

AI and assessment and feedback in foreign language education

Professor Mary Richardson

<https://profiles.ucl.ac.uk/51771-mary-richardson>

Dr Catarina Correia

<https://profiles.ucl.ac.uk/71735-catarina-f-correia>

The role of assessment in learning

Assessment regulates learning: by capturing **what was achieved** (summative) and by informing **what next in learning** (formative).

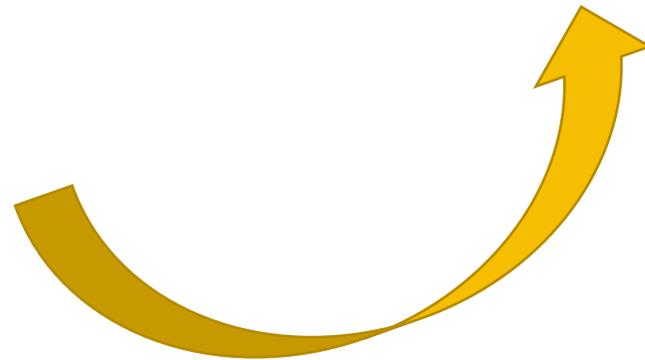
Feedback bridges the learning gap (knowledge and skills) between where the learner is (know/ can do) and where they need to go (need to know/be able to do)

Good feedback involves both **judgement** and **information about next steps**, is learner-centred, and a dialogue between the giver and the receiver of feedback.

The impact of feedback depends on quality of the message, as well as how the receiver processes and responds to it.

Teacher/student and peer/student relationship.

Moving from the tsunami view to surfing the wave



The role of AI in assessment

AI is used as a **tool to assess** (e.g. reality simulator in medical examinations; or an adaptive test) and can become **an assessor** (e.g. automated marking and feedback chatbot).

AI has conflated what was the traditional separation between the assessor who designs the assessment and makes the judgements (human) and the instrument that is used to assess (object).

AI becomes a partner in the assessment ecology, enhancing our ability to:

- 1) use assessment to improve learning - e.g. chatbots for personalised feedback, and data management to support the pedagogical decision-making.
- 2) use assessment to report on learning - e.g. data management to report on summative outcomes.

AI and the student: who gets to be in the driving seat?

Self-regulated learners who use genAI assessment information in intelligent ways to enhance their learning

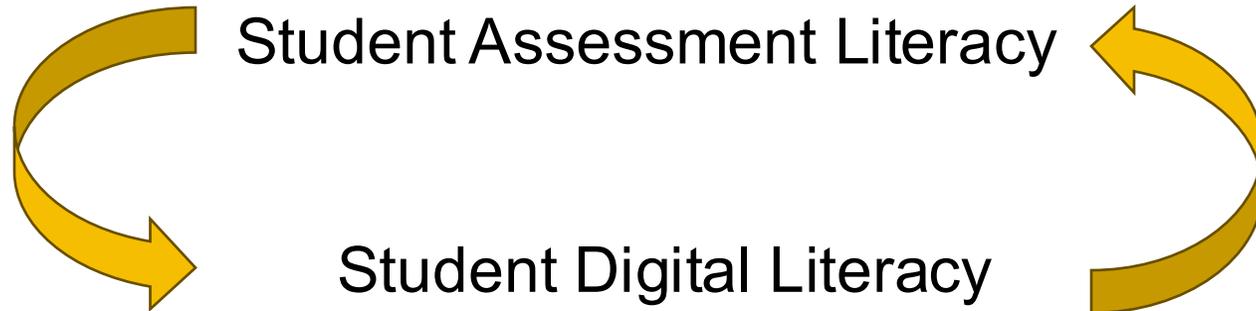
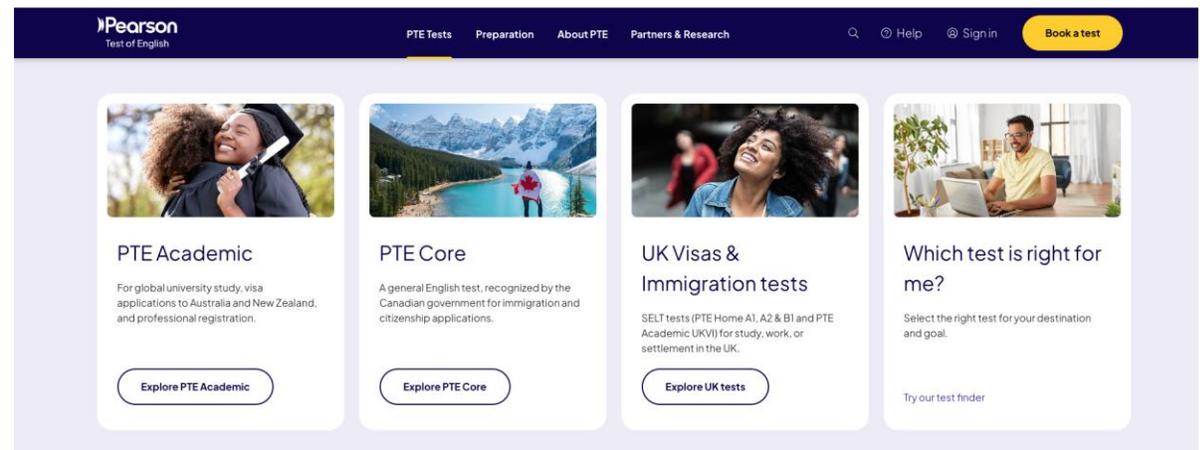


Image generated by MSCopilot (smart GPT-5)

A case: PTE Academic test

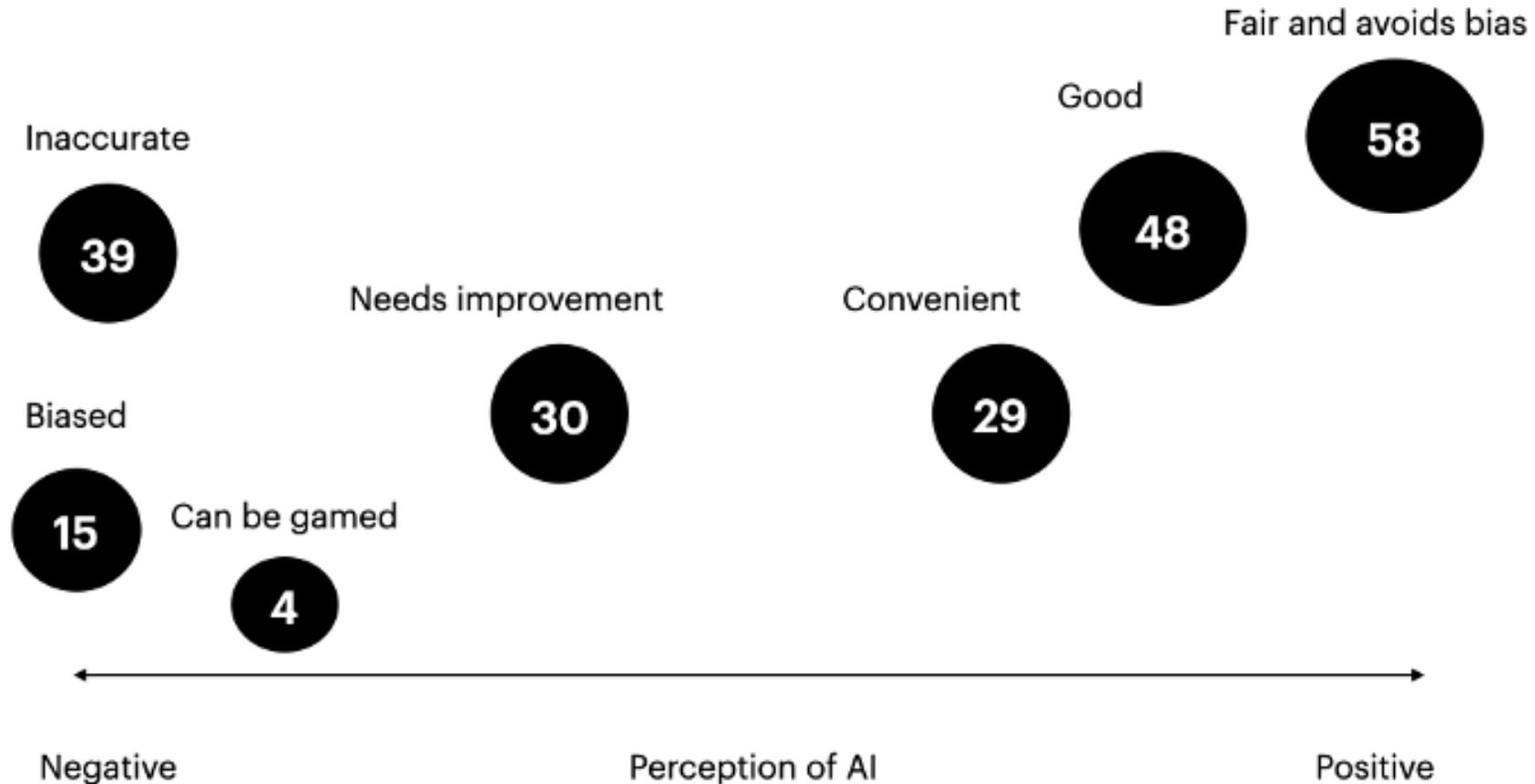
1. What is known about the role [their perception of planning for and taking the test] of the candidate in the AI-led language testing setting of PTE Academic?
2. How can we characterise the experience of candidates in AI-led language test domains?
3. To what extent can we use candidates' feedback to inform testing with AI components?

<https://www.pearsonpte.com/>



The screenshot shows the Pearson Test of English website homepage. The header includes the Pearson logo, navigation links for PTE Tests, Preparation, About PTE, and Partners & Research, along with search, help, and sign-in options, and a 'Book a test' button. The main content area features four cards: 'PTE Academic' (for global university study, visa applications to Australia and New Zealand, and professional registration), 'PTE Core' (a general English test recognized by the Canadian government for immigration and citizenship applications), 'UK Visas & Immigration tests' (SELT tests for study, work, or settlement in the UK), and 'Which test is right for me?' (a test finder tool). Each card has an 'Explore' button.

General views about AI use in tests



Reflections on the test-taking

I felt like I was going to go to prison. They body checked me, checked my documents, I couldn't have my jacket on. I'm not going to see the light of day again! IV15

'... in the background a girl answering not the same question, but also something about relatives. Her answer was like 'sister' or 'niece'. I could hear everyone around me, it was distracting and I lost focus a little bit.' IV19

IV4 was very specific claiming that PTE requires 'concentrating on the screen '...that requires a lot of 'working memory and a 'different level of stress'.

Reflections on understanding of AI

'Did the system listen to me properly? Has my voice reached them or not?' IV4

'I don't believe my assessment results were correct although I passed the exam comfortably. I was seeing low scores in sections where I am the most confident and sure of my capabilities and performance in the exam.' IV12

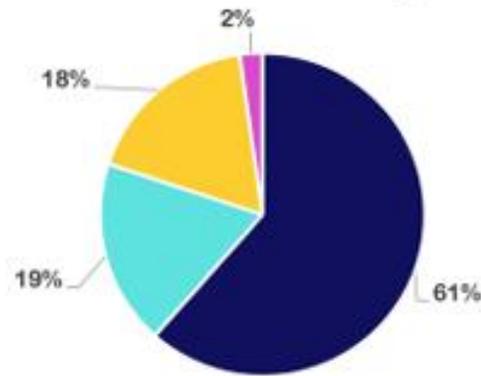
IV17 said that non-human functions for marking would mean the test was less biased and fairer; working on screen means *'...avoiding human examiners [is a good thing], especially for those who are stressed, introverted and nervous'*

Recent research from PTE

From Barrow (2025)

Chart 5. Perceptions of objectivity of AI scoring compared to human scoring

I believe AI scoring is more objective than human scoring



• Agree • Strongly agree • Disagree • Strongly disagree

...the study reveals a general trust in AI scoring systems, with test takers appreciating the objectivity and fairness they offer compared to only human scoring. However, there is a consensus on the need for a balanced approach, combining AI and human oversight to ensure accuracy and reliability, particularly for speaking tasks. (p23).

Useful studies – Open Access

Barrow, E. (2025). *Test Taker Perspectives: Pearson Test of English*.
<https://www.pearsonpte.com/research>

Richardson, M., & Clesham, R. (2021). Rise of the machines? The evolving role of AI technologies in high-stakes assessment [Article]. *London Review of Education*, 19(1).
<https://doi.org/10.14324/LRE.19.1.09>

Richardson, M., Gray, S. L., & Popov, J. (2023). *Computer-based tests and machine marking: Candidates' perceptions and beliefs about the test taking experience*.
<https://www.pearsonpte.com/research>

Zhang, Y., & Liu, Y. (2025). Designing ChatGPT-mediated feedback activities in EFL writing: a design-based study of the dialogic feedback triangle. *Assessment and Evaluation in Higher Education*, 0(0), 1–21. <https://doi.org/10.1080/02602938.2025.2571846>