Designing transparent evaluation criteria and rubrics with AI

Related curriculum: M4-T2-MLC4

1 ANALYSE the task	2 NEGOTIATE criteria	3 AUDIT with AI	4 TEST with evidence

GROUP ACTIVITY: RUBRIC CLINIC WITH USER TESTING

Activity: The teaching team redesigns a rubric from the student perspective: first it agrees on which evidence really matters, then it uses AI as a clarity auditor and finally tests the rubric with sample cases before validating it.

Duration: 60 minutes

Work format: Groups of 4 or 5 teachers, preferably from the same programme, professional family or teaching team.

Materials: A real assessable task, a generative AI tool, a rubric template, two anonymised student evidence samples or fictional cases and a shared document.

Key idea: a transparent rubric is not only for grading; it also helps students understand what good work means before they submit it.

Roles within the group

Role	Main responsibility
Facilitator	Keeps time and ensures that all voices participate.
Curriculum guardian	Checks that the criteria are aligned with learning outcomes, assessment criteria or competences.
Student advocate	Checks whether the language would be understandable for students before starting the task.
AI operator	Enters the agreed prompts and saves the relevant responses.
Rapporteur	Records agreements, changes and AI suggestions that are not accepted.

Activity instructions

1. Choose the task: The group selects a real task it already assesses: presentation, project, workshop practice, report, professional simulation, final product or portfolio.

2. Name the evidence: Before discussing marks, the group answers: what will students have to produce or demonstrate to provide evidence of learning?

3. Silent criteria: Each teacher individually writes three criteria they consider essential. Then the group compares them and groups repeated or very similar criteria.

4. First teacher agreement: The group agrees on 4 or 5 criteria and removes those that assess untaught aspects, generic attitudes or elements that cannot be observed.

5. AI audit: AI is asked to review the criteria as if it were an assessment expert: it must detect ambiguities, overlapping levels, bias, non-transparent language and criteria that are difficult to evidence.

6. User testing: The group applies the rubric to two real anonymised evidence samples or two fictional cases. If two teachers would interpret the same descriptor differently, that descriptor must be rewritten.

7. Student version: The final rubric is written in understandable language and includes a short opening note: what will be assessed, what will not be assessed and how students can use it to improve before submission.

8. Shared closing: Each group presents one important change made thanks to AI and another change decided by the professional judgement of the teaching team.

Group work template

Element	Teaching team agreement
Task or product to be assessed	
Related learning outcome, competence or curricular criterion	
Observable evidence students will submit or demonstrate	
Aspects that will NOT be assessed to avoid confusion	

From teacher criteria to transparent criteria

Initial criterion	Problem detected	Criterion reformulated for students
	Is it too general? Can it be observed? Does it depend on personal preferences?	
	Would students understand what they need to do to improve?	
	Is it aligned with what has been taught and with the task?	
	Does it avoid assessing aspects external to learning?	

Suggested prompt for the AI audit

Act as an expert assessment advisor for Vocational Education and Training. Review this task and these criteria. Detect ambiguous criteria, descriptors that are difficult to observe, overlapping levels, possible bias and aspects that students might not understand. Suggest improvements, but justify each change. Do not invent learning outcomes that are not included in the information provided.

Rubric user testing

Evidence reviewed	Doubt or disagreement between teachers	Adjustment made to the rubric
Evidence A		
Evidence B		
Criterion with the most disagreement		
Descriptor that needs clarification		

Final product

The group will submit: a revised student-facing rubric, a short transparency note and a justification of three decisions made during the process.

Final submission	Must include
Transparent rubric	4 or 5 criteria, differentiated levels, understandable language and weights if appropriate.
Note for students	What will be assessed, what will not be assessed, how the rubric will be used and how it can help improve.
AI use record	Prompt used, proposal received, accepted changes, rejected changes and reason.

Criteria for assessing the activity

- The rubric is aligned with the task and the learning to be assessed.
- The criteria describe observable evidence rather than general impressions.
- The levels distinguish the quality of the work without overlapping.
- The language can be understood by students before they complete the task.
- The teaching team documents what AI contributed and which decisions were human.

Suggested answer key for the facilitator

There is no single correct answer. The activity is well completed if the group shows that it has moved from intuitive or barely visible criteria to shared, observable criteria that students can understand.

- AI should appear as a tool for review and comparison, not as the sole author of the rubric.

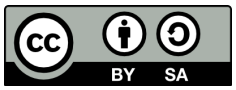
- Testing with evidence should lead to real adjustments in the descriptors.
- The final version should allow different teachers to interpret the criteria in a similar way.
- The rubric should help students know how to improve before submitting, not only after receiving the mark.

Classroom transfer closing

Final question	Group answer
Which phrase in the rubric could create doubts for students?	
Which criterion has improved the most after the AI audit?	
Which human decision should not be delegated to AI?	

Copyright © 2026 VITA Consortium

This work may be reused and adapted under the terms of the Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).



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 National Agency. Neither the European Union nor the National Agency can be held responsible for them.