AI for assessment

AI for assessment

Generative AI can support aspects of assessment design by helping learning designers draft assessment ideas, generate examples, propose question formats, and refine instructions or criteria language. It can be a useful drafting partner, but it should not be relied upon uncritically for validity, fairness, or alignment.

Assessment design has direct consequences for workload, evidence quality, learner experience, and academic integrity. For that reason, AI should be used as a support tool for exploring or refining assessment ideas, not as an authority on whether an assessment is educationally sound.

Where AI can help

AI can assist with:

- drafting assessment task ideas
- generating question stems
- proposing rubric language
- refining assessment instructions
- generating examples of evidence or performance
- suggesting formative assessment activities
- checking whether an assessment appears aligned to stated outcomes

These uses are most valuable when AI is helping to surface options and language, rather than determining assessment quality by itself.

Good practice

When using AI for assessment:

1. **Begin with the learning outcomes**
   Ensure the assessment is tied to what learners are actually expected to know or do.
2. **Use AI to generate options**
   Ask for alternative task formats, clearer wording, or possible criteria rather than assuming the first output is suitable.
3. **Review validity and authenticity**
   Check whether the task genuinely measures the intended outcome and reflects meaningful evidence of learning.
4. **Refine the language**
   AI can help make instructions, questions, and rubric statements clearer, but clarity should not come at the expense of precision.
5. **Check fairness and level**
   Make sure the assessment is appropriate for the learner group, the course level, and any programme or industry expectations.

Risks and limitations

AI can produce assessment content that appears polished but is weak in important ways. Common problems include:

- shallow alignment to learning outcomes
- poor or inconsistent cognitive level
- unrealistic or generic task contexts
- criteria that sound good but are difficult to apply reliably
- invented standards or references
- assessment ideas that are too broad, too narrow, or too easy to game

These risks make human review essential, especially for summative assessment.

Example uses

Example 1: Improving assessment instructions

A designer has a strong task idea but the instructions are too wordy or unclear. AI can suggest a cleaner structure with:

- purpose
- task instructions
- expected output
- submission requirements
- success criteria summary

The designer then checks that the final wording remains accurate and institutionally appropriate.

Example 2: Drafting rubric language

AI can help convert rough notes into draft criteria language, such as turning “shows understanding of audience and purpose” into more explicit descriptors. The designer still needs to test whether the rubric is usable and aligned.

Relationship to other design work

Assessment design should remain connected to writing learning outcomes and objectives, course level alignment, summative assessment planning, and tasks and activities.

AI-generated illustration of assessment and rubric design with criteria, evidence planning, and struc

Example: an AI-generated assessment design image showing rubric drafting and thinking about authentic evidence of learning.

Practical guidance

AI can be helpful in assessment design when it is used to:

- generate alternatives
- improve wording
- draft structures
- stress-test alignment
- produce first-pass rubric or question language

It should not be treated as the final authority on assessment quality, validity, or compliance.

Use AI in assessment design as a drafting and review aid, while keeping final responsibility for alignment, fairness, and quality with the learning designer.

Revision #5
Created 2024-08-14 23:13:39 UTC by Chester Whitwell
Updated 2026-03-17 01:02:38 UTC by Tasker