

TALKING IT OUT: USING INTERACTIVE ORAL ASSESSMENT TO AUTHENTICATE LEARNING IN AN AGE OF GEN AI

Colum A. Foley, & James Byrne
Dublin City University (Ireland)

Abstract

Generative Artificial Intelligence (GenAI) has intensified concerns regarding academic integrity and the validity of traditional higher education assessment. This paper presents two case studies examining interactive oral assessment approaches designed to strengthen assessment authenticity in AI-rich environments. The first explores a face-to-face group presentation incorporating spontaneous questioning and collaborative reasoning among undergraduate students. The second examines a fully online Individual Interactive Oral Assessment (IOA) involving adaptive one-to-one dialogue with postgraduate learners. Drawing on theories of authentic assessment, formative assessment, service quality, and learning-oriented assessment, the paper argues that dialogic assessment formats provide a sustainable alternative to surveillance-based responses to GenAI. Findings indicate that assessment designs centred on interaction, contextualised reasoning, and live explanation reduce opportunities for AI-assisted substitution while enhancing student engagement and confidence. The paper concludes that authentic oral assessment can preserve validity, integrity, and educational value in contemporary higher education.

Keywords: *Interactive oral assessment, Generative Artificial Intelligence, authentic assessment, academic integrity, higher education.*

1. Introduction

Generative Artificial Intelligence (GenAI) tools are increasingly embedded within students' academic practices, enabling rapid generation of text, arguments, and summaries. Although these technologies may support learning, they challenge assumptions surrounding authorship, originality, and the reliability of traditional written assessment. Unsupervised coursework completed asynchronously now provides limited evidence of students' independent understanding. Institutional responses have largely focused on AI-detection tools and restrictive policies. However, detection systems remain unreliable and ethically contested, while prohibition-based approaches risk limiting innovation and damaging trust. GenAI has therefore exposed a broader issue within higher education: the overreliance on product-focused assessment that privileges outputs over learning processes. This paper argues that assessment integrity should be addressed through authentic assessment design rather than surveillance. Two case studies are presented to demonstrate how dialogic and interactive assessments can preserve validity by requiring students to explain, justify, and apply understanding in real time.

1.1. Authentic assessment

Authentic assessment involves tasks that reflect real-world professional practice and require students to apply knowledge meaningfully within contextualised environments (Gulikers et al., 2004). Gulikers et al.'s five-dimensional framework (task, physical context, social context, assessment form, and criteria) provides a useful structure for evaluating authenticity in assessment design. Assessments that score highly across these dimensions are more likely to promote transferable learning and resist generic AI-generated reproduction. Within AI-rich environments, authenticity functions as a mechanism of assessment resilience. Tasks involving live communication, professional judgement, and contextual explanation reduce opportunities for AI substitution while foregrounding students' own reasoning processes.

1.2. Formative assessment and dialogue

Nicol and Macfarlane-Dick (2006) emphasise the importance of formative assessment in supporting self-regulated learning. Central principles include feedback dialogue, opportunities for

reflection, and clarity regarding assessment expectations. Interactive oral assessments align closely with these principles by allowing assessors to adapt questioning, probe understanding, and provide immediate feedback. Dialogic assessment formats also blur the distinction between formative and summative assessment, enabling assessment encounters to simultaneously support learning and evaluate performance.

1.3. Learning-oriented assessment and service quality

Race (2001) argues that effective assessment should actively contribute to learning through engagement, articulation, and feedback. Oral assessments operationalise these principles by requiring students to explain, justify, and refine understanding in real time. Additionally, Parasuraman et al.'s (1985) service quality model offers a complementary perspective by framing assessment as a relational experience. Dimensions such as reliability, responsiveness, and assurance are particularly relevant within interactive assessments, where fairness and student confidence are shaped through interpersonal engagement.

2. Methodology

This paper adopts a qualitative case study approach to examine two assessment designs implemented across contrasting higher education contexts. The cases were selected to demonstrate complementary approaches to dialogic assessment: one group-based and face-to-face, and the other individual and fully online. Data sources included assessment design documentation, reflective accounts from teaching staff, and anonymised student feedback collected through module evaluations. The analysis was intended to generate transferable insights into assessment resilience within AI-rich environments rather than statistically generalisable findings.

3. Case study one: Face-to-face group presentation

3.1. Context and design

The first case study involved final-year undergraduate students completing a face-to-face group presentation within a capstone module. Students collaboratively investigated a complex disciplinary problem before delivering a structured presentation followed by an interactive question-and-answer session. Assessors engaged students in spontaneous questioning that required clarification, elaboration, and collaborative reasoning. Although students were permitted to use GenAI tools during preparation, no scripts or pre-generated responses could be relied upon during the live assessment.

3.2. Authenticity and integrity

The assessment aligned strongly with Gulikers et al.'s (2004) framework for authentic assessment. The task mirrored professional communication practices, emphasised collaboration, and foregrounded real-time reasoning. Assessment criteria prioritised conceptual understanding, coherence, and responsiveness rather than rehearsed content reproduction. The interactive questioning phase proved particularly significant for academic integrity. Students were required to explain choices, respond to counterarguments, and negotiate meaning collaboratively, substantially reducing the feasibility of AI-generated substitution.

3.3. Educational value

Consistent with Race's (2001) learning-oriented perspective, students reported that preparing for and participating in the presentation deepened their understanding and increased confidence. The assessment was perceived as rigorous but fair, with integrity emerging from demonstrated competence rather than surveillance or compliance measures.

4. Case study two: Individual Interactive Oral Assessment (IOA)

4.1. Context and design

The second case study explored a fully online Individual Interactive Oral Assessment (IOA) conducted with postgraduate learners. Each student participated in a scheduled one-to-one oral assessment lasting approximately 25 minutes. The structure combined pre-published thematic prompts with adaptive questioning tailored to individual responses. Students were encouraged to draw on professional or disciplinary contexts, ensuring that responses remained personalised and situated.

4.2. Formative and integrity functions

The IOA operationalised Nicol and Macfarlane-Dick's (2006) formative assessment principles by enabling immediate feedback, clarification, and dialogue during the assessment encounter. Assessors could

probe understanding, request examples, and address misconceptions in real time. From an integrity perspective, personalised questioning made generic AI-generated responses ineffective. Students' ability to articulate contextualised reasoning became the primary evidence of learning, reinforcing both validity and reliability.

4.3. Student experience and service quality

Applying Parasuraman et al.'s (1985) framework, the IOA demonstrated strong alignment with reliability, responsiveness, and assurance. Standardised assessment structures supported consistency, while adaptive dialogue created a supportive and transparent environment. Student feedback indicated high levels of trust in the process and recognition of the assessment's authenticity.

5. Discussion

The case studies demonstrate that assessment resilience in the GenAI era is strengthened through interaction, contextual understanding, and live reasoning. Rather than treating AI solely as a threat, these approaches recognise its role within preparation while reserving summative judgement for independently demonstrated understanding. The findings also challenge deficit-oriented approaches to academic integrity by reframing integrity as the presence of credible evidence of learning rather than the absence of misconduct. Although concerns regarding workload and scalability remain, strategically implemented dialogic assessments may reduce time spent addressing academic misconduct cases while improving assessment authenticity and student trust.

6. Implications for Practice

Several implications emerge for educators and institutions seeking to redesign assessment within AI-rich environments:

- Prioritise assessment formats requiring interaction, explanation, and contextual judgement.
- Explicitly integrate ethical GenAI use into assessment guidance.
- Evaluate assessment quality in terms of learner experience alongside psychometric robustness.
- Invest in staff development focused on dialogic assessment literacy.

These principles support assessment approaches that remain educationally meaningful while responding constructively to technological change.

7. Conclusion

Generative Artificial Intelligence has intensified existing challenges surrounding assessment authenticity and academic integrity within higher education. However, it has also created opportunities to reconsider how learning is evidenced and evaluated. The case studies presented demonstrate that authentic, dialogic assessment designs can strengthen integrity, validity, and educational value without relying on surveillance-based approaches.

Interactive oral assessments provide a constructive and sustainable response to the pedagogical disruptions introduced by GenAI by foregrounding explanation, reasoning, and contextual understanding. Rather than representing a defensive reaction to AI, these approaches reorient assessment toward practices that better reflect how learning occurs and should be recognised in contemporary higher education.

References

- Gulikers, J. T. M., Bastiaens, T. J., & Kirschner, P. A. (2004). A five-dimensional framework for authentic assessment. *Educational Technology Research and Development*, 52(3), 67–86. <https://doi.org/10.1007/BF02504676>
- Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199–218. <https://doi.org/10.1080/03075070600572090>
- Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1985). A conceptual model of service quality and its implications for future research. *Journal of Marketing*, 49(4), 41–50. <https://doi.org/10.1177/002224298504900403>
- Race, P. (2001). *The lecturer's toolkit: A practical guide to learning, teaching and assessment* (2nd ed.). London: Routledge Falmer.