Can AI Replace Teachers? A Conceptual Analysis of Instructional Capabilities and Limitations

Sasitharan A/L Manokaran, Yoharani A/P Francis

Faculty of Business, UNITAR University College Kuala Lumpur, Malaysia

DOI: https://doi.org/10.47772/IJRISS.2026.101000578

Received: 07 October 2026; Accepted: 12 October 2026; Published: 11 November 2026

ABSTRACT

Generative artificial intelligence has renewed debate about whether technology can replace teachers. This conceptual article examines that question by distinguishing the performance of instructional tasks from responsibility for the wider educational role. A theory-informed synthesis of selected research and institutional guidance is used to develop a framework for allocating teaching responsibilities in higher education. Technological pedagogical content knowledge provides the theoretical foundation, while human-centred AI guidance informs the governance perspective. The framework proposes that suitability for delegation depends on task boundedness, evidence of learning, error consequences, learner readiness, and the availability of professional oversight. AI may support explanation, practice, and preliminary feedback, yet success on these tasks does not establish equivalence across assessment judgment, contextual adaptation, relational support, and institutional accountability. Recent experimental research illustrates both the potential of carefully designed AI tutoring and the risk of confusing assisted performance with independent learning. Six conceptual propositions and a task-allocation matrix are presented. The article recommends evaluating AI-supported arrangements against explicit learning and accountability criteria rather than assuming either universal replacement or universal human superiority. No primary data is reported. The contribution is a conditional account of when AI may substitute for instructional activities and when a teacher-led arrangement remains warranted.

Keywords: artificial intelligence; teacher roles; higher education; instructional delegation; TPACK; independent learning; human oversight

INTRODUCTION

The question of whether AI can replace teachers appears straightforward until the meaning of replacement is examined. Teachers explain concepts, design practice, assess achievement, organise learning environments, respond to uncertainty, and exercise professional judgment. An AI system that performs one of these activities successfully has demonstrated a task capability. It has not necessarily demonstrated the capacity to assume the full educational role. This distinction matters in higher education because institutions may judge innovation through access, speed, and cost while students experience it through the quality of their learning. A quick explanation may be useful, but educational success also involves retaining knowledge, transferring understanding, and recognising an incorrect answer. Similarly, an automated assessment score may be convenient without being sufficiently defensible for a consequential decision.

The emerging literature gives grounds for neither a blanket rejection nor an unrestricted endorsement of substitution. Kasneci et al. (2023) discusses opportunities and challenges of large language models in education. Crompton and Burke (2023) review the broader field of AI in higher education. These sources establish a varied educational landscape rather than a single intervention with uniform effects. Different systems, subjects, and implementation arrangements require separate evaluation. Therefore, in this article its asks: To what extent can AI perform teachers' instructional roles, and what limits are its suitability for replacing teachers? Its contribution is a framework that connects demonstrated task capability with the conditions under which responsibility may reasonably be delegated. The proposed framework concerns

higher education and includes generative AI tutoring, while recognising that AI also encompasses systems beyond conversational language models.

The originality of this article lies in its conditional, risk-sensitive instructional allocation framework. Unlike

TPACK, which explains the integration of technological, pedagogical, and content knowledge, the proposed framework explicitly addresses whether an instructional responsibility should be delegated, under what evidence threshold, with what consequences if an error occurs, and to whom accountability remains assigned. Unlike broad human-centred AI guidance, it offers an activity-level decision structure linking pedagogical outcomes to risk and governance. This is a conceptual extension rather than an empirically validated model.

Aim and Conceptual Approach

The aim is to develop criteria for distinguishing task substitution, teacher augmentation, and the replacement of an educational role. Three objectives guide the analysis: identify instructional activities that may be delegated; explain why evidence of task performance is insufficient for full role equivalence; and propose conditions for evaluating teacher–AI arrangements.

This article uses a selective, theory-informed conceptual synthesis. Foundational theory, peer-reviewed studies, and official guidance were chosen for their relevance to teaching knowledge, AI capabilities, learning outcomes, and responsibility. The synthesis is not a systematic review. No exhaustive search, formal study screening, pooled effect estimate, or empirical test of the framework was conducted. Research findings are used to illustrate the problem, while the allocation criteria and propositions are the article's theoretical development. Policy documents contribute normative guidance; they cannot establish experimentally that a particular arrangement improves learning. The framework should therefore be interpreted as a basis for evaluation and subsequent research, rather than an empirically validated replacement model.

Defining Replacement and Instructional Capability

Task substitution occurs when a system performs a specific activity previously undertaken by a teacher, such as producing initial practice questions. Teacher augmentation occurs when AI contributes to an activity while a teacher designs, checks, adapts, or interprets it. Role replacement occurs when responsibility for a substantial educational function is transferred from the teacher to another arrangement. These categories are analytically distinct. An institution may automate first responses to common questions while retaining academic oversight. It may use AI to draft feedback while the lecturer approves the final response. Neither arrangement necessarily eliminates the teacher's role, even if it reduces time spent on a particular task.

Capability must also be defined in relation to outcomes. A system may generate a plausible explanation, identify a common error, or maintain a conversation. Evaluating it requires asking whether the explanation is correct, whether the learner understands it, and whether that understanding persists when assistance is withdrawn. Fluency, learner satisfaction, and independent achievement are different measures. Responsibility introduces another distinction. Producing an answer is an observable function. Being accountable for a grade, resolving an appeal, or deciding how to accommodate a learner involves authority and institutional processes. AI can inform such processes, but the presence of useful output does not itself determine who should hold responsibility. This article treats teacher leadership in consequential decisions as a governance recommendation, not a claim about every jurisdiction's legal requirement.

WHAT THE EVIDENCE DOES AND DOES NOT ESTABLISH?

AI as an Instructional Resource

AI can be used to generate examples, provide alternative explanations, and support repeated practice. Whether those functions are educationally valuable depends on accuracy and design. Chan and Hu (2023)

report students' perceptions of benefits and challenges of generative AI in a Hong Kong university sample. Perceived usefulness provides information about acceptance and experience; it does not demonstrate that teachers can be replaced. AI-assisted support may also broaden the time during which students can obtain a response. However, availability should not be confused with accessibility for every learner. Language, interface usability, cost, disability access, and digital confidence can shape whether an apparently available service is usable. These conditions are proposed evaluation criteria within the present framework.

Evidence From Designed Tutoring

Kestin et al. (2025) reports a randomized study of a research-designed AI tutor in undergraduate science education. Students using that tutor achieved stronger learning outcomes in less time than students in the compared active-learning condition. This is relevant evidence that carefully constructed AI instruction can be effective for course content. The inference must remain bounded. A designed tutor embedded in an existing course is not equivalent to an unconfigured chatbot, and an experimental comparison of selected learning activities does not test replacement of the entire teacher role. The study therefore supports evaluating specific instructional arrangements rather than declaring universal equivalence between teachers and AI.

Assisted Performance and Independent Learning

Bastani et al. (2025) conducted a field experiment in high-school mathematics that distinguished a general-purpose AI interface from a tutor with learning safeguards. Unrestricted AI assistance improved practice performance but was associated with poorer subsequent performance when access was removed. The guarded tutor largely mitigated that negative effect. The study concerns high-school learners, so it is used here as a mechanism-relevant example rather than direct evidence about university populations. Taken together, these studies make outcome selection central. Performance during assistance and learning demonstrated independently can diverge. The present framework therefore requires unassisted retention or transfer measures before treating successful task completion as evidence for instructional substitution. It also treats pedagogical configuration as a substantive part of the intervention.

Theoretical Foundation

Technological pedagogical content knowledge links technological knowledge with subject knowledge and pedagogical knowledge (Mishra & Koehler, 2006). It offers a useful basis for examining AI because selecting a tool cannot be separated from what is being taught and how learners are expected to develop understanding. Technical competence alone does not determine educational suitability.

Applied to the present question, the framework suggests that a teacher's expertise includes choosing explanations for a particular concept, interpreting student misconceptions, and deciding how technology fits the learning sequence. An AI system may contribute to each process, but its appropriate role depends on the content and pedagogy rather than its general conversational ability. The article extends this perspective by adding delegation criteria. It asks not only whether technology can support an activity, but also what evidence and oversight are required before an activity is assigned to it. This extension is the author's conceptual proposal; it should not be attributed to the original TPACK framework.

Miao and Cukurova (2024) provide a complementary competence perspective through UNESCO's framework for teachers. Its attention to human-centred thinking, AI ethics, applications, pedagogy, and professional learning situates AI integration within teachers' development. Miao and Holmes (2023) similarly foreground human agency and inclusion. Together, these sources motivate a governance-sensitive analysis without proving that any allocation rule is universally optimal.

Automation research distinguishes the technical ability to perform a function from the appropriate level of human involvement. Parasuraman et al. (2000) argue that automation can operate at different stages and levels, making delegation a design decision rather than a binary choice. Lee and See (2004) further show that appropriate reliance depends on calibrated trust: both over-reliance and under-use may undermine

performance. Applied to teaching, the central question is not simply whether AI produces a correct response, but whether students and educators can identify when to rely on it, verify it, or intervene.

Professional expertise also involves jurisdiction over complex, uncertain work (Abbott, 1988). Teachers interpret incomplete evidence of understanding, balance competing goals, and make situated decisions that cannot be reduced to answer generation. Human–AI collaboration should therefore be assessed as a configuration of complementary capabilities and responsibilities rather than an assumption that human review automatically guarantees safety. Algorithmic governance scholarship cautions that seemingly technical decisi- ons can obscure institutional context and distribute harms unevenly (Selbst et al., 2019). These perspectives extend the article beyond technological integration toward the governance of delegated professional work.

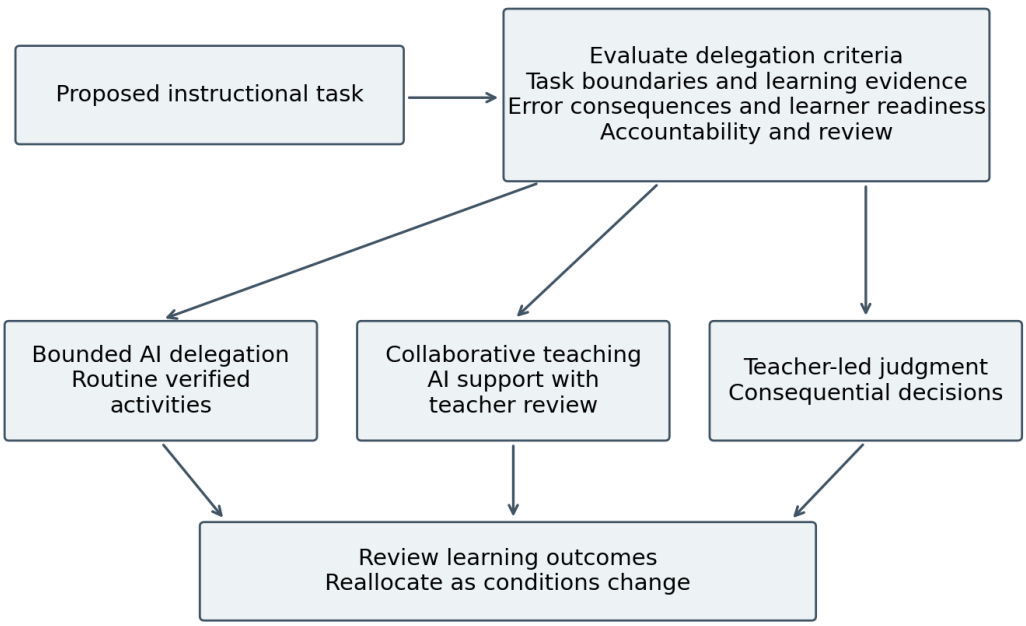
A Framework for Allocating Instructional Responsibilities

The framework begins with the activity rather than the occupation. Each proposed use is evaluated against five questions. Is the task sufficiently bounded for its output to be checked? What evidence shows learning rather than assisted performance? What happens if the output is wrong? Are learners prepared to use and question the system? Can an accountable educator review outcomes and respond to problems? These criteria lead to three provisional arrangements: bounded delegation, collaborative teaching, and teacher-led judgment. The arrangements are conditional rather than fixed rankings of technology and people. A low-consequence practice activity may qualify for delegation after testing, whereas a complex assessment judgment may require collaborative support with a teacher retaining responsibility.

Table 1 Proposed allocation of instructional responsibilities

Activity	Proposed arrangement	Criterion for delegation
Initial examples and practice	Bounded delegation	Content checked and objective explicit
Explanations and feedback	Collaborative teaching	Accuracy and independent learning reviewed
Complex assessment decisions	Teacher-led judgment	Valid criteria and accountable review
Course adaptation	Collaborative teaching	Learner needs interpreted in context
Institutional decisions	Teacher-led judgment	Authority and responsibility assigned

Figure 1 Proposed conceptual framework



Note. The figure represents a conceptual proposal requiring evaluation. Allocation categories are conditional

and may change after review.

The figure represents a proposed decision structure, not an empirical causal model. Institutions should revisit the allocation when the tool, learner population, learning outcome, or consequence of error changes. A system validated for introductory practice should not automatically be adopted for professional competence assessment.

Risk-sensitive allocation adds three explicit dimensions to the original five criteria: the severity and reversibility of potential error, the likelihood that error will be detected before harm, and the identity of the person or institution authorised to correct or appeal a decision. Low-risk, reversible practice tasks may be delegated after content testing and periodic sampling.

Medium-risk personalised feedback requires documented teacher review triggers, including uncertainty, repeated misconceptions, and student complaints. High-risk decisions involving progression, certification, accommodation, or disciplinary consequences should retain identifiable human decision authority, transparent records, and an accessible appeal route. These are proposed governance thresholds, not empirically established universal cut-offs.

A practical allocation record should document the task and intended outcome; the AI model and configuration; evidence from unassisted learning; error severity, reversibility, and detectability; learner accessibility; reviewer and escalation pathway; and the date for reassessment. A successful pilot should not automatically justify expansion into a more consequential activity.

Empirical Propositions and Normative Governance Principles

Empirical propositions (P1–P5): These claims can be investigated through comparative designs, independent learning assessments, error audits, and measures of learner and teacher behaviour.

P1. Tasks with explicit success criteria and readily checkable outputs are more suitable for bounded AI delegation than tasks requiring interpretation of competing educational goals.

P2. AI-supported task completion provides stronger grounds for delegation when accompanied by evidence of independent retention and transfer than when evaluated only through assisted performance.

P3. The consequences of error increase the need for professional review, even when average system performance is satisfactory.

P4. Teacher AI competence and instructional alignment shape the relationship between technical capability and learning outcomes.

P5. Learner readiness and access conditions moderate the effectiveness of AI delegation; the same system may require different levels of support across student groups.

Normative governance proposition (P6): This principle states a defensible institutional standard and requires ethical and policy justification, not only statistical testing.

P6. Accountability and review arrangements are necessary criteria for defensible delegation, alongside instructional effectiveness and efficiency.

P1–P5 are empirically testable propositions; P6 is a normative governance principle. Empirical studies can estimate learning outcomes, error rates, and oversight workload, whereas normative analysis must justify authority, fairness, transparency, and rights of review. Evidence of improved performance alone cannot settle whether an institution ought to transfer decision authority.

DISCUSSION

Why task automation changes rather than settling the teacher role

If AI provides routine explanations effectively, teachers may devote more attention to analysing students' reasoning, facilitating discussion, and designing learning activities. However, this redistribution of work is not automatic. Checking generated materials, resolving mistakes, and helping students use the system can create new responsibilities. Efficiency claims must include those activities.

A hypothetical economics course illustrates the issue. An AI tutor may provide practice identifying shifts in demand and supply. The lecturer sets the problem sequence, checks explanations, reviews common misconceptions, and evaluates students' independent interpretation of a new case. The AI has substituted for parts of practice support while the teacher continues to coordinate the educational process. This example is illustrative and is not a reported intervention. The question is therefore how work and responsibility are redistributed. Counting automated responses would miss whether students developed the intended understanding and whether educator workload decreased overall. The allocation framework makes those outcomes visible.

Assessment and Professional Judgment

Assessment requires evidence that corresponds to intended learning outcomes. TEQSA (2023) identifies assessment reform as a response to challenges intensified by generative AI. In the present framework, AI may support assessment preparation or preliminary analysis, but consequential judgments require scrutiny of validity, consistency, and the opportunity for review. A teacher's decision can also be flawed. Human leadership should therefore not be treated as a guarantee of quality. Moderation, explicit criteria, and appeal processes remain relevant for both human and AI-supported assessment. The purpose of oversight is to create a defensible process rather than assume that one actor is infallible.

Teacher judgment is interpretive rather than merely corrective. A lecturer may need to distinguish a misconception from a language barrier, anxiety, unfamiliar disciplinary conventions, or an inaccessible assessment format. Such decisions depend on contextual reasoning about learner history, course objectives, disciplinary standards, and available support. Pedagogical expertise includes selecting when to scaffold, when to challenge, and when to withhold immediate answers so that students practise independent reasoning. AI-generated recommendations can inform these choices, but apparently persuasive outputs do not remove the need to examine their assumptions.

Ethically, consequential educational judgments require attention to procedural fairness, privacy, unequal error burdens, explainability sufficient for review, and meaningful opportunities to contest decisions. Human oversight is not a ceremonial approval step: reviewers need time, competence, evidence, and authority to disagree with the system. Conversely, human judgment may itself be inconsistent or biased; accountable processes should therefore include rubrics, moderation, documented reasons, and audit of both AI-assisted and human decisions.

Governance and Implementation

Chan (2023) proposes an AI education policy framework with pedagogical, governance, and operational dimensions. This helps explain why replacing a task involves more than buying a tool. Institutions must define permitted uses, train staff, maintain access arrangements, and specify who responds when the system gives unsuitable guidance.

An initial implementation should use a bounded setting with clear outcomes and review points. Expansion should depend on demonstrated learning and workable support processes. Cost estimates should include

staff preparation, ongoing checking, technical support, and changes in assessment. These are proposed implementation principles, not cost-effectiveness findings from the present article.

Stress-testing the framework in difficult implementation cases

Consider first an AI tutor that gives fluent but subtly incorrect clinical or engineering advice. Even if average accuracy is high, the potential severity and low detectability of an error make unsupervised delegation unsuitable; domain-expert validation and escalation are needed. Second, an automated essay-feedback tool may appear useful overall but systematically misinterpret multilingual students' phrasing. Group-disaggregated error audits, accessible contestation, and lecturer moderation become prerequisites for scaling. Third, a university may propose AI-based grading of borderline progression decisions. Here reversibility, procedural rights, and institutional authority favour teacher-led judgment even when the tool offers consistent preliminary scores. A fourth case concerns a large introductory course where AI supplies round-the-clock practice. Bounded delegation may be reasonable when explanations are checked, students complete unassisted transfer tasks, and the institution monitors dependency and access barriers. These contrasting examples illustrate that the same technology may occupy different allocation categories depending on task stakes, population, error detectability, and oversight capacity. They are analytic stress tests, not observed case-study findings.

Research Agenda

Future research should compare instructional arrangements while keeping the learning objectives and assessment criteria consistent. A feasible study could compare teacher-led practice, AI-supported practice with teacher oversight, and bounded AI practice. Independent post-tests and delayed transfer tasks would distinguish short-term assistance from durable learning.

Researchers should record system configuration, content, learner characteristics, and teacher involvement. An intervention described simply as AI tutoring cannot be replicated or interpreted adequately without that information. Outcome measures should also include error incidence, student help-seeking, accessibility, and total educator time. Mixed methods could explain why arrangements work differently. Interviews may show whether students question AI output or copy it, while teaching observations may reveal additional monitoring work. Comparative institutional case studies could examine how accountability is assigned and how students challenge a mistaken decision. Sample size should be justified for the primary outcome, and any allocation to experimental conditions should preserve appropriate educational support. Changes in tools over time should be documented. A favourable result in one discipline or institution would warrant further testing rather than a universal replacement claim.

Future empirical work should preregister comparisons across risk categories and report false-positive and false-negative error patterns, reversibility, escalation frequency, and differential outcomes for learner subgroups. Governance research should separately examine whether affected students understand AI involvement, can obtain a reasoned explanation, and can appeal an adverse decision. Such mixed evidence would test both the predictive value and practical feasibility of the proposed framework.

LIMITATIONS

This synthesis is selective and the framework has not been validated. The criteria combine empirical questions with normative judgments about responsibility. Different institutions may reasonably assign responsibilities differently, and the article does not specify a universal legal or professional standard. The experimental examples differ in population and intervention. Their findings cannot be pooled informally or used to infer a single effect of AI. Moreover, professional experience, subject knowledge, and instructional quality vary among teachers. The proposed analysis must compare defined arrangements rather than an ideal teacher with an unspecified system.

The article concentrates on instructional allocation. The detailed relational mechanisms through which teachers may support emotional well-being are outside its main scope. They require their own analysis and

measures. This boundary allows the present article to make a focused contribution to the replacement debate.

CONCLUSION

AI may substitute for instructional activities when outputs are sufficiently reliable, learning is demonstrated independently, and responsibility is clearly assigned. That conclusion does not establish that the full teacher role has been replaced. Educational decisions also involve contextual interpretation, assessment validity, and accountable responses to problems. The article contributes a conditional allocation framework grounded in task characteristics, learning evidence, learner readiness, and oversight. Future research should test specific teacher–AI arrangements across subjects and learner groups. The practical question becomes what can be delegated under which conditions, with what educational result and who remains responsible.

REFERENCES

1. Abbott, A. (1988). The system of professions: An essay on the division of expert labor. University of Chicago Press.
2. Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö., & Mariman, R. (2025). Generative AI without guardrails can harm learning: Evidence from high school mathematics. Proceedings of the National Academy of Sciences, 122(26), Article e2422633122. https://doi.org/10.1073/pnas.2422633122
3. Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. International Journal of Educational Technology in Higher Education, 20, Article 38. https://doi.org/10.1186/s41239-023-00408-3
4. Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. International Journal of Educational Technology in Higher Education, 20, Article 43. https://doi.org/10.1186/s41239-023-00411-8
5. Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. International Journal of Educational Technology in Higher Education, 20, Article 22. https://doi.org/10.1186/s41239-023-00392-8
6. Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences, 103, Article 102274. https://doi.org/10.1016/j.lindif.2023.102274
7. Kestin, G., Miller, K., Klales, A., Milbourne, T., & Ponti, G. (2025). AI tutoring outperforms in-class active learning: An RCT introducing a novel research-based design in an authentic educational setting. Scientific Reports, 15, Article 17458. https://doi.org/10.1038/s41598-025-97652-6
8. Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. Human Factors, 46(1), 50–80. https://doi.org/10.1518/hfes.46.1.50_30392
9. Miao, F., & Cukurova, M. (2024). AI competency framework for teachers. UNESCO. https://www.unesco.org/en/articles/ai-competency-framework-teachers
10. Miao, F., & Holmes, W. (2023). Guidance for generative AI in education and research. UNESCO. https://doi.org/10.54675/EWZM9535
11. Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. Teachers College Record, 108(6), 1017–1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x
12. Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans, 30(3), 286–297. https://doi.org/10.1109/3468.844354
13. Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S., & Vertesi, J. (2019). Fairness and abstraction in sociotechnical systems. Proceedings of the Conference on Fairness, Accountability, and Transparency, 59–68. https://doi.org/10.1145/3287560.3287598

14. Tertiary Education Quality and Standards Agency. (2023). Assessment reform for the age of artificial intelligence. https://www.teqsa.gov.au/guides-resources/resources/corporate-publications/assessment-reform-age-artificial-intelligence