STIAEM

Slow Tech Innovation
AI, Ethics and
New Teaching Methods

Moving Beyond Rote Text Reproduction: Methodologies for Real-World Tasks

Co-funded by the European Union

2025-1-IT02-KA220-SCH-000365213

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Education and Culture Executive Agency (EACEA). Neither the European Union nor EACEA can be held responsible for them... Project reference code: 2025-1-IT02-KA220-SCH-000365213

Università di Foggia
Studi Umanistici, Lettere, Beni Culturali, Scienze della Formazione

İSTANBUL - ESENLER
İLÇE MİLLİ EĞİTİM MÜDÜRLÜĞÜ

1474

STIAEM

Slow Tech Innovation
AI, Ethics and
New Teaching Methods

Co-funded by
the European Union

2025-1-IT02-KA220-SCH-000365213

eduOpen

CC BY-NC-SA 4.0 International

This license requires that reusers give credit to the creator. It allows reusers to distribute, remix, adapt, and build upon the material in any medium or format, for noncommercial purposes only. If others modify or adapt the material, they must license the modified material under identical terms.

Università di Foggia
Studi Umanistici, Lettere, Beni Culturali,
Scienze della Formazione

OER 6.3.2: Moving Beyond Rote Text Reproduction: Methodologies for Real-World Tasks
Reference Module: Module 6 - Classroom Practice and Authentic Assessment
Related Video-Pill: Video 6.3 - Designing Authentic Assessment Tests
Authoring partner: University of Foggia - Italy
Topic / Thematic Focus: Beyond rote reproduction: methodologies for real-world tasks

Introduction and Learning Objectives

This resource completes, on the operational level, the pathway begun with OER 6.3.1 and accompanies video-pill 6.3 'Designing Authentic Assessment Tests'. If the previous resource defined what makes a test authentic, this one illustrates its redesign methodologies, showing how to move beyond tasks of pure textual reproduction and how to build complex tasks rooted in the real world.

The text connects to the European DigCompEdu framework (Redecker, 2017), with particular reference to Area 4 - Assessment (4.1 Assessment strategies, 4.3 *Feedback* and planning) and to Area 5 - Empowering Learners, and is consistent with the principles of Universal Design for Learning (CAST, 2024) in the diversification of the forms of expression and assessment.

By the end of the reading, the teacher will be able to:

- recognise why tasks of reproduction have lost their capacity to measure in the era of generative models;
- apply the redesign blueprint of Villarroel and colleagues to transform a traditional test into an authentic task;
- govern the real constraints (time, subjectivity of marking, large classes) with sustainable organisational strategies.

1. The collapse of tasks of reproduction in the era of generative models

The starting point of this resource is an observation that is by now hard to dispute. Large language models carry out with extraordinary efficiency the operations that traditional schooling has long taken as the privileged object of assessment: summarising a text, synthesising concepts, expounding notions, answering multiple-choice questions. In a few seconds, and with impeccable fluency. A radical consequence follows: any test based on the pure reproduction of information has ceased to measure what it intended to measure.

The problem is not only technical, but pedagogical. If we reward reproduction, we are in fact teaching how to delegate to the machine: the student learns that the task consists in providing a correct output and that such an output is today obtainable without mobilising one's own cognitive resources. The task of reproduction thus becomes an incentive to uncritical delegation, exactly the opposite of the formation of thought that the school should pursue.

Consider, for concreteness, the classic task of summarising a chapter. Until a few years ago it measured, albeit imperfectly, the student's ability to select the relevant information, to rank it and to reformulate it in their own words. Today the same task no longer discriminates: those who have understood and those who have not read the text can submit a correct, fluent and indistinguishable piece of work. The test, in other words, has ceased to observe the cognitive process that it intended to assess.

Reversing the perspective does not mean fleeing technology, but bringing assessment back to its proper object: not the product, but the cognitive process that generates it. This is the direction along which the methodology of redesigning tests presented in the following paragraphs moves.

2. The blueprint of Villarroel: contextualising the stimulus and making it traceable

The reference model for the redesign is the one proposed by Villarroel and colleagues (2018; 2019), who translate the principles of authentic assessment into an operational blueprint for the teacher. The first intervention consists in contextualising the stimulus, making it situated. In place of the abstract request - describe the causes of a historical event - a concrete performance is proposed: prepare a documentary panel for an exhibition aimed at your peers, choosing and ordering the sources I provide. Knowledge remains necessary, but is transformed into the instrument of a tangible performance.

The second intervention consists in making the task traceable and extended over time. The test is broken down into stages - research, draft, critical revision, final version - and the student is asked to document the pathway with a logbook, noting the logical steps, the discarded roads, the corrections made. In this way the teacher observes thought in action, and not only the final result; and the student develops a self-regulatory competence that no traditional test elicits.

An example can clarify the scope of this intervention. If the task consists in writing a newspaper article on a civic topic, traceability requires the student to document the choice of the angle of treatment, the sources consulted and discarded, the revisions the text underwent. The teacher does not assess only the finished article, but the quality of the reasoning that produced it: the relevance of the sources, the coherence of the choices, the ability to revise one's own work critically.

The third intervention concerns the advance sharing of the criteria and the function of *feedback*, which constitutes the theme of the next paragraph.

3. *Feedback* as feed-forward: transparent rubrics and assessment of the process

Villarroel and colleagues attribute a decisive role to the transparency of the criteria and to the prospective function of *feedback*. Rubrics are not handed out after the test, as an instrument for justifying the grade, but before, as a map of the expected work: the student must know in advance which dimensions will be assessed, with which level descriptors and with which quality standard. In this sense *feedback* is not a retrospective judgement on the product, but a prospective indication that orients the next step: it is feed-forward, the looking ahead that turns assessment into a guide for learning.

On this level, generative models can play a supporting role, but not one of substitution: they can help the teacher draft descriptive *feedback*, generate examples at different levels of quality, produce variants of a task. The validation of the judgement - the decision about what is adequate, relevant, of quality - remains, however, inalienably human: no algorithm can assume the pedagogical responsibility of assessment. The authentic assessment of the process requires, consequently, specific tools: analytic rubrics with observable descriptors, metacognitive logs, process portfolios.

4. Multidisciplinary tasks and authentic products: what AI cannot carry out

The decisive character of tasks for the real world is their multidisciplinary openness and their irreducibility to a single answer. An authentic task asks the student to establish original connections between different fields, to mobilise creativity and intuition, to make argued decisions in ambiguous situations. And it is precisely in these territories that generative artificial intelligence shows its constitutive limits.

The authenticity of a task depends, according to Gulikers, Bastiaens and Kirschner (2004), on the coherence between five dimensions: task, context, expected outcome and criteria.

A language model produces plausible answers on the basis of statistical regularities learned from enormous textual corpora; it does not possess a physical or social anchoring in our class, it did not conduct the investigation, it did not collect the data, it did not

experience the interpretive conflict that emerged in the classroom. Rooting the task in these situated elements makes it structurally resistant to automated substitution: the model's output, however fluent, cannot satisfy a performance that requires what it does not have.
A naive reading must nevertheless be avoided. It is not a matter of making the tests impossible for the machine in a sort of technical contest, but of reorienting them towards what has pedagogical value: the ability to think in and about the context, to argue, to build and defend a product. Resistance to automation is a consequence of this value choice, not its aim.
Concrete examples help translate the principle into practice. An authentic multidisciplinary task could ask students to design and present a plan for the enhancement of a local cultural asset, integrating history, geography, civic education and economics: students must collect data in the field, interview witnesses, argue for design choices and defend them before an audience. No generative output, however fluent, can replace the situated experience that that task requires.

5. Governing the real constraints: time, subjectivity, large classes

The design of authentic tests clashes, in everyday practice, with unavoidable constraints, which must be named honestly. Time, first of all: designing and marking authentic tasks requires a greater investment of work than that required by multiple-choice items. Subjectivity, secondly: an open task introduces variability between markers, and with it the risk of inequity. Logistics, finally: large classes make the management of complex individual products burdensome.
These constraints are met with tried-and-tested strategies, not with surrender. First of all, with a balanced assessment plan: not every test must be a great authentic task; quick low-stakes tests alternate with a limited number of focal authentic tasks distributed over the year. Secondly, with sharing: test models and rubrics are designed and exchanged among colleagues. Subjectivity is then reduced with detailed analytic rubrics, descriptors of observable behaviours, model examples at the different levels and sessions of cross and blind marking on a sample of pieces of work. In large classes, finally, the load is lightened with short presentations in rotation and with peer assessment, which is in itself a learning experience.
A final consideration concerns peer assessment and cross marking: far from being merely an expedient to reduce the teacher's load, they constitute occasions of authentic learning, since they ask students to apply the criteria of quality to their classmates' work and, by reflection, to their own. Thus authentic assessment generates a virtuous circle between assessing, learning and self-regulating.
As Wiggins (1990) observes, the quality of learning also depends on the quality of the teachers' work: the apparently high costs of authentic assessment are offset by the gains in terms of professional development and student learning. Direction, more than perfection, is what counts.

Real constraint	Sustainable operational strategy
Time (burdensome design and marking)	Balanced assessment plan: quick low-stakes tests + few focal authentic tasks over the year; sharing of models among colleagues
Subjectivity (variability between markers)	Analytic rubrics, observable descriptors, model examples at different levels, cross and blind marking on a sample
Large classes (logistical load)	Short presentations in rotation, peer assessment, sampling of pieces of work

Visual Scaffolding

MOVING BEYOND ROTE REPRODUCTION:
A BLUEPRINT FOR AUTHENTIC ASSESSMENT IN THE AI ERA

THE CRISIS OF ROTE REPRODUCTION

THE MEASUREMENT GAP IN THE AI ERA

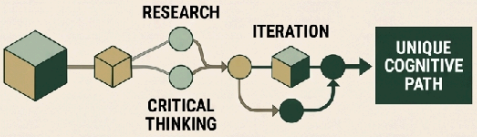
Tasks based on pure information reproduction (summarizing, conceptual synthesis) no longer discriminate between students who understand the material and those who simply use AI.



Reward for Uncritical Delegation: When schools reward rote output, they inadvertently teach students to delegate cognitive resources to machines rather than mobilizing their own thought processes.

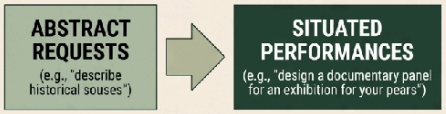
FROM PRODUCT TO PROCESS

Assessment must pivot from observing the final "indistinguishable" piece of work to observing the unique cognitive path the student took to create it.



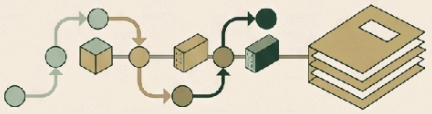
THE VILLARROEL REDESIGN BLUEPRINT

CONTEXTUALIZE THE STIMULUS



Replace abstract requests with situated performances.

IMPLEMENT TRACEABILITY & PROCESS LOGS



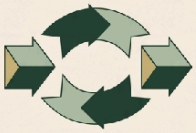
Break tasks into stages—research, draft, and critical revision—requiring students to maintain logbooks documenting their logical steps and discarded ideas.

TRANSPARENT ANALYTIC RUBRICS



Share rubrics before the task begins so they serve as a map for learning rather than just a justification for a final grade.

FEEDBACK AS "FEED-FORWARD"



Use feedback as a prospective tool that orients the next learning step, rather than a retrospective judgment of what has already been done.

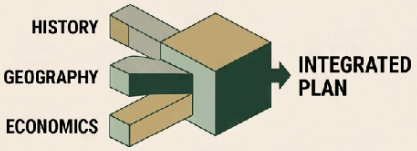
ANCHORING TASKS IN THE REAL WORLD

SITUATED AUTHENTICITY



AI lacks physical and social anchoring; tasks rooted in specific classroom investigations, field data, or local community contexts are structurally resistant to automation.

MULTIDISCIPLINARY CHALLENGES



Designing a plan for a local cultural asset requires integrating history, geography, and economics—a task requiring human intuition and local experience.

MANAGING TIME & SUBJECTIVITY

REAL CONSTRAINT	SUSTAINABLE OPERATIONAL STRATEGY
TIME INVESTMENT	Balanced plan: few focal authentic tasks + quick low-stakes tests; share models with colleagues.
MARKER SUBJECTIVITY	Use analytic rubrics with observable descriptors and conduct cross-marking sessions on samples.
LARGE CLASS SIZES	Utilize short presentations in rotation, peer assessment, and selective sampling of work.

Conclusions and Metacognitive Reflection Prompts

This resource has translated the principles of authentic assessment into operational methodologies for moving beyond textual reproduction. The collapse of tasks of

reproduction in the era of generative models makes it necessary to shift attention from the product to the process; the blueprint of Villarroel provides the key interventions - contextualising the stimulus, making the task traceable, sharing the criteria in advance, using *feedback* as feed-forward; and the anchoring to authentic and multidisciplinary products makes the test resistant to automated delegation. The constraints of time, subjectivity and class size do not justify surrender, but require conscious organisational choices.

To transfer these core ideas into one's own practice, the following reflection prompts are proposed:

- Take a traditional test in your subject and redesign it according to the blueprint: which elements of the stimulus did you contextualise, and which traces of the process did you make observable?
- In what way could you use a generative model as support for drafting descriptive *feedback* without delegating to it the validation of the judgement?
- Which strategy for reducing subjectivity (analytic rubrics, model examples, cross and blind marking) would you consider most sustainable in your context, and with which colleagues could you share it?

Technical Metadata Sheet

Metadata Field	Specification / Parameter Value
Identifier	OER 6.3.2
Title	Moving Beyond Rote Text Reproduction: Methodologies for Real-World Tasks
Creator	University of Foggia - Italy
Description	Complementing OER 6.3.1, the resource illustrates redesign methodologies that move beyond tasks of pure textual reproduction, which have lost measurement capacity in the era of generative models. It applies the redesign blueprint of Villarroel and colleagues to turn a traditional test into an authentic task and treats feedback as feed-forward through transparent rubrics. The practical focus is on governing real constraints — time, subjectivity of correction, large classes — with sustainable organisational strategies.
Keywords	authentic tasks; rote reproduction; redesign blueprint; Villar\u00f3el; feed-forward feedback; rubrics; multidisciplinarity; assessment constraints
Language	EN (English)
Affiliation	University of Foggia - Italy
Contributor	STIAEM Transnational Scientific Board (peer review, Rhodes validation protocol)
Media Format	Text; Infographics
Genre	Training Manual; Educational Guide
URL	
Citation	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://link.springer.com/article/10.1007/BF02504676 Villarroel, V., Bloxham, S., Bruna, D., Bruna, C., & Herrera-Seda, C. (2018). Authentic assessment: Creating a blueprint for course design. Assessment & Evaluation in Higher Education, 43(8), 1287-1301. https://doi.org/10.1080/02602938.2017.1412396 Villarroel, V., Boud, D., Bloxham, S., Bruna, D., & Bruna, C. (2019). Using principles of authentic assessment to redesign written examinations and tests. Innovations in Education and Teaching International, 58(1), 38-49. https://doi.org/10.1080/14703297.2018.1564882 Wiggins, G. (1990). The case for authentic assessment (ERIC Digest ED328611). ERIC Clearinghouse on Tests, Measurement, and Evaluation. https://eric.ed.gov/?id=ED328611 Redecker, C. (2017). European framework for the digital competence of educators: DigCompEdu (Y. Punie, Ed.). Publications Office of the European Union. https://doi.org/10.2760/159770 CAST. (2024). Universal Design for Learning guidelines version 3.0. CAST. https://udlguidelines.cast.org/ European Commission. (2022). Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and

	learning for educators. Publications Office of the European Union. https://doi.org/10.2766/153756
Educational Level	Secondary School
Associated Module / MOOC Video Lesson	Module 6 - Classroom Practice and Authentic Assessment / Video 6.3 - Designing Authentic Assessment Tests
Learning Objectives	recognise why reproduction tasks have lost measurement capacity in the era of generative models; apply the redesign blueprint of Villarroel and colleagues to turn a traditional test into an authentic task; govern real constraints (time, subjectivity of correction, large classes) with sustainable organisational strategies.
Rights	Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0)
Date Available	13/09/2026
Date Modified	

References and Webography

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://link.springer.com/article/10.1007/BF02504676

Villarroel, V., Bloxham, S., Bruna, D., Bruna, C., & Herrera-Seda, C. (2018). Authentic assessment: Creating a blueprint for course design. Assessment & Evaluation in Higher Education, 43(8), 1287-1301. https://doi.org/10.1080/02602938.2017.1412396

Villarroel, V., Boud, D., Bloxham, S., Bruna, D., & Bruna, C. (2019). Using principles of authentic assessment to redesign written examinations and tests. Innovations in Education and Teaching International, 58(1), 38-49. https://doi.org/10.1080/14703297.2018.1564882

Wiggins, G. (1990). The case for authentic assessment (ERIC Digest ED328611). ERIC Clearinghouse on Tests, Measurement, and Evaluation. https://eric.ed.gov/?id=ED328611

Redecker, C. (2017). European framework for the digital competence of educators: DigCompEdu (Y. Punie, Ed.). Publications Office of the European Union. https://doi.org/10.2760/159770

CAST. (2024). Universal Design for Learning guidelines version 3.0. CAST. https://udlguidelines.cast.org/

European Commission. (2022). Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators. Publications Office of the European Union. https://doi.org/10.2766/153756