

Ready, or Finished.

Completion tells you somebody reached the end of a course. It cannot tell you whether they can do the job you retrained them for.



Juliette Denny

Learning scientist · Author of *Learning to be Human* · Founder, Growth Engineering & Zavmo

The argument in one page

Reskilling has become the central promise of workforce strategy. Move people out of work that is disappearing and into work that is growing, and the organisation survives the transition intact. It is a good promise. The problem is the evidence we accept that it has been kept, which is almost always a completion rate. This paper argues that completion is not weak evidence of readiness. It is not evidence of readiness at all.

- ① **The stakes have moved.** Around **39%** of the core skills workers use are expected to change by 2030, and 63% of employers name skills gaps as the top barrier to transformation.¹ Reskilling is no longer a programme running quietly beside the business. For many roles it is the business continuity plan.
- ② **Completion cannot separate three different people.** The person who did not learn it, the person who learned it and cannot apply it, and the person who can now do the job all produce the same record: finished. A measure that cannot distinguish those three is not a weak measure. It is the wrong instrument.
- ③ **The failure happens after the course, where we are not looking.** Decades of research on the transfer of training find that what predicts whether a skill is used is largely what happens once the learner is back: the opportunity to perform, the support of the manager, the conditions they return to.³ Completion is recorded at exactly the moment before any of that has happened.
- ④ **We know this, and we measure it anyway.** **98%** of L&D functions say they want to measure business impact. **24%** have the budget.² The gap is not ignorance. It is that the harder measurement has never been funded as its own line.
- ⑤ **Readiness is demonstrable, and the standards to demonstrate it against already exist.** Britain maintains public descriptions of what competent performance looks like in an occupation.⁷ Judged work against those descriptions is evidence. A tick against a module is not.

What follows sets out what completion actually records, why the substitutes people reach for do not survive contact with the question, what readiness looks like when it is evidenced properly, and what it costs to start.

SECTION 01

What a completion rate really records

A completion is a timestamp. It says that a person opened a thing, moved through it, and arrived at the end. In most systems it means precisely that and nothing more, because that is all the underlying standard was built to carry.

This is worth stating plainly, because the number is usually presented as an outcome. SCORM, the standard most corporate learning still runs on, had its final revision published in 2009.⁶ Its data model was designed to answer whether a learner started, whether they finished, and sometimes what they scored on a quiz at the end. It was a reasonable design for the catalogue-and-course world it was built for. It is now being asked a question it was never built to answer.

The three people it cannot tell apart

Imagine three people who have each just completed the same reskilling programme.

The first did not really learn it. They moved through the material, recognised the right answers on the assessment, and retained very little. The second learned it genuinely and can talk about it well, but has not once done it in the conditions of the real job. The third can do the job.

All three appear in your report as one number. If the programme covered four hundred people, that number is 100% and it is presented to a board as a success. The distribution underneath it is unknown, and the organisation has no way of finding out except by discovering it in production, which is the most expensive place to discover anything.

39%

of core skills expected to change by 2030¹

66%

faster skill change in the most AI-exposed occupations⁵

531,200

skill-shortage vacancies, over a third of all vacancies⁴

The rate of change is the reason this has stopped being a philosophical objection. When the skills in a role turned over slowly, a generous view of completion cost you little. At the current rate, an organisation that cannot distinguish readiness from attendance is redeploying people into critical work on the strength of a timestamp.

SECTION 02

The substitutes, and why they do not hold

Everybody working in this field already senses that completion is thin, so several substitutes are in common use. Each is better than nothing. None of them answers the question.

The end-of-course assessment

A test at the end of a module measures recall and recognition under exam conditions, immediately after exposure. That is a real measurement of something. It is not a measurement of whether the person can do the work three weeks later, under interruption, with an ambiguous input and no answer bank. Assessment design can narrow that gap, but a knowledge test taken at the point of maximum recency will always flatter.

The satisfaction score

The happy sheet measures reaction, which Kirkpatrick placed at the first of four levels precisely because it is the easiest to collect and the least informative about outcome.⁸ It tells you whether people enjoyed it. Enjoyment is worth having and it is not capability.

The self-rating

Asking somebody to rate their own capability after training produces a number quickly. It also runs directly into the fact that the people least able to do a thing are frequently the least well placed to judge that they cannot. A self-rating is a good measure of confidence, and confidence and competence move together only some of the time.

The manager's impression

This is the best of the four, because a manager is at least watching real work. It is also unstandardised, uneven across a large organisation, and heavily dependent on how much time the manager actually spends observing. As the sole evidence for a workforce transition it does not survive an audit, and it does not aggregate.

None of these should be thrown away. Reaction data catches a badly delivered programme early. A manager's judgement is indispensable. The argument is narrower than it may sound: not one of them, and not all four together, tells you whether a named person can now perform a named task to a defined standard.

SECTION 03

Where reskilling actually fails

If completion were merely incomplete evidence, this would be a smaller problem. The reason it matters is that it is recorded at the exact moment before the thing that determines success has happened.

The research literature on transfer of training is unusually consistent on this. Across many studies, the strongest and most reliable predictors of whether a trained skill actually gets used are not features of the course. They are features of what the learner returns to: whether there is an opportunity to perform the skill, whether the supervisor supports its use, and whether the working environment makes using it possible.³

Read that alongside a completion rate and the mismatch is stark. We stop measuring at the end of the course. The evidence says the outcome is largely determined afterwards.

What that means in practice

It means a reskilling programme can be well designed, well delivered, enthusiastically received and completed by everybody, and still fail, because the people came back to roles that gave them no chance to use it for four months. It also means the reverse: a mediocre programme can succeed because the work was waiting and the manager insisted.

Neither of those outcomes is visible in the report the board is shown. Both are visible if you are capturing what people do after the course rather than whether they reached the end of it.

SECTION 04

What readiness looks like as evidence

The alternative is not a bigger survey. It is a change in what gets recorded, and the standards for it already exist.

Judged work, against a public standard

Evidence of readiness is a piece of real work, produced in the conditions of the actual role, judged against a description of competent performance that was not written by the person marking it. Britain maintains exactly those descriptions: National Occupational Standards for occupational performance, and a regulated qualification framework with levels, learning outcomes and assessment criteria.⁷ Using them costs a mapping exercise and buys a standard that survives a change of manager.

Verbs, not scores

The technical means also exists and is no longer new. xAPI, ratified as IEEE 9274.1.1 in 2023, records statements of the form actor, verb, object.⁶ It can record that a person analysed a case, produced a plan, evaluated an option or revised a decision after feedback.

The discipline is in the verbs. A system that emits "viewed" and "completed" in a modern format has changed its plumbing and not its measurement. The verbs that carry evidence of readiness are the ones near the top of Bloom's revised taxonomy: apply, analyse, evaluate, create.⁹ Those are the levels at which somebody demonstrates they can do the job rather than describe it.

Sampled, not universal, at least at first

The objection is always cost, and it is a fair objection. Judged evidence is more expensive than a tick. The answer is that it does not have to be collected for everybody on day one. A reskilling cohort of four hundred can be evidenced properly for the forty whose readiness carries the most risk, and the distribution that reveals will tell you more than the four hundred completions did.

98%

of L&D functions want to measure business impact²

62%

believe they are able to²

24%

have the budget allocated to do it²

Three moves, in order

Stop reporting completion as an outcome

Keep collecting it. It is a useful operational number and it tells you whether a programme is running. Move it out of the outcome section of the report and into the delivery section, where it belongs. This costs nothing and it changes the conversation immediately, because the outcome section is then visibly empty and somebody has to ask what goes in it.

Define readiness for one role before you build anything

Take the single role your reskilling depends on most. Write down what a person must be able to demonstrate to be ready for it, against the national standard for that occupation rather than an internal opinion.⁷ This is a week of work and it is the piece most programmes never do, which is why they end up measuring attendance by default.

Measure the ninety days after, not the last day

The transfer evidence says the outcome is decided after the course.³ So put the measurement there. Whether the person had the opportunity to perform, whether the manager supported it, whether the work was actually waiting. Those three questions, asked at ninety days, will explain more of your variance than any change to the course itself.

A CLOSING NOTE

The kindest thing is to know

There is a version of this argument that sounds like an attack on L&D, and it is not one. I have run these programmes. The reason completion became the measure is not laziness. It is that it was the only thing the technology could reliably capture, and once a number exists it gets reported, and once it gets reported it becomes the target.

What has changed is that the constraint has lifted. The standards exist. The means of capturing what somebody actually did, as they do it, exists and is ratified. The cost has fallen far enough that judged evidence is a budgeting decision rather than an impossibility.

And the case for doing it is not really about proof for the board, though it helps there. It is about the person. Telling somebody they are reskilled when they are not is not a kindness. It puts them into a job they have not been equipped for and lets them discover the gap in front of their new colleagues. Knowing who is ready, and who needs another eight weeks, is the more respectful answer as well as the more accurate one.

Where this thinking is being built

This is the problem I am building **Zavmo** to solve. It teaches against national occupational standards and the regulated qualification framework, and it records what a person actually did rather than whether they reached the end. The evidence accumulates as a by-product of the learning, per person and per skill, which is what makes readiness something you can see rather than something you assume.

If you are running a reskilling programme and cannot currently tell readiness from attendance, that is the conversation I would like to have.

Read more and get in touch at juliettedenny.com · zavmo.ai



About the author

Juliette Denny is a learning scientist and the author of *Learning to be Human*. She is the founder of Growth Engineering, an award-winning learning-technology company, and of Zavmo, an AI-native learning companion built on the science of how people actually learn. Her work sits at the intersection of pedagogy, behavioural science and technology, with a single throughline: designing learning that treats people as capable of growth.

REFERENCES & SOURCES

- 1 World Economic Forum, *Future of Jobs Report 2025* (2025). 39% of core skills expected to change by 2030; 63% of employers cite skills gaps as the top barrier to business transformation.
- 2 GP Strategies & Watershed, *Measuring the Business Impact of Learning* (2025). 98% of surveyed L&D functions want to measure business impact, 62% believe they are able to, 24% have budget allocated to it.
- 3 Blume, B., Ford, J. K., Baldwin, T. & Huang, J., *Transfer of Training: A Meta-Analytic Review*, Journal of Management (2010). Opportunity to perform, supervisor support and work environment are among the most consistent predictors of transfer.
- 4 Department for Education, *Employer Skills Survey 2022* (2023). 531,200 skill-shortage vacancies, 36% of all vacancies; 1.72 million workers affected by skills gaps within their existing workforce.
- 5 PwC, *2025 Global AI Jobs Barometer* (2025). Skills sought by employers changing 66% faster in the occupations most exposed to AI.
- 6 IEEE 9274.1.1–2023, *Standard for Learning Technology: Experience API (xAPI)* (2023), and Advanced Distributed Learning, *SCORM 2004 4th Edition* (final revision 2009).
- 7 Ofqual, *Register of Regulated Qualifications*, and the UK's National Occupational Standards, maintained by sector bodies. Public descriptions of qualification content and of competent occupational performance.
- 8 Kirkpatrick, D., *Evaluating Training Programs: The Four Levels* (1994; framework first published 1959). Reaction, learning, behaviour, results.
- 9 Anderson, L. & Krathwohl, D. (eds), *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives* (2001). The revised taxonomy and its active verbs.
- 10 Skills England, *Annual Skills Report 2026* (2026). "The education pipeline alone will be insufficient to meet employers' needs"; employers asked to demonstrate the return on their training investment.

© 2026 Juliette Denny. This white paper may be shared and cited with attribution. The statistics herein are drawn from the sources above; figures were current at the time of writing (September 2026). juliettedenny.com