specific Candidate Score Card, and AI-assisted evaluation, has not itself been independently studied or validated in published research. What is validated, extensively, is the underlying method: structured, job-relevant work samples scored consistently by trained evaluators.

We believe the strongest, most honest case for Proof Before Hire is this: it applies genuinely well-established hiring science, in a form manufacturers haven't seen before, to a problem costing them real money today.

Hire on Evidence, Not Hope

Every hiring decision is a bet. The question is whether you're betting on how someone interviews, or on what you've actually watched them do. Proof Before Hire is built to give you the second option.

To see how this would work for a specific role on your team, schedule a Proof Before Hire consult with The 5-Star Approach.



References

The Manufacturing Institute. Diversity, Equity, and Inclusion Benchmarking in Manufacturing (May 2023).

<https://www.themanufacturinginstitute.org/wp-content/uploads/2023/06/May-2023-Diversity-Equity-Inclusion-Benchmarking-in-Manufacturing-1.pdf>

NIST Manufacturing Extension Partnership. A Manufacturer's Guide to Finding and Retaining Talent.

<https://www.nist.gov/feature-stories/manufacturers-guide-finding-and-retaining-talent>

SME / Tooling U-SME. Best Practices in Learning & Development.

https://www.sme.org/globalassets/tooling-u-sme/resources/white-papers/11644_tu_sme_best_practices_in_learning_development_revised_09_14.pdf

SHRM Foundation (Elaine D. Pulakos). Selection Assessment Methods (2005).

https://www.shrm.org/content/dam/en/shrm/topics-tools/news/hr-magazine/assessment_methods.pdf

Society for Industrial and Organizational Psychology (SIOP). Statement on the Use of Artificial Intelligence.

<https://www.siop.org/wp-content/uploads/2024/06/SIOP-Statement-on-the-Use-of-Artificial-Intelligence.pdf>

Society for Industrial and Organizational Psychology (SIOP). Uniform Selection Guidelines Statement.

https://www.siop.org/wp-content/uploads/2025/12/UniformSelectionStatement_121525.pdf

U.S. Office of Personnel Management (OPM). Work Samples and Simulations.

<https://www.opm.gov/policy-data-oversight/assessment-and-selection/other-assessment-methods/work-samples-and-simulations/>

Harvard Business Review. The CEO of Automattic on Holding “Auditions” to Build a Strong Team (2014).

<https://hbr.org/2014/04/the-ceo-of-automattic-on-holding-auditions-to-build-a-strong-team>

PostHog. Engineering Hiring Handbook (SuperDay). <https://posthog.com/handbook/people/hiring-process/engineering-hiring>

Lenny's Newsletter. Adding a Work Trial to Your Interview Process (Gumroad, Linear, and others).

<https://www.lennysnewsletter.com/p/adding-a-work-trial-to-your-interview>