

Standards for Evaluating Authentic Assessment Innovations

The Alternative Assessment and Formative Feedback Toolkit (AAFFT) project was developed as part of a two-year CLI Fellowship at Erasmus University Rotterdam. Its purpose is to help lecturers, program coordinators, and educational developers design, implement, and evaluate assessment practices that go beyond traditional exams and written tests.

Authentic and alternative assessment approaches are increasingly recognized as powerful drivers of deep learning, student engagement, and employability skills. They enable students to demonstrate not only what they know, but also how they can apply knowledge, solve complex problems, collaborate effectively, and reflect critically. These skills are essential in academic, professional, and civic contexts. At the same time, formative feedback practices provide the scaffolding that allows students to learn from the process, improve iteratively, and take ownership of their learning journey.

The toolkit is designed to be practical, evidence-informed, and adaptable. It brings together:

- Checklists and templates to support the design of assessments and feedback activities.
- Guidelines and exemplars showcasing good practices across disciplines.
- **Evaluation standards and tools that help measure the quality and impact of innovative assessment strategies.**
- Protocols for reflection and continuous improvement so that both staff and students can actively learn from their experiences.

In this context, the standards would allow lecturers, program coordinators, or evaluators to systematically check if authentic assessment practices meet quality expectations.

STANDARD SCALE

1	2	3	4	N/A
The standard is not adequately addressed or implemented in the initiative	The standard is partially addressed or implemented, & improvements are needed.	The standard is well addressed & implemented, meeting expectations. Improvements could enhance the initiative.	The standard is exceptionally addressed and implemented, exceeding expectations. No improvements are needed.	Not applicable

According to the above-mentioned information, a rating of 1 & 2 is considered insufficient evaluation, while 3 & 4 are considered sufficient evaluation. The criteria for identifying the standards below included **measurability** (the standards must allow the collection of meaningful data so that progress and outcomes can be evaluated consistently), **applicability** (the standards are designed to be flexible and relevant across different disciplines, programs, and teaching contexts), **evidence-based** (each standard reflects insights from credible academic literature and established best practices), **ethical** (the standards take into account the well-being, rights, and safety of all stakeholders), and **continuous improvement-oriented** (rather than serving as a static checklist, the standards are designed to encourage reflection, feedback, and iterative enhancement of teaching practice).

Standards to evaluate the impact of the education initiative

Alignment with Intended Learning Outcomes

The assessment tasks are directly linked to course/program outcomes

Rating scale

1

2

3

4

N/A

Justification:

Authenticity & Real-World Relevance

Tasks mirror professional or societal challenges relevant to the discipline

1

2

3

4

N/A

Justification:

Transparency & Clarity

Criteria, rubrics, and expectations are clearly communicated to students

1

2

3

4

N/A

Justification:

Evidence-Informed Design

The assessment approach is informed by research and/or established good practices

1

2

3

4

N/A

Justification:

Inclusivity & Accessibility

The assessment accommodates diverse learners and ensures equitable participation

1

2

3

4

N/A

Justification:

Standards to evaluate the impact of the education initiative

Rating scale

Feedback Integration

The design ensures opportunities for formative feedback, reflection, and improvement

1

2

3

4

N/A

Justification:

Skill Development

The assessment fosters 21st-century skills (critical thinking, creativity, collaboration, communication, digital literacy)

1

2

3

4

N/A

Justification:

Progress Monitoring

Student progress is observable and measurable throughout the process

1

2

3

4

N/A

Justification:

Transfer of Learning

The assessment enables students to apply knowledge and skills across contexts

1

2

3

4

N/A

Justification:

Sustainability & Long-Term Impact

The design supports skills and knowledge that remain relevant beyond the course

1

2

3

4

N/A

Justification:

Standards to evaluate the impact of the education initiative

Rating scale

Use of Technology & Innovation

Technology is integrated in meaningful, value-adding ways (e.g., e-portfolios, AI scaffolds, simulations)

1

2

3

4

N/A

Justification:

Student Empowerment

Students are active participants (choice, self-assessment, peer-feedback) in the assessment process

1

2

3

4

N/A

Justification:

Ethics & Integrity

The design addresses fairness, transparency, and academic integrity, including considerations of AI

1

2

3

4

N/A

Justification:

Faculty Readiness & Support

Instructors receive adequate training, resources, and support to implement the innovation

1

2

3

4

N/A

Justification:

Stakeholder Engagement

Where relevant, industry, community, or external partners are engaged in assessment design or feedback

1

2

3

4

N/A

Justification: