Training Metrics Aren't Capability Metrics

Idea In Short

A completion certificate proves someone attended a course, not that they can do the job on Monday, yet most organizations report completion rates as though they answered that question. The fix is not more training hours, it is a separate record, one filed against a person rather than a course, carrying a task, a date and the signature of someone qualified who watched the work done right. Industries where errors cost lives, aviation and powered industrial trucks among them, already run on demonstrated competence rather than logged hours. Leaders who want an honest answer can run a simple test this afternoon: count the people signed off to perform one critical task unsupervised, then compare that number against the completion rate for the course meant to cover it. The gap between those two numbers is the organization's real training report.

A completion certificate is mostly a dated attendance slip, sometimes with a quiz score that took three attempts. What it cannot tell a shift manager is whether the person on it can do the job on Monday, which is all the manager was asking. Most organizations report the slip anyway, as though it answered that question. That is not harmless bookkeeping. It helps explain why so many firms find their skills gaps in incident reports, long after the dashboard went green.

Forty hours per employee

United States [US] organizations with more than 100 employees spent $102.8 billion on training in 2025, nearly 5% more than in 2024. The hours went the other way, with the average employee logging 40, down from 47. If hours measured skill, the workforce would have shed a seventh of its competence in a single year. Nobody believes that, though nothing in those figures disproves it. Spend and hours describe inputs. Reading capability into them is like grading a hospital by the trays its kitchen sends upstairs.

Much of the profession knows this and measures the near end anyway. Roughly 80% of the

779 practitioners in a 2019 ATD survey said their organizations collected reaction scores and checked what people learned. The numbers drop fast after that, to 54% for on-the-job behavior and 38% for business results. Measurement thins out exactly where capability begins.

Changing the unit of record

Fixing this starts with what gets written down. A completion is filed against a course. A capability belongs to a person. Its record carries the worker's name, the specific task, a date and the signature of somebody qualified who stood there and watched it done right. Learning platforms often see only their own courses, which is why any decent primer on workforce skills tracking sends certificates and floor evaluations somewhere else. Set up like that, the record tells a night-shift supervisor which three people can run line four tonight, something no percentage can do.

What pilots and forklift drivers share

Industries where errors cost lives were among the first to act on this difference. In Europe, pilot training has been moving away from minimum hours and toward demonstrated competencies for the better part of a decade. Time logged at the controls proves that someone was present at the controls. Nobody earns a driving license by watching the theory videos, either.

Workplace safety law sometimes encodes the same principle. A forklift driver in the United States cannot be cleared on classroom time alone. The federal rules for powered industrial truck operators require classroom instruction and hands-on practice, then a workplace evaluation of the driver, repeated at least every three years. An accident or near miss sends the driver back for refresher training sooner. Many other jobs have nothing like that, and in those a plant operator or lab technician can be signed off with course records and little else. Leaders responsible for those roles could borrow the forklift structure, named evaluator and incident-triggered retests included.

How completion targets distort behavior

None of this makes training metrics worthless. In regulated work, a completion record is a legal fact, and an auditor asking whether every operator received the hazard briefing

deserves a yes with a date beside it. Trouble starts when that proof of delivery gets promoted, usually without anyone deciding to promote it, into a claim about readiness.

Once completion becomes a target, especially one tied to somebody's bonus, people find the shortest path to the green checkmark. The compliance video plays to an empty desk in a background tab. Quizzes get retaken until the right letter turns up, which on a four-option question rarely takes long. This is Goodhart's law at desk level, a measure that stopped telling the truth once people were paid to move it. A fourth-attempt pass does tell you something about persistence. It cannot show whether that person could diagnose a failing pump at three in the morning.

Smile sheets, as trainers call the end-of-course feedback forms, have a related weakness. A 2008 meta-analytic review of trainee reactions found the scores swing with how a course is delivered, especially the instructor's style and the amount of human interaction, more than learning outcomes do. A five-star review of a fire-safety session can share a file with a fumbled extinguisher drill a month later without any contradiction.

The research on transfer and forgetting

Researchers have a name for the step that completion records skip. They call it transfer, a dry word for whether anything from the course carries over to the job and persists there. Since the late 1980s, the research has kept pointing beyond the classroom, to the learner's motivation and the job they walk back into, especially whether anyone lets them try the new skill before it fades. That leaves much of the result with line managers, who rarely attend the course and are seldom asked, a month later, what changed.

One caveat cuts against this article's own case. The most repeated statistic in the field, that barely a tenth of training spending ever reaches the job, owes much of its academic standing to a 1988 review that took it from a 1982 article in a practitioners' journal. It has turned up in trade press coverage for years, sometimes with a passing nod to doubts about its basis. Critics of training metrics often lean on a figure with less behind it than the completion counts they mock. Numbers about learning deserve skepticism, the damning ones included.

There is a timing problem too. The certificate is dated on what is probably the learner's best day with the material. Memory slides downhill from there, a fact Hermann Ebbinghaus

pinned down in the 1880s by drilling himself on nonsense syllables. The method sounds quaint, but one of the authors of a 2015 replication in PLOS ONE repeated it on himself and got much the same curve. Job skills hold up better than nonsense, though not forever. A 2025 meta-analysis of procedural skills puts the loss at about half of the initial training gains, gone within six and a half months if the skill goes unused, at least on accuracy measures. Emergency shutdowns are the classic case, rarely practiced, with training records that stay green throughout.

Building capability records that hold up

Capability is mostly built somewhere other than the classroom. In Kolb's experiential learning cycle, a fixture of countless training slides, people try something, look back at how it went and adjust their thinking before the next attempt. A course can start that loop. It closes, if at all, on the floor, under the eye of someone who has done the job for years.

The practical version, for anyone running training or operations, is smaller than it sounds. A food plant on three shifts might start with lockout and allergen changeovers, two tasks where one bad day injures someone or stops a line. Someone writes down what doing each properly looks like, and a qualified assessor signs and dates it after watching. Lockout might then get rechecked quarterly, while the email-security module keeps its annual slot. Plants that tackle every role at once tend to bury those tasks under hundreds of trivial ones. Completion records still belong in the file. They simply stop being the headline.

The training industry has every incentive to keep the old scoreboard. Seats and completions are easy to sell and easier to invoice, and a provider paid per learner gains little from showing that half the learners never needed the course. The scoreboard is shifting, slowly, as corporate learning comes under pressure from finance teams to show outcomes before contracts renew. A finance department is an odd engine for honesty about human skill. It may still be the most effective one available.

Who signs for competence

The deeper obstacle is accountability. A capability record needs a named person to sign a judgment, that this employee can perform this task and was observed doing it on this date. A completion record rarely asks anyone to judge anything. When a signed judgment proves wrong, a supervisor answers for it, whereas a completion that proves wrong usually has no

author at all. Budgets tend to get the blame when companies avoid capability records, though this lopsided exposure may explain the avoidance better.

Any leader can run this test in an afternoon, though few will enjoy the result. Choose one critical task. Count the people signed off to do it unsupervised on each shift, then put that number next to the completion rate for the course meant to cover it. The distance between those two figures is the organization's real training report. Leaders who run the comparison now will have something worth defending at the next audit, and those who skip it may still learn the number later, from an incident report.

Summary

Training metrics and capability metrics answer different questions, and conflating them is not a rounding error, it is the gap that shows up later in an incident report. Spend, hours and completion rates describe what an organization delivered. A capability record, signed by someone qualified who watched the work done correctly, describes what an employee can actually do. Neither record is worthless, a completion log still satisfies an auditor's hazard-briefing question, and aviation and industrial-truck regulation show that competence-based verification works at scale once an organization commits to it. What changes the outcome is accountability, since a signed capability judgment has an author who answers for it, while a completion rate that proves wrong usually has none. Leaders who want the truth can get it in an afternoon, by comparing who is certified against who is actually signed off to work unsupervised.