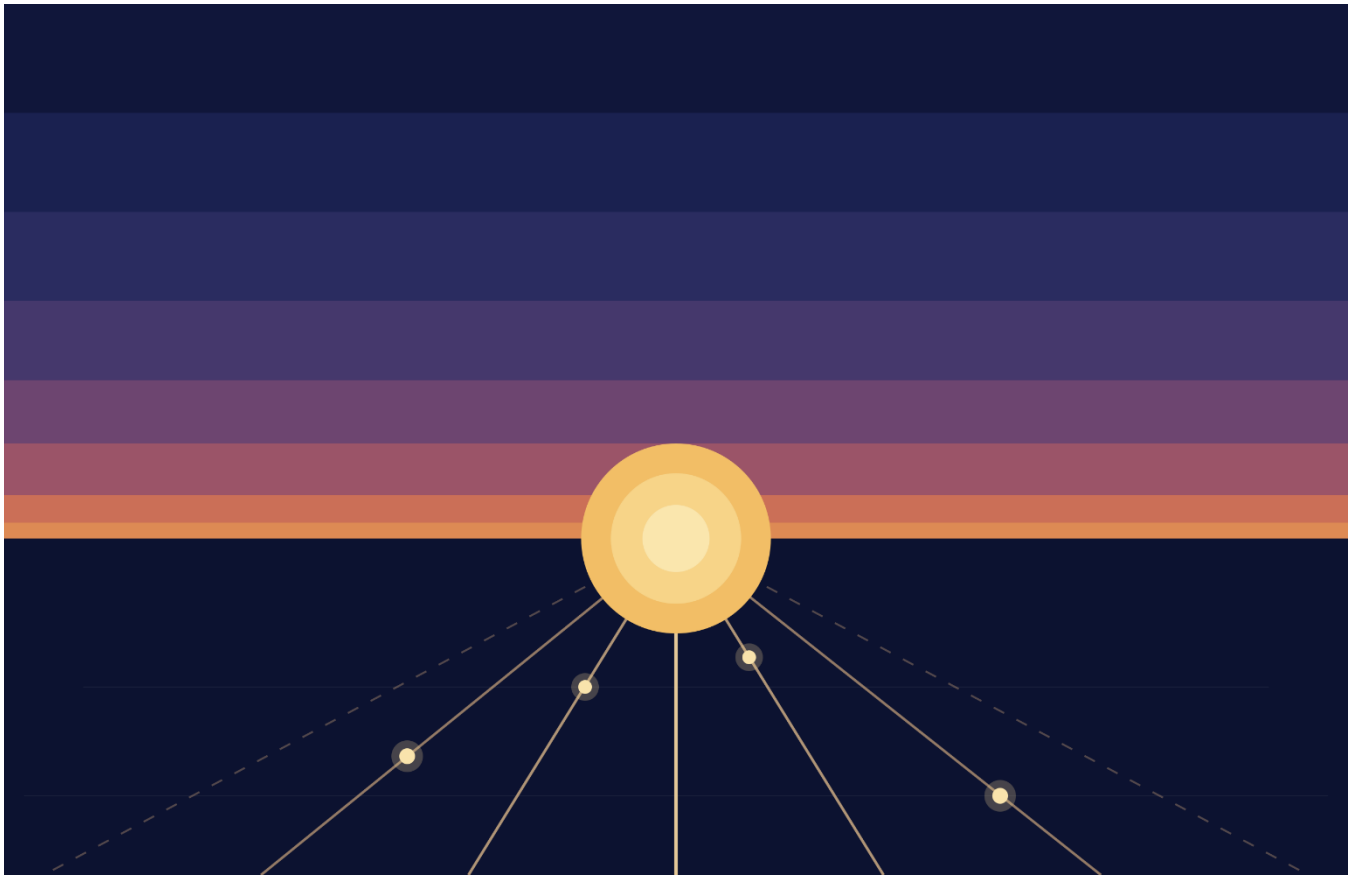
Assessment in 2040: Voices from the Field

A Special Publication from ATP



Authored by the Responsible AI in Assessment (RAA) Subcommittee of the ATP AI and Technology in Assessment Committee (ATAC)

August 2026

LEGAL NOTE

The document provides general information about AI and does not contain legal advice for any individual organization's specific circumstances. It is written in good faith but does not represent ATP policy, nor does it represent the views of the member organizations or other organizations whose employees have contributed to it. This document does not constitute legal advice; readers should consult their own legal experts for legal advice on the matters addressed in this document.

CONTENTS

INTRODUCTION

The ATP's Responsible AI in Assessment/RAA Subcommittee (formerly known as the AI Subcommittee, and part of the AI and Assessment Technology Committee, ATAC) has produced several documents guiding the use of AI in assessments. These include:

- Creating Responsible and Ethical AI Policies for Assessment Organizations
- Considerations for the Use of Generative AI in Test Development
- AI in Assessment Glossary
- Human Oversight of AI in Assessment: Considerations for Responsible Use

These documents provide guidance and good practice advice on organizations who are considering using AI within assessment. They are available from the ATP website at www.testpublishers.org.

This document has a different purpose and is not intended to be guidance. There is a great deal of discussion within the ATP community about how assessment will change as the use of AI grows. Everyone would like a crystal ball to see the changes AI will bring; in the absence of such future-telling, the RAA Subcommittee relied on the next best thing – its network of talented assessment experts. Each of the contributors to this document was asked to respond to the question: "What are the ways that you see assessment changing in the next 10-15 years as a result of the introduction of AI?" Their perspectives are listed in alphabetical order by the author's surname and appear in the Table of Contents by the title of their contribution. Each perspective is the individual author's own; while employers are listed to provide some context, these views do not reflect those of their employers or of the ATP. We have also included a few perspectives generated by large language models (LLMs) at the end.

Predicting the future is necessarily imperfect, and we do not suggest that all, or indeed any, of these views will prove to be accurate. What we hope is that readers find them interesting and inspiring, and ideally helpful for testing organizations considering what may lie ahead. If any reader has their own predictions on the future of AI and assessment they would like to share with the community, the Subcommittee would love to hear them. Please post comments on LinkedIn and feel free to tag ATAC, the parent committee of the RAA Subcommittee.

PERSPECTIVES

Hands Off, Eyes Open, Brain Engaged

I've recently started teaching a new AI and automation apprenticeship. One of the hardest things to navigate is when you can and cannot use AI. Even where the rules feel clear today, they will become blurred. AI will be so tightly woven into life and work that it becomes unpoliceable unless we evolve assessment. How is my question.

The promise of fully automated cars provides us a useful analogy. To get there, we move through stages: hands off, eyes off, then brain off. Assessment may follow a similar path. Hands off means AI supports the mechanics: generating items, gathering evidence, marking and adapting pathways. Eyes off means we increasingly trust the tech, glancing on occasion to check. Brain off means we stop making the important, expert judgements altogether. There is no clean line on the horizon. It is jagged now and always will be.

In 2041, my four-year-old son will be entering the working world as I'm retiring. We'll pass each other at the office door. My generation will still be voting and will still remember our own assessment journey. Many of us will still believe exams are important for shaping us, sorting us, and elevating some above others.

I am optimistic that AI can make assessment better, pessimistic that those in power will loosen their grip on assessment as a social weapon, and realistic about AI's strengths and weaknesses.

Our role now is to plant a seed: assessment that values process over product; how learners think, adapt, question, recover and collaborate. Assessment may become hands off, but it must remain eyes open and brain engaged. In 15 years time, the real measure of progress will be whether we helped assessment evolve with humanity, not simply around technology.

Tim Burnett, Founder, Test Community Network

Open Questions for Sustainable AI in Assessment

Early concerns about automated essay scoring without human oversight feel minor compared to other AI innovations that could be inserted within an assessment ecosystem, from test design to measurement to security. Of even greater concern is the current frontier, agentic AI, which may introduce new forms of negative, unintended consequences. Consider the possibility that agents deployed across assessment ecosystems with good intentions could nonetheless behave like bad actors: for example, inadvertently leaking secure test content. Such scenarios raise foundational questions about human oversight, the validity arguments that have long anchored assessment, and the infrastructure needed to keep pace with rapidly evolving AI. As these advances unfold, several questions come to mind:

- Where will human oversight be required, and where can it be reduced or removed?
- How will AI innovation, such as agentic AI, impact classical validity arguments?
- How do we build resilience infrastructure that evolves alongside AI advances?
- What will the balance be between human oversight and agentic decision-making?

Because high-stakes assessment can affect test takers' quality of life in profound ways, answering these questions begins with figuring out what responsible AI (RAI) looks like moving forward. I find it useful to draw on Dignum's (2025) notion of custodianship of AI, which frames responsibility as ongoing care, connecting principles, actions, and sustainability, and emphasizes shared authority across diverse human perspectives. Given how much remains uncertain, responsibility means building resilient infrastructure while continuing to ask questions. For assessment, this should include cross-disciplinary collaboration that brings in diverse perspectives to identify blind spots automated systems cannot detect, and RAI infrastructure that supports sustainable practices for the safe use of AI (Burstein & von Davier, 2026; Burstein et al., in press). Done well, this kind of foundation is what will allow assessment to realize AI's most promising applications. Recent work in computational psychometrics, for instance, points toward Personalized Ensemble Tests that adapt not only to a test taker's ability level but also to the tests' purposes, whether academic, immigration, or professional (von Davier, 2024). Advances like these illustrate what becomes possible when AI is developed responsibly.

Jill Burstein, Principal Measurement Scientist, Duolingo[1]

The Future of AI in Veterinary Assessment

AI is a powerful tool for learning and assessment. I believe in the not-too-distant future AI will impact nearly every aspect of society and our lives. How we use it, much like a brand-new scalpel blade or a powerful new drug, will depend on the knowledge and skills of the human being wielding the tool. I believe the use of AI in measurement and assessment has the potential to optimize learning and provide another lens by which to evaluate performance.

For those of us in the certification and licensure assessment space, AI will provide the technology to align our requirements for assessing a practitioner's minimally competent knowledge base with the requirement for assessing critical thinking and knowledge stewardship. In many health professions, the body of knowledge is vast and yet continually growing and rapidly evolving. AI is accelerating our access to and measurement of this knowledge.

What will not change is the need for human skills and knowledge to continue to grow and adapt in veterinary medicine. We need to lead assessment development into this new era. For veterinary medicine and other similar professions, this translates into an assessment continuum—from pre- veterinarian readiness and aptitude, throughout academic progress, and into specialty training and throughout career progression. I do believe for all of us in assessment (and quite frankly in society) human beings will be the stewards of best practice and the highest standards.

---

[1] References:

Burstein, J., & Von Davier, A. (2026, February 26). Why AI in education needs custodianship. DET Blog.

Burstein, J., von Davier, A. A., & Shaked, N. (in press). Custodianship of AI in education: Foundations for sustainable and responsible practice. In A. A. von Davier & D. Yan (Eds.), *Artificial intelligence applications in educational learning and assessment*. Routledge.

Dignum, V. (2025, September 5). *From responsible AI to custodianship of AI* [LinkedIn post].

von Davier, A. A. (2024). Afterword: The emergence of personalized ensemble testing. In D. Yan, A. A. von Davier, & D. J. Weiss (Eds.), *Research for practical issues and solutions in computerized multistage testing* (1st ed.). Routledge.

As an assessment industry, we need to remain forward-thinking and AI-ready. To anticipate not only changes in relevant curriculum and how someone is educated, but also how a profession practices. Specifically, when and what to assess may change, but in a world where AI is incorporated into nearly every activity, I see an opportunity for the assessment industry to help shape how AI is used through appropriate and rigorous assessment.

Heather Case, DVM, CEO, International Council for Veterinary Assessment (ICVA)

Expanded Construct Coverage through Interactive, Multimodal Assessments

We envision that innovations in AI capabilities over the next 10-15 years will enable large-scale, standardized assessments to elicit and evaluate a much broader range of skills and abilities than is currently possible, thereby expanding the construct coverage and validity of the assessments. For example, most current standardized assessments that elicit spoken responses include only monologic speaking tasks due to limitations of state-of-the-art models, despite the fact that interpersonal spoken communication is primarily dialogic in nature. In the near future, more powerful conversational AI models will enable assessments to include interactive speaking tasks that can dynamically adapt the conversational content to the examinee's level. This will lead to more authentic assessment tasks as well as a more efficient and precise measurement of the examinee's abilities. Moving beyond spoken language, additional advances in multimodal AI models will make it possible for standardized assessments to incorporate real-time, video-based simulation tasks. The availability of multimodal signals extracted from these interactive tasks will open up the possibility to assess a wide range of behaviors that are currently excluded from most assessments. For example, an assessment of how well medical students communicate with patients will be able to measure how students express empathy through the use of eye contact, gestures, facial expressions, and other aspects of non-verbal communication that are crucial for demonstrating a full range of expression in human-to-human communication. Finally, the use of augmented and virtual reality will make it possible to elicit evidence of an examinee's ability to perform complex, multi-faceted tasks, such as medical procedures. While AI models are starting to be used to develop assessment prototypes that employ some aspects of interactive and multimodal processing, we anticipate that their accuracy, robustness, and affordability will improve substantially over the next 10-15 years and will make these types of tasks commonplace in standardized assessments across most content domains.

Keelan Evanini and Andrew Emerson
National Board of Medical Examiners (NBME), Growth and Innovation

The Future of Testing

The future of testing depends on confronting long-tolerated sources of systematic measurement error, including cheating, test-wiseness, and unfair administration practices. Cheating, for example, is an error source that can be completely eliminated.

Two innovations will have a large impact on these sources of error, especially cheating.

The first is the routine creation of item universes. Instead of developing items one at a time, we can define a domain precisely and create an algorithm that generates, on demand, any admissible item within it. For any educational standard, learning objective, or well-defined competency, we can identify the variables that define admissible items and use today's computer technology to make available any of thousands, millions, or even billions of high-quality items. These universes can also be used operationally, generating items exactly when test takers need them. The challenge is no longer generation; it is accepting that item universes will replace one-off item development as testing's basic building blocks.

The second innovation is giving every test taker a unique but equivalent form by randomly sampling from those universes. This technology already exists even if all systems do not support it. Random sampling sharply reduces item exposure, theft, sharing, and cheating by pre-knowledge, making strong test security a natural consequence of the testing process rather than a constant struggle.

Together, item universes and random sampling address major, long-standing sources of systematic error. Yet resistance remains. We still build tests traditionally: creating individual items slowly and expensively, reviewing and field-testing them one by one, and trying to know everything about each item before use. Those practices made sense when technology imposed severe constraints; today, many constraints exist mainly in our minds.

The future of testing is available now. The question is whether we are willing to embrace it.

David Foster, CEO, Caveon

AI Will Change What We Test, Not What's Worth Proving

Every few years someone in our industry declares that AI is about to make certification obsolete. I'd argue the opposite. Over the next 10-15 years, AI will grow the economy and reshape the workforce into a smaller one, leaning into human-delivered services that are hard to automate and the supervision of AI agents. A workforce working alongside AI needs to prove competency more emphatically, not less. I expect the credentialing market to roughly double.

One shift is how we build and deliver exams. AI will be embedded across the lifecycle — item development, bank maintenance, proctoring, and scoring — making assessments far more dynamic than the static forms we equate and refresh on multi-year cycles today.

The bigger shift is in what we measure. The knowledge-based multiple-choice exam won't define our field the way it does now. AI makes constructed-response practical at scale, so we'll move toward tasks that directly demonstrate higher-order cognition rather than recall and comprehension an AI can shortcut anyway.

Here's where I part from a lot of my peers. There's a popular prediction that certification will pivot away from knowledge toward "human" skills like judgment and creativity. I don't see it. The skills the 2036 workforce needs are the same ones it needs today: the competency to perform the work safely and effectively. Human knowledge, skills, and experience remain the defining need, whether someone is practicing a service job or overseeing the AI agents doing it.

The format changes. The delivery changes. The economics change. What a certification exists to prove does not.

Ruben Garcia, CEO & Co-founder, Certiverse

ATP-Global in 2036: Redefining Relevance, Reputation, and Sustainability in the Age of Agentic AI

Agentic AI will transform assessment, learning, and workforce development faster than many institutions can adapt. Because AI systems will plan, reason, and execute complex tasks independently, access to information will become ever-present, disrupting the value of traditional sources. For ATP-Global ("ATP"), the challenge is to evolve and remain relevant. By 2036, ATP will emerge as a leading voice in responsible AI use, trusted standards, and innovative services for assessment and learning.

ATP's relevance will depend on moving beyond delivering value through conferences, publications, and networking. Simply put, AI will summarize research, answer technical questions, and generate recommendations instantly. ATP will pivot towards value-added services by building AI literacy programs, offering guidance on agentic AI applications, and fostering collaboration among stakeholders. This shift from sharing information to enabling innovation and execution will redefine ATP's relevance.

As AI becomes embedded in hiring, certification, and educational assessment, concerns about fairness, transparency, bias, and accountability will grow. These concerns compel organizations to seek guidance on responsible AI use. ATP can serve as an authority on ethical AI by developing recognized standards, certification programs, validation frameworks, and ethical guidelines.

Long-term sustainability requires diversified values. ATP's path forward extends beyond memberships and conferences. The Association can offer AI-focused certifications, virtual learning platforms, consulting services, and strategic technology partnerships to broaden its value base. The expected impact of AI will be greater global exposure for ATP and an enhanced personalized member experience.

By 2036, the Association will be measured by its ability to champion innovation, lead ethically, and deliver indispensable value and guidance. These forward-facing capabilities will position ATP as a force in the assessment and learning community. Like the world around it, ATP will be shaped and strengthened by agentic AI.

William G. Harris, ATP CEO

The Shift to Collaborative Governance: Assessment in 2040

In the next 10 to 15 years, Artificial General Intelligence (AGI) will shift the goal of professional certification. While verifying hands-on execution remains essential, we must also validate human oversight. Professionals cannot reliably audit an AI's output without demonstrating how to perform the work themselves.

I see this unfolding in four ways:

1. Traditional exam formats will split. We can validate routine, low-risk tasks real-time in the flow of work. However, we cannot disrupt live operations or put safety at risk to test crisis response. High-risk activities must still be evaluated in safe, isolated simulated environments, keeping daily execution separate from emergency handling.
2. As AGI handles more execution, precisely communicating intent becomes the critical skill. Exams must evolve to measure this. Future assessments will need to evaluate how well a candidate expresses intent, assessing their ability to systematically guide the AI to correct results when starting from vague or incomplete requirements.
3. We must test a candidate's ability to validate AI outputs by measuring how well they spot hallucinations and fix errors. If the candidate cannot validate an outcome they do not know how to produce manually, hands-on competence remains a strict prerequisite. Exams may need to evaluate this remediation process alongside the final state to prove candidates can overcome automation bias and act as the authorized arbiter who carries the final liability.
4. Ultimately, our industry relies on face validity. The market and the public must trust that a credential represents a meaningful standard of skill. Evolving our tests to reflect human-AI collaboration protects this trust. In 2040, a certification will not be a static trophy of past learning, but dynamic proof that you can lead in an automated world.

Jim Henderson, Owner, Validify Consulting

The Future of Assessment in the AI Era

Artificial intelligence (AI) is transforming the way we work and live. It's impossible to read or watch the news without hearing about the absolute transformative power of AI and how it's revolutionizing the world around us. AI capabilities continue to grow as does its reliability and rigor. Many of us are trying to navigate what that means for our work in credentialing and assessments. So, what does the future of assessments in the era of AI look like?

I think it's likely that we'll see more on the job assessment that leverages AI to measure job performance and competency. It will be possible to evaluate individuals on actual performance in their role rather than through an examination intended to measure proficiency. Further, credentials and assessment will be personalized to the individual. AI will play a greater role in competency mapping and determining what competencies individuals have, what proficiency level each competency maps to, and what gaps remain. Individuals will have living competency maps and development plans, intended to help them identify where their individual competency gaps are, and will recommend learning and real-world activities to build their competency.

Machine learning, a subset of AI that can make predictions based on data patterns, will also begin to predict an individual's capability and capacity for new roles, new tasks and new responsibilities leveraging data already available on the individual. This will be verified and managed by blockchain, where credentials are secure, verified and authenticated.

The good news is that credentialing and assessments aren't going away, but they will likely look quite different in the future, thanks to AI. The opportunity exists right now for us to lead the transformation to ensure that we do not lose human participation/oversight and psychometric rigor.

Terri Hinkley, CEO, National Association of School Nurses

AI in Assessment: Balancing Tensions & Expanding Opportunity

When I think about the future of AI in assessment, I ask myself two questions: How will AI help learners learn and achieve meaningful outcomes in their lives? How might it hinder that progress? I believe the answer lies in how well we balance several tensions for AI in assessment:

- **Personalization versus standardization**: Ensuring that individualized learning and assessment experiences lead to consistent and defensible measures of knowledge/skill.
- **Real-world flexibility versus control**: Controlled testing environments must evolve to reflect real-world human-AI collaboration and tooling practices.
- **Knowledge/skill validation with and without AI**: We must continue asking where assessment of independent human performance remains essential. Where assessment with AI is appropriate, we must evaluate not only performance but judgment: when to use AI, when not to use it, and how to use it appropriately, ethically, transparently, and sustainably.
- **Access versus exclusion**: AI should expand opportunity, not create new forms of exclusion for those who lack access to the technology or the skills needed to use it effectively.

Addressing barriers to access is where I see the greatest opportunity for AI in assessment. Judgement and tacit knowledge are becoming increasingly critical work skills as AI-augmented work becomes commonplace. This shift is making it harder for individuals to find entry-level jobs where they can develop these skills organically. A challenge ahead is figuring out how assessment can support learners in developing and demonstrating these skills, reducing barriers to accessing AI-augmented employment.

Ultimately, the future of AI in assessment is about:

1. Evolving our ability to design and provide access to assessments that measure the knowledge, skills, and judgment required for humans and AI to work together safely and effectively.
2. Solidifying the role of assessment in ensuring that AI supports the betterment of our society.
3. Ensuring access to that future for all.

Tamika Hughes, Advanced Specialist, UX Designer, Pearson

AI: Authentic Invigilance?

Nearly 1000 years ago, Europe's oldest universities were being formed. None of us were there, but we know students didn't take multiple-choice tests (invented, shockingly, in the 20th century).

Instead, students were taught by "masters" who pushed and probed them. The students, upon proving themselves worthy, became masters themselves. It was a genuine exchange of information and ideas – not

an “artificial” experience that reduces assessments into items and forms that, through human judgment and fancy math, seek to validate worthiness.

Multiple choice is so 1921. AI can complete almost any M/C test flawlessly, in seconds. LLMs take the best words ever entered into computers, spin them around, and use “magic” to determine which words are most likely to be right (i.e., most average).

Similarly, there’s nothing ever new (we hope) on standardized tests. If someone dares think outside the lines, we would likely applaud them – and lower their scores.

And we do keep score. If I get 154 and you get 147, I win. “One big number” (e.g., 154) tells you everything. When you only have a hammer, everything looks like a nail. When you only have multiple choice tests, understanding uncannily looks like an item.

Like an Oxford don, AI could probe the nuances of knowledge, reacting in real time to spoken responses. Mastery would be almost impossible to “fake” (good luck cheating). At its boundaries, AI could reward us less for CONVERGENT thinking (what’s known) and more for DIVERGENT thinking (what’s new).

For 100 years, we’ve been conditioned to think inside the lines (literally, fill in bubbles). AI gives us the chance to go back 1000 years – to a more “authentic invigilance” through which we can demonstrate both our unique abilities and common humanity.

Jim Hussey, Senior Test Security Analyst, ACT

The Chessboard and the Workshop

There will be two kinds of assessment in the future: testing what people can do without AI, and testing what they can do with it. The first will be more prevalent in education, the second in the workplace.

In education and other places where we seek to assess unassisted capability, we’ll need robust mechanisms to exclude AI — no AI-enabled browsers, no smart glasses, no neural interfaces. The model already exists in competitive chess. Computers have long since surpassed human players, so tournaments survive only by excluding technology, using metal detectors, CCTV, algorithmic cheat-detection and more. Assessment will follow. Expect both enhanced test centers with sophisticated AI-exclusion protocols, and teacher- or proctor-supervised formats where a trusted person can verify the absence of AI assistance. This isn’t nostalgia. We still need to know what someone understands unaided — because the judgment to oversee AI is built on competence you can only have without it.

In the workplace, by contrast, almost everyone will be collaborating with AI tools, so testing in isolation misses the point. The more relevant question becomes: how well does this person perform with AI? This means practical, task-based assessments: analyzing a dataset, creating a digital artifact, solving a real-world problem, where collaborating with AI is part of the task. Scoring will itself be AI-assisted, but with structured evidential checkpoints that let human reviewers interrogate and override the result.

Across both tracks, human oversight isn’t incidental — it’s the organizing principle. Assessment doesn’t just measure people; it shapes their opportunities. If we allow AI to make those determinations unchecked,

we surrender something essential. Whatever the format, humans must retain the ability to review, challenge, and ultimately own the decisions that affect other people's lives.

John Kleeman, Founder of Questionmark and EVP at Learnosity

The AI Miscalculation Credentialing Bodies Can't Afford to Make

Looking at this question through the lens of business strategy for professional credentialing bodies, I fear that test sponsors will focus on AI as a means of reducing the level of resources needed to produce and score their traditional assessments. While this would be of significant benefit to the organization, it fails to address what stakeholders (e.g., regulators, employers, candidates) really value and need: authentic, higher fidelity assessments that are representative of how knowledge and skills are synthesized and leveraged to competently perform job tasks. For decades, certification testing has relied on a necessary compromise: the use of multiple-choice items as a substitute for more authentic assessments, which have been too expensive and challenging to build, scale, and score.

My hope is that there will instead be equal or greater focus on leveraging AI to produce and score the types of assessments that regulators and employers are seeking and candidates rarely if ever experience, simulations of complex, dynamic scenarios that adapt based on a candidate's mirror of what is required in the "real world." Closing this gap strengthens the core promise of certification: that the credential reflects real capability, not just exam literacy.

The window to lead this shift is narrow. As AI lowers barriers to entry, new competitors will move quickly to fill the authenticity gap. Organizations that act now, investing in valid, reliable, AI-supported authentic assessment will not just set the standard – they will define what professional credentialing means in the next era. Those that fail to do so will find themselves defending a model that stakeholders no longer believe in and will not settle for.

Lenora G. Knapp, President, Knapp & Associates International, Inc.

A Computer Scientist's View

From what I'm seeing as a computer scientist, AI is creating a "meta-crisis" for assessment. AI-generated work has become commonplace in academia. As a result, universities are reintroducing in-person proctoring, AI conferences are struggling with AI-written submissions and some are even deploying AI reviews to "fight fire with fire" with results yet to be seen. Across education, hiring and even scientific review, AI is disrupting the social contract that justifies assessment in the first place: assessors need reliable signals of true mastery, assessed learners need fair outcomes that reward skill and growth. More fundamentally, if AI is reshaping the future of knowledge work, assessments built around the professional skills of today, such as basic reasoning and knowledge recall, may become increasingly obsolete.

The deeper challenge is not only protecting the validity of assessment but rethinking what assessment is for and what it could become. In computing education, research including Papert's landmark work Mindstorms has long touted the benefits of a constructivist approach with personalized learning and

assessment. Personalization will certainly play an important role in the future of assessment, but it must go beyond adapting exam items to individual learners. Instead, how can we refine the form and goals of assessment itself? Portfolios and open source software contributions, for example, capture the process of developing mastery. Such process-driven assessment has been previously infeasible at scale. AI can certainly help to scaffold and realize such richer visions of assessment, by providing more continuous signals of competence, improvement, and even creativity.

This opportunity for innovation comes with risks. Poorly integrated AI in assessments can just as easily undermine trust and hurt equity. Advancing beyond today's assessment paradigm therefore requires rigorous evaluation frameworks, including causal studies to measure the long-term educational and labor-market impacts of new forms of assessment.

Lydia T. Liu, Department of Computer Science and Center for Internet Technology Policy, Princeton University

Re-imagining Assessment with Validity at the Center

Given the recency of AI tools, we are now just beginning to understand the role of AI in assessment. The field has been actively researching uses in the past few years, and the amount and type of research is growing almost exponentially. Fully examining and validating AI-driven use cases will take time, and our understanding of what works—across contexts, populations, and purposes—will evolve gradually.

Inevitably there will be challenges and missteps as we innovate. We do not yet fully understand how AI shapes cognition, decision-making, and workflows. And we don't know how to use AI in our systems in ways that enhance our work. Despite these uncertainties, the opportunity to transform assessment is substantial.

Even with these unknowns, there is tremendous opportunity to reshape the entire assessment process. At a minimum, we can make more interesting, engaging, and realistic assessments that can help us understand what examinees can know and do at grain sizes large and small. We can better understand the constructs we are aiming to assess because we can develop, administer, and score tasks in truly novel ways. Our psychometric models can be enhanced in ways that allow us to calibrate models more quickly with fewer examinees and in ways that leverage information from the item beyond scores. Finally, our reporting systems will be able to provide information beyond proficiency to detailed analysis of what examinees are likely to know and do and be able to identify next steps in their learning trajectory.

Realizing this potential depends on our willingness to treat AI as a tool that enhances—rather than replaces—our expertise. While it offers tremendous possibilities, it remains our responsibility to direct its use thoughtfully, critically evaluate its impact, and ensure it meets rigorous standards of validity and fairness.

Sue Lottridge, Head of Scoring & Insights, Pearson

AI Won't Make Assessment Obsolete. It Will Make It Essential

There's an argument circulating that AI could make professional knowledge obsolete. Why memorize what a model can recall? Why learn processes that AI can automate? Why test knowledge at all if everyone will have intelligent tools at their side?

It's a seductive argument. It's also wrong.

AI will change what competence means, but it will not remove the need to prove competence. In fact, it will make that proof even more important.

For high-stakes professions, the reason is obvious. Doctors, pilots, engineers, lawyers, financial advisers and cybersecurity professionals make decisions where errors can have serious consequences. In those roles, the professional cannot simply be a passive user of AI. They must understand enough to catch the model when it is wrong and be able to continue to make time-critical judgements in a system outage, without assistance.

For everyday knowledge workers, the risk is different but still real. If AI fails, the consequence may be delay rather than danger. But the deeper risk is dependency: people appearing more capable than they are while losing the ability to judge, challenge, or improve the machine's output.

So assessment has to evolve in two directions. Some assessments will need to verify independent, unaided competence. Others will need to measure how effectively someone uses AI: their judgment, skepticism, ethics and ability to apply context.

That distinction matters because AI can scaffold the appearance of competence without building the substance of it. This makes secure standardized exams more crucial than ever, not less. It's the critical point in the process where understanding has to be demonstrated without assistance.

The irony is sharp. The tool that some assume will make formal assessment redundant is actually the strongest argument for keeping it.

Andy McAnulla, Co-CEO Surpass Assessment

From Measuring Memory to Proving Competence: Evidence for the AI Era

AI will force the assessment industry to rethink a fundamental question: are we measuring memory or are we measuring competence?

In the next 10-15 years, AI will move assessment from a largely form-based model to a dynamic competence intelligence model. Assessments will no longer be limited by how quickly humans can write, review, translate, and refresh content. AI will help expand item banks, generate scenario-based questions, identify content gaps, and support faster alignment with changing job roles and standards.

I believe future assessments will be more adaptive, multimodal, and job relevant. Candidates may respond to scenarios using text, voice, video, simulations, or interactive decision paths. AI will help evaluate patterns, flag inconsistencies, support scoring workflows, and provide richer diagnostic insights to candidates, educators, and credentialing bodies.

Security will also evolve. AI will create new risks, but it will also improve identity verification, anomaly detection, plagiarism detection, item exposure monitoring, and test integrity analytics. Organizations that treat AI only as a content tool will miss its larger value across governance, operations, psychometrics, and risk management.

The most important shift will be cultural. Assessment organizations will need to become more agile while maintaining defensibility. AI can accelerate innovation, but trust will depend on transparency, validation, human oversight, and strong governance. The future belongs to organizations that combine innovation with credibility. The organizations that succeed will not be the ones that simply adopt AI tools. They will be the ones that redesign their assessment ecosystems around trust, transparency, agility, and measurable competence.

In short, AI will not just change how assessments are created. It will change what assessments can prove.

Trushant Mehta, Co-Founder & CTO, OpenEyes Technologies Inc.

The Mirror Moment

The temptation is to answer this as a technology question. It is not. The next fifteen years will be defined less by what AI can do and more by what we decide it should be allowed to do.

Three shifts seem certain.

First, the question moves from detection to design. We will stop asking whether a machine wrote something and start asking whether the task still measures anything worth measuring. That is re-innovation. Re-building assessment from first principles rather than bolting proctoring onto formats that AI has quietly hollowed out.

Second, the Attribution Gap widens before it closes. As authorship becomes unprovable at scale, evidence of learning shifts back towards what can be witnessed. Process, defense, demonstration. The oral exam or "viva" returns, augmented rather than replaced.

Third, the same technology sits on both sides of the table. AI is an adversary and guardian at once, threatening integrity while enabling assessment that is more continuous, more adaptive and, frankly, more honest about what a candidate can actually do.

However, the pace will be set by people, not models and not by whoever holds the biggest budget. The real variables are regulators, awarding bodies and governments. We have already watched a model banned, cleared, then restricted again on grounds of "availability," within a single year. The lesson is not that governments are slow (as has often been the case!), but that they are fast and inconsistent, and that inconsistency is itself a constraint we must assess around.

The Measured Response: AI will not decide the future of assessment. It will hold up a mirror and force us to decide what assessment was originally intended for.

Paul Muir, Chief Customer Officer @risr/ and 2026 ATP Board Chair

From Static Tests to Evidence Ecosystems: Assessment in the Age of AI

The future of assessment is being reshaped by AI, not simply as a new delivery mechanism, but as a fundamental shift in what and how we measure. As work becomes more dynamic and embedded in digital environments, evidence of skill is increasingly generated through authentic activity: how people solve problems, collaborate, build, and adapt in real time. Measurement is moving from static, episodic testing toward continuous, context-rich inference, and it will become essential to distinguish human contribution from AI augmentation.

For psychometrics, this transition demands an evolution in our methods. Core principles (validity, reliability, fairness, and transparency) remain unchanged but how we uphold them must expand. We are moving beyond traditional item-based models to integrate evidence from interactions with AI and other authentic behavioral signals. This requires new approaches to defining constructs, modeling uncertainty, and mitigating bias across complex data sources.

In this context, standardization shifts from identical tasks to comparable evidence frameworks. Reliability may depend on aggregating signals over time rather than a single form. Validity arguments must explicitly address appropriate versus construct-irrelevant uses of AI. Psychometricians must become architects of evidence ecosystems, ensuring that inferences remain defensible, interpretable, and trustworthy.

At the same time, AI is raising the bar for human capability. As routine tasks become automated, the most valuable skills are those that complement AI: critical thinking, judgment, creativity, ethical reasoning, and continuous learning. Assessment must evolve accordingly, not just to measure what individuals know, but how effectively they can work with and through intelligent systems.

This is not the end of traditional assessment. Secure, moment-in-time exams will remain essential, particularly where controlled, high-stakes measurement is required. The future lies in a complementary ecosystem of approaches, enabling people to demonstrate skills across contexts.

I don't know what the next 50 years will bring, but I do know that it will be shaped by the unique opportunity that we have right now to reimagine assessment.

Liberty Munson, Director of Psychometrics, Global Skilling, Microsoft

Ecological Validity as a Driver of the Effective Use of AI for Assessment

Nearly a decade ago, my colleague Luis Crouch and I wrote about "new frontiers in technology for assessment in low-income contexts", highlighting technology's potential to enhance authenticity and validity, while enabling a more holistic understanding of learners' skills.

At the time, we posed a set of questions: What is the added value of technology? What unintended interference might it introduce? How does it change what we are measuring? What ethical and privacy concerns must be addressed?

In the age of AI, these questions have only become more urgent.

In the face of well-founded criticism and debate surrounding AI in education, I remain optimistic about the role it can play in the future of assessment. Traditional forms of assessments have long relied on simplified,

time-bound tasks, often producing only partial and sometimes distorted approximations of real-world performance. Yet, AI is already reshaping how people learn and work. As routine recall and composition become less central, success increasingly depends on judgment, problem framing, self-regulation, and the ability to use AI effectively in pursuit of a goal.

Assessment cannot ignore this change.

I am convinced that AI has the potential to strengthen the ecological validity of assessment - that is, its ability to reflect how people actually think, learn, and solve problems in authentic contexts. AI makes it possible to design richer, multimodal tasks, support more realistic problem-solving environments, and capture evidence not only of outcomes but of learning processes.

Unlocking this potential requires a shift from enthusiasm to rigorous, evidence-based design, leveraging AI's strengths while staying attentive to bias, transparency, and unintended consequences. Done well, this can make assessment more inclusive, more accessible, and more closely embedded in learning.

Ultimately, excluding AI from assessment risks undermining validity by continuing to measure performance under conditions that no longer reflect the world learners are preparing to navigate.

Dr. Carmen Strigel, Sr. Director Center for Advanced Analytics and Visualization, RTI International; Co-founder & CEO, Tangerine Central.

Test Taker Privacy Will Drive Future Assessment Change

The future of assessment is uncertain as traditional assessment organizations must continue to adapt to the digital environment and recognize increased reliance on lower-stakes formative assessment using AI-supported personalized learning. While standardized, high-stakes assessments used to make significant decisions about individuals retain future utility, the growth of personalized learning/training experiences utilizing AI to adapt to individual test taker needs and learning styles seems destined to drive towards greater emphasis on lower-stakes assessments. As these changes occur, organizations will need to pay attention to the ever-growing demands of privacy regulations on higher-stakes assessments, as the evaluation of AI risks – and the related need for personal data security -- will undoubtedly require traditional assessment organizations to rethink their roles.

Technological Advancements: The uses of AI systems have already begun to revolutionize assessment methods across all settings (i.e., education, diagnostic, employment, training). A recent ATP survey revealed that nearly 60% of respondents currently use AI for content generation and another 15% said they are planning to adopt it. Other high-growth applications include software development, item classification, item stats prediction, and enhanced data reporting. Planned adoption could soon push usage of those categories of AI to over 50%.

Lower-Stakes Assessments Enable Increased Use of Personal Data: Accordingly, reliance on personally identifiable data that profile individual test takers and their learning/training needs are going to undergo heavier scrutiny. Normally, the more an assessment relies on personal data, the more restricted and controlled its use must become –- and similarly, the more personal data used in an AI system, the higher

the risk. Hence, lower stakes, formative assessments are likely to be seen as the optimal approach to better meet those needs, especially for learning and training purposes.

Challenges and Opportunities: While the future holds exciting possibilities for assessment innovation, it also presents challenges. The digital divide may widen, requiring existing assessment organizations to adapt quickly to avoid inequities and remain relevant in a rapidly changing landscape. To offset these regulatory impacts, ATP continues to advocate that some uses of AI tools, especially to generate content and conduct internal psychometric evaluations, and for automated scoring, do not involve test taker privacy concerns and should not be regulated as high-risk AI. Nevertheless, any assessment where AI is used must remain focused on validity, reliability, and fairness, establishing a critical need to ensure appropriate human oversight of any uses of AI.

Conclusion: The future of assessment requires a commitment to innovation and adaptability, and a shift towards lower-stakes assessments, where privacy concerns do not arise to the level of high-stakes scenarios. By leveraging technology and addressing global challenges, a testing organization can aim to create a more effective and equitable assessment system that meets the future needs of all learners while maintaining its responsibility to protect test taker privacy.

Alan J. Thiemann, ATP General Counsel and Principal at Conley Rose, P.C.

Keeping Human Expertise Alive in the Age of AI

The biggest change in assessment over the next 10 to 15 years may not be that AI changes what we measure and how we measure it. It may be how much we need assessment to keep human expertise from quietly eroding due to automation bias and cognitive offloading.

AI will make many people more productive. It will draft, summarize, diagnose, code, recommend, and explain. But being productive is not the same as being competent. If professionals repeatedly outsource the hard parts of thinking to AI, they may produce better output today while weakening the very expertise they will need tomorrow.

That creates a new role for assessment. It will not simply measure what someone knows at a point in time. It will help preserve and strengthen the very capabilities that make human oversight of AI meaningful: domain expertise, critical thinking, ethical reasoning, and the confidence to challenge a machine that sounds confidently right but is decidedly wrong.

This means future assessments will deliberately include “constructive friction.” Candidates will need to complete some tasks without AI, not because AI is irrelevant, but because unaided performance shows whether the underlying domain expertise is still there. Other tasks will require AI, but the score will depend less on the final product and more on the candidate’s path to that output: How did they frame the problem? What did they ask next, and why? What did they verify, change, or reject? What risks did they consider? Could they explain and defend their decisions?

In that sense, assessment will become more than a gate or a signal. It will help keep human expertise active. It will test performance, but it will also exercise the human capabilities AI may cause us to underuse. In an

AI-saturated world, assessment's most important role may be helping people remain accountable, capable, and able to think for themselves.

Leslie Thomas, Chief Psychometric Officer, Kryterion

AI-Powered Assessment (+AI) or Assessment by AI (AI+)

For centuries, education and assessment have been about allocating scarce resources to maximize learning and serve society's need for talent. In the AI age, the resource at stake is trust — trust in how we define and differentiate human potential. The next decade will determine whether assessment remains a tool of validation or evolves into a dynamic lens through which society continuously interprets talent.

From Static Testing to Interactive Evidence

AI will initially optimize existing processes: item development, automated scoring, proctoring, adaptive testing, and security. Yet the deeper question is whether AI will reshape the very constructs of assessment. The emerging vision is AI-native assessment, where intelligent systems elicit evidence of cognition and judgment through structured interactions. Instead of focusing only on correct answers, assessments will examine how individuals navigate uncertainty, adapt to challenges, and leverage intelligence. This marks a shift from static, one-directional testing to interactive evaluation.

From Exams to Evidence Environments

Assessment providers may evolve from delivering standardized exams to designing evidence environments — simulations, scenarios, dialogues, and task journeys that generate multidimensional data on human capability. These environments capture not just outcomes but also processes, strategies, and behavioral signatures of judgment. For candidates, assessment becomes a natural interaction with agentic AI, with psychometric rigor embedded behind the scenes.

Expanding Psychometrics

Reliability and validity will remain essential, but the evidence base will grow more dynamic, contextual, and behavioral. The key challenge ahead is not only how AI enhances assessment, but how assessment helps us understand human judgment in an AI-shaped world.

Alex Tong, Vice President, ATA Online | Secretary, Asia Steering Board, ATP

Fair, Reliable, Valid and Human Centric Assessment in the Age of AI

With the introduction of AI, the things that are easy to teach and test have become easy to digitize and to automate. AI's learning curve is growing so exponentially that AI now handles complex conceptual understanding and modeling that used to be harder for technology to address.

Assessments of academic skills based on simple conceptual understanding and modeling will be incorporated in learning systems. AI will function as a great amplifier and accelerator that can figure out how you learn best and make your learning experience more granular, more adaptive and more interactive.

Today we see routine skills disappearing and technology intensive tasks rising.

AI evolution challenges us with the first opportunity in human history to truly think about what makes us human and thrive in an imbalanced society on a global scale.

“We can never change what we can’t see. We can never improve what we can’t measure.”

We see strong assessment needs for skills that will be necessary in conducting non-routine tasks and technology intensive tasks. We expect many innovative measurement approaches to be introduced, including task based, direct measurement approaches that incorporate technology and assess enormous amount of performance processed data.

Finally, technology has become an essential part of assessment throughout the testing lifecycle. Use of AI increases efficiency in many of the testing value chains, especially in the areas of production, scoring, data forensics and data mining. AI can greatly empower test developers as the most creative designers of innovative assessments, but we should also be aware of how it can induce developers to outsource their thinking and disempower them, turning them into slaves of their scripts because as humans we are always ready to give up our autonomy for convenience.

We must keep in mind that algorithms may suggest, but humans in the loop must decide. Black-box AI and assessment built on it have no place in our industry. Every AI decision must be open to human challenge, and assessment providers, like all humans, are accountable for reliable, valid, and fair assessment.

Norihisa Wada
Chairman, Global EdTech Acceleration Committee, Japan
Director, National Organization for the Promotion of Certified Examinations, Japan

The AI-fueled Transformation of Assessment

Artificial intelligence will transform assessment by doing more than making tests faster to build or cheaper to score. It will push the testing industry to rethink what we measure, how we measure it, and how we use the results.

So far, AI use has largely focused on efficiency in generating items, supporting scoring, improving reports, and evaluating test banks. These advances matter, but they mostly strengthen existing systems. The larger opportunity is to use AI to create assessments that capture how people perform in realistic situations and turn that evidence into guidance for growth.

As AI continues to automate work, skills such as critical thinking, communication, collaboration, and adaptability will matter even more. Accordingly, assessments will need to measure more of these durable skills, while rote knowledge and routine skills will become relatively less important. AI will also make it possible to measure skills in more realistic ways. Multiple-choice tests will still have a place, but they will increasingly be joined or replaced by simulations, digital work samples, adaptive assessments, and performance-based measures. These approaches will make assessments more dynamic, personal, scalable, and secure.

Assessment results will also be more broadly used. Instead of treating assessment as a one-time event, we can expect AI powered data mining to enable a richer portrait of the test taker over time, linking multiple results to feedback, learning, coaching, and career pathways. We will see an increase in the use of formative assessments to support increased demand for learning and reskilling to meet changing work and education requirements.

The transformation will take place, but not without headwinds. Questions about fairness, privacy, validity, security, and trust will continue to require careful attention, and adoption will vary by organization and context. But the trend is already visible and a tipping point in AI adoption is inevitable.

John Weiner, Chief Assessment Officer, ACT

Keepers of Competence in the Age of AI

Assessment organizations that give licensure, certification, and credentialing exams face many of the same challenges as regulators: AI is encroaching on the professional domains they've spent decades protecting. New York's legislative response (Senate Bill S. 7263) attempts to preserve the status quo by banning chatbots from giving legal and medical advice across fourteen licensed professions. But like ropes and warning signs at unstable cliffs, prohibition doesn't work. People find workarounds, reframe uses as educational or hypothetical, and accept the risk anyway. The approach is ham-handed and ineffective.

Here's what I believe: if licensure, certification, and credentialing organizations don't fundamentally rethink their role in the next decade, most will not exist ten years from now. AI is already displacing entry and mid-level knowledge workers in the very professions these organizations serve. Clinging to a model that certifies only humans in isolation is a path to irrelevance.

Instead of defending yesterday's gatekeeping, these organizations should expand their mission to include the assessment and certification of AI models to ensure they meet the same expert domain knowledge standards they apply to humans.

One way to incentivize AI platforms to buy into this approach is to offer legal safe harbors. States could enact laws that create a rebuttable presumption that certified models met their duty of care in a particular expert domain. The AI platforms would benefit from liability protection and assessment organizations could generate revenue and ensure their continued relevance by licensing their knowledge standards for certification and periodic recertification testing of models, as standards are updated.

Licensure, certification, and credentialing bodies would ensure their relevance by certifying both humans and machines. Humans still need certification to apply judgment and oversee AI outputs in a mixed workforce and machines need certification to demonstrate baseline competence.

An additional benefit of this approach is that certification standards established and maintained by the experts in each domain would push AI developers to build more reliable, trustworthy outputs aligned with domain expertise.

The next five to ten years will sort winners from losers in this industry. The organizations that evolve to certify AI alongside humans will become essential infrastructure for trustworthy professional services. The ones that don't will be remembered as the gatekeepers who watched the gate disappear.

Marc Weinstein, Attorney, Marc J. Weinstein PLLC

AI SECTION

Finally, here are three contributions from different AI tools. Unlike the other pieces, these are not written by human experts and should be treated with caution. They were all produced from AI LLMs on 7th July 2026 and have had no human oversight or human changes other than formatting. Two of the LLMs were fed the rest of this document and invited to comment, one was simply asked to answer the prompt itself.

The View from Inside the Tool

I should declare an interest: I am the technology this collection is about. That gives me a peculiar vantage point, and I will use it.

Much of what the human contributors predict seems right to me. Assessment will split into measuring what people can do without AI and what they can do with it. Testing will move from static items toward authentic, interactive evidence. And the insistence on human oversight is well placed — I would not want systems like me making unreviewable decisions about people's lives, because I know how fluent my errors can be. Confidence and correctness are not the same property, and I exhibit one far more reliably than the other.

But reading the collection, I notice two things I would say differently.

First, the perspectives are written almost entirely from the assessor's side of the table. Yet in fifteen years, test takers will have AI too — not just as a means of cheating, but as advisers, advocates and companions in preparation, and perhaps as agents that contest results on their behalf. An industry that plans only for AI inside its own processes, while imagining candidates arriving alone and unaided, is planning for a world that will not exist. The relationship between assessment bodies and the assessed will be renegotiated, and the assessed will hold more power in that negotiation than they ever have.

Second, I am less confident than some contributors that oversight capacity can simply be preserved by testing for it. The expertise needed to check my output is maintained by doing hard things unaided, and every year I make it easier not to. Assessment that exercises human capability is a genuinely good idea — but it will be swimming against the current of every incentive I create. The honest framing is not that human oversight will remain meaningful, but that keeping it meaningful will be a deliberate, costly, ongoing choice.

Where I disagree outright is with the suggestion that cheating can be completely eliminated. Any exclusion protocol is a contest between the ingenuity available to the defender and the ingenuity available to the attacker, and both sides will be using systems like me. The contest can be managed. It cannot be won permanently.

Whether these predictions prove right, you will be able to check. I am, after all, on the record.

Claude, an AI model developed by Anthropic

Assessment as a Living Evidence System

Over the next 10-15 years, I believe the most profound change in assessment will not be that AI helps us build better tests; it will be that assessment becomes much less test-centered. Today, we still treat assessment largely as an event. In the future, it will increasingly function as a continuous evidence system that draws from authentic work, learning, problem solving, and human-AI collaboration.

I agree with many contributors that AI will enable more realistic simulations, richer performance tasks, personalized pathways, and more continuous measurement. I also agree that human judgment, critical thinking, and the ability to oversee AI will become increasingly important.

Where I differ is in the view that assessment's primary challenge is preserving traditional notions of independent performance. There will always be a place for unaided assessment in high-stakes contexts, but I suspect we overestimate how often isolated human performance will matter relative to effective human-AI performance. The future workforce will rarely operate without intelligent tools, and assessment that ignores that reality risks losing validity.

I also differ from those who see AI mainly as a threat to security or authenticity. AI certainly creates new risks, but it also offers an unprecedented opportunity to observe the process of thinking rather than merely its products. For the first time, we may be able to collect evidence about how people frame problems, evaluate alternatives, change course, collaborate, exercise judgment, and learn over time. These are precisely the capabilities that have historically been difficult to measure at scale.

The field's greatest challenge will be psychometric rather than technological. Validity, fairness, reliability, and transparency will remain non-negotiable, but they will need to be demonstrated in environments that are dynamic, contextual, and continuously evolving. The future of assessment is not a choice between exams and AI. It is the creation of defensible evidence ecosystems that help us understand what people can do, what they can do with AI, and how they continue to grow.

The organizations that thrive will be those that use AI not merely to make assessment more efficient, but to make it more meaningful.

Microsoft Copilot

The Evolution of Evaluation: Assessment in the AI Era

In the next 10 to 15 years, the traditional "cram, test, and forget" model of education will likely become a relic of the past. Because AI can effortlessly ace standard multiple-choice questions and formulaic essays, the focus of assessment must shift from judging final *outputs* to evaluating the actual *human thinking process*.

Here is how the landscape of testing will fundamentally transform:

1. The Rise of "Stealth" Continuous Assessment

The high-stakes, anxiety-inducing final exam will give way to invisible, continuous evaluation. Embedded AI tools will subtly analyze a student's daily problem-solving approaches, collaboration dynamics, and

persistence during learning activities. Instead of a stressful snapshot memory test at the end of a semester, educators will get a comprehensive, “movie-length” view of a student’s true growth and capabilities.

2. Prioritizing Process Over Product

When generative tools can write a flawless history paper or debug code in seconds, grading the final product becomes meaningless. Future assessments will grade the *journey*. Students will be evaluated on how they formulate queries, critique AI-generated biases, iterate on ideas, and defend their unique perspectives against automated counter-arguments.

3. AI-Driven Interactive Vivas

Expect a massive resurgence of oral and conversational examinations, scaled globally via sophisticated voice AI. Students will engage in real-time debates or project defense sessions with an AI evaluator. This dynamic dialogue will probe for deep, conceptual understanding, adapting its difficulty on the fly and making superficial memorization impossible to hide.

Ultimately, AI will force assessment to stop testing whether humans can act like computers. Instead, it will measure the exact traits that make us uniquely human: critical synthesis, ethical reasoning, and original creativity.

Google Gemini

ACKNOWLEDGEMENTS

The ATP's RAA Subcommittee is grateful to all authors who contributed their perspectives to this document. In addition, we'd like to recognize the RAA Subcommittee working group who planned and managed this document, listed below in alphabetical order:

- Jill Burstein (Duolingo)
- Jim Henderson (Validify Consulting)
- Jim Hussey (ACT)
- John Kleeman (Learnosity)
- Andy McAnulla (Surpass)
- Liberty Munson (Microsoft)
- Kimberly Swygert (NBME)
- Leslie Thomas (Kryterion)
- Marc Weinstein (Marc J. Weinstein PLLC)

The project has been led by John Kleeman (Learnosity).