

# CO-LAB Guidelines for Assessing Collaborative Learning in the Classroom

AUTHOR: LUIS VALENTE (RESEARCHER AND CONSULTANT, UNIVERSITY OF MINHO'S INSTITUTE OF EDUCATION, ICT COMPETENCE CENTRE)



## A - General Assessment Guidelines

The general guidelines in Section A of this document apply to all [CO-LAB scenarios](#) and provide information on three key points: The purpose of assessment; what to assess; and how to assess.

### A1 – The purpose of assessment

Determining the purpose of assessment should be the first step in the design of any assessment of learning. In this way we can ensure the effectiveness of the assessment tools and have a better understanding of their intrinsic value. The different types of assessment below, each have a specific purpose:

- (a) Formative assessment, allowing to redefine improved strategies and processes of learning, by involving student groups and teachers, using simple assessment instruments, well-targeted and clearly defined objectives;
- (b) Self-assessment, centred on individual and personal reflection, focusing on commitment, motivation, engagement in teamwork and overcoming personal goals. Group goals can be realized with checklists or progress level scales (e.g. I cannot do this yet, I can do this alone, I am fully capable etc.) applied over a pre-set timeline (e.g. before a new topic), systematically (e.g. weekly) or at particular milestones while performing projects or tasks (e.g. at the beginning, in the middle, at the end of the project or task);
- (c) Summative assessment, aiming to provide a formal academic record of students' learning progress and used to establish proficiency levels.

Once the purpose has been established, assessment will be closely related to skills and knowledge (what to assess) and to the type and complexity of instruments to be used (how to assess). Considering the purpose, we can adopt a flexible assessment model to provide useful data for formative teacher intervention and to induce the adoption of self-regulatory actions of the student and the group.

### A2 - What to assess

We must consider the choice of assessment model according to the project or task to which it applies, which could include:

- (a) The individual and collective learning outcomes;
- (b) Teamwork and group goals including individual commitment;
- (c) The methodological process adopted to implement and complete the project or task;
- (d) The product or;
- (e) A mix of targets.

However, considering we are seeking to develop collaborative learning scenarios, the emphasis should be put on assessing what leads to the development of team work skills and the results or knowledge achieved by the group, much more than evaluating individual successes.

When assessing knowledge it is useful to consider the four categories proposed by Anderson and Krathwohl (2001): factual; conceptual; procedural and; meta-cognitive. The factual and conceptual categories constitute low level knowledge, relating to the recognition of facts and enumerations (in other words, the 'what'). The procedural and metacognitive categories on the other hand constitute high level knowledge, relating to skills needed to solve new problems or to make use of existing knowledge in new situations (in other words the 'how'). The procedural and meta-cognitive categories are related to creativity, entrepreneurship, and the planning and design of projects, for example.

Factual knowledge includes the core elements of curriculum content, or what students should know about each topic, such as details, specific data, facts, dates, etc. It requires memorization, it is verifiable by direct comparison, and it can be easily assessed by students themselves or by their peers.

Conceptual, or declarative knowledge, refers to the understanding of interrelationships between pieces of knowledge. This type of knowledge enables us to understand the organization of items into categories, to





identify hierarchies and structures, to recognize elements by observing specific characteristics, to recognize models, to understand generalizations, etc.

Procedural knowledge refers to being able to do something by applying criteria, algorithms, appropriate techniques and methodologies, and using different skills and knowledge or combining them in different and innovative ways. Procedural knowledge is required to run research projects, to establish relationships between different data and to understand cause-effect relationships.

Metacognitive knowledge includes the evaluation of knowledge itself, its nature, the construction process, its importance and its effects or consequences. Metacognition is the sum of what a student has discovered, perceived or apprehended about a particular subject. Metacognitive knowledge is invoked to explain phenomena or events studied before or, for example, in the production of explanatory theories and creative expression.

### A3 - How to assess

The benefit of assessment relates directly to how effective it is. Its internal and external credibility, legitimacy and reliability recognized by stakeholders has a significant impact on its value and potential. Therefore, deciding when to assess and how to assess are interdependent aspects. For example, if we are to make an assessment at the end of each class or workshop, the instruments will have to be simple and not too long. Complex instruments could provide us with biased data or hasty information. Perhaps a checklist would be adequate, but a written reflection would probably be of little use.

On the other hand, if our goal is to provide information to the learners themselves about their progress, we need instruments to collect information that we can display in the shape of scales, trend lines or graphics. Here our choice of instruments include Likert scales, checklists, models based on rubrics and quizzes with detailed feedback.

In those situations where the purpose of the assessment is to provide further information, particularly to the teacher or for external elements to the learning process, it might be appropriate, for example, to opt for infographics, mind maps, portfolios or online written documents such as blogs or multimedia presentations. Any of the assessment methods may involve the use of digital media tools, but not necessarily. In the case of using digital media, in some cases, statistical analysis and content analysis tools can be useful. In other cases we can adopt rubrics to assess specific items or learning outcomes.

#### A3.1 Advice on assessment with rubrics

The use of rubrics is a speedy and interesting way to collect information on progress in learning and provide feedback to stakeholders. However, the construction of rubrics does not usually excite teachers, mainly because they think it is complex and difficult to achieve. However, this does not have to be the case, and we hope the guidelines below will help you in this task.

##### A.3.1.1 Let's begin by asking: What is a rubric?

Although in the literature very elaborate definitions can be found, we would like to adopt a more pragmatic definition: A rubric is a scoring tool that lists the criteria for a piece of work and articulates gradations of quality for each criterion, from excellent to poor (Goodrich, 1996). Rubrics therefore allow us to build a table with indicators of what we want to assess - the criteria - and a grading scale consisting of verifiable quality descriptors - progress, quality or proficiency levels.





### A.3.1.2 Structure of a rubric

Table 1 – Sample structure of a classic rubric

List of criteria	Quality descriptors activity or outcome			
	Highest level (4)	Medium + level (3)	Medium - level (2)	Lower level (1)
Criterion A	Descriptor A4	Descriptor A3	Descriptor A2	Descriptor A1
Criterion B	Descriptor B4	Descriptor B3	Descriptor B2	Descriptor B1
Criterion N	Descriptor N4	Descriptor N3	Descriptor N2	Descriptor N1

### A.3.1.3 Why use rubrics?

Firstly, the use of rubrics is important because they are very clear instruments for assessment. They are flexible and effective for self- and peer-assessment. Rubrics are also important because they allow students to be engaged in their design. In this way, students identify themselves better with the assessment objectives, with the criteria and scales to measure these criteria, and teachers can give formative feedback if the descriptors are sufficiently clear and objective. Goodrich (1996) states that "rubrics can improve student performance, as well as monitor it, by making teachers' expectations clear and by showing students how to meet these expectations". The most common argument in support of the use of rubrics is that they help to define what we mean by quality.

### A.3.1.4 How to build a rubric

The first advice we can give is to involve students. Indeed, when they participate in defining the criteria and identifying quality descriptors that help them to understand whether the criteria have been met, the rubric becomes an indispensable tool. Sometimes, teachers tend to focus on the task and not on learning outcomes, and this is a significant mistake. As Brookhart (2013) points out, "the biggest mistake teachers make when they use rubrics with performance assessment is that they focus on the task, the product, and not the learning outcome or proficiency the task is supposed to get students to demonstrate." In this sense, it is essential to define the most appropriate and important criteria that is meant to be assessed. "These should not, generally, be characteristics of the task itself, but rather characteristics of the learning outcome the task is supposed to indicate."

The simple question we should ask ourselves is - What features of students' work provides us with evidence that they have acquired certain skills or knowledge? Then, by decomposing the question into different characteristics, as suggested by Susan Brookhart, we can assemble the following table.

Table 2 - Desired Characteristics of Criteria for Classroom Rubrics (Brookhart, 2013, p. 25)

Characteristics	
The criteria are...	Explanation
Appropriate	Each criterion represents an aspect of a standard, curricular goal, or instructional goal or objective that students are intended to learn.
Definable	Each criterion has a clear, agreed-upon meaning that both students and teachers understand.
Observable	Each criterion describes a quality in the performance that can be perceived (seen or heard, usually) by someone other than the person performing.
Distinct from one another	Each criterion identifies a separate aspect of the learning outcomes the performance is intended to assess.
Complete	All the criteria together describe the whole of the learning outcomes the performance is intended to assess.
Able to support descriptions along a continuum of quality	Each criterion can be described over a range of performance levels.



Once the specific criteria is set, we need to ask ourselves about the quality of student work to be demonstrated at every level, ranging from the highest to the lowest. This question should have four answers to match as many quality levels. To make this task easier, we can turn back to Susan Brookhart’s expertise, as shown in the following table.

**Table 3 - Desired Characteristics of Descriptions of Levels of Performance for Classroom Rubrics (Brookhart, 2013, p. 28)**

Characteristics	
The descriptions of levels of performance are ...	Explanation
Descriptive	Performance is described in terms of what is observed in the work.
Clear	Both students and teachers understand what the descriptions mean.
Cover the whole range of performance	Performance is described from one extreme of the continuum of quality to another for each criterion.
Distinguish among levels	Performance descriptions are different enough from level to level that work can be categorized unambiguously. It should be possible to match examples of work to performance descriptions at each level.
Center the target performance (acceptable, mastery, passing) at the appropriate level	The description of performance at the level expected by standards, curriculum goal, or lesson objective is placed at the intended level on the rubric.
Feature parallel descriptions from level to level	Performance descriptions at each level of the continuum for a given standard describe different quality levels for the same aspects of the work.

### A.3.1.5 Strategies to build a rubric

Although it is of course important that each teacher adapts their assessment strategies to suit the learners in their own class, the curriculum and their familiarity with building specific assessment tools, we can nevertheless present some suggestions.

Let’s firstly consider an open strategy directly involving pupils put forward by Heidi Goodrich (1996), which consists of the following seven steps:

1. *Look at models: Show students examples of good and not-so-good work. Identify the characteristics that make the good ones good and the bad ones bad.*
2. *List criteria: Use the discussion of models to begin a list of what counts in quality work.*
3. *Articulate gradations of quality: Describe the best and worst levels of quality, then fill in the middle levels based on your knowledge of common problems and the discussion of not-so-good work.*
4. *Practice on models: Have students use the rubrics to evaluate the models you gave*
5. *Use self- and peer-assessment: Give students their task. As they work, stop them occasionally for self- and peer-assessment.*
6. *Revise: Always give students time to revise their work based on the feedback they get in Step 5.*
7. *Use teacher assessment: Use the same rubric students used to assess their work yourself.*

Another option involving students, is suggested by Rebecca Anderson (1998), and refers to a four-by-four methodology:

1. *In small groups, students identify four characteristics of good quality in the task they are completing.*
2. *Each group writes four characteristics on the chalkboard or overhead.*
3. *One reporter from each group discusses one characteristic from the four characteristics, perhaps the one the group discussed the most or had the greatest passion about. As the facilitator, instructors can help identify similarities and differences between the groups’ criteria.*



4. *After each group has shared its criteria, the instructor asks the entire class to reach consensus about four criteria to be used when grading the task.*
5. *In small groups, students write four descriptors with a corresponding score for each of the four criteria. The scores range from 1 to 4, with 4 being the highest score.*
6. *After each group has shared its criteria, the class reaches group consensus.*

Susan Brookhart also proposes a deductive (top-down) or inductive (bottom-up) approach to constructing rubrics. In the top-down approach, the assessment criteria are built on the curriculum guidelines, so, prior to the task. It is necessary, therefore, that the goals are clearly defined. Under this approach, Brookhart suggests:

1. *Create (or adapt from an existing source) a conceptual framework for achievement.*
2. *Write general scoring rubrics using these dimensions and performance levels.*
3. *For teacher scoring, you may adapt the general scoring rubrics for the specific learning goal for the performance you will be scoring.*

In either case (whether the rubrics remain general or are adapted to more specific learning goals), use the rubrics to assess several students' performances, and adapt them as needed for final use.

In the bottom-up option, the rubric is typically built after completion of the activity. Some students work is used to define the assessment framework. This strategy is suitable for situations where performance descriptors are not easily definable prior to the activity. Brookhart suggests:

1. *Get a dozen or more copies of students' work.*
2. *Sort, or have students sort, the work into three piles: high, medium, and low quality work.*
3. *Write, or have students write, specific descriptions of why each piece of work is categorized as it is.*
4. *Compare and contrast the descriptions of work and extract criteria or dimensions.*
5. *For each of the criteria identified in step 4, write descriptions of quality along the dimensions, for as many levels as needed.*

### A3.2 The simplicity of a checklist

In our day-to-day lives, we use checklists often without clearly understanding them or thinking deeply about them. They can in fact be very useful instruments for assessing students. The simpler they are, the higher their value.

As stated by Atul Gawande (2010) "avoidable failures are common and persistent, not to mention demoralizing and frustrating, across many fields" and "the reason is increasingly evident: the volume and complexity of what we know has exceeded our individual ability to deliver its benefits correctly, safely, or reliably." To reduce these common errors we need simple and effective strategies to assess and validate our knowledge. These simple strategies involve the use of checklists. Checklists provide a kind of cognitive network, helping us to identify flaws of attention, memory and accuracy. Yet they also have limits, so it is decisive to "identify which kind of situations checklists can help and which ones they can't."

### A3.3 What should a checklist look like?

There are good and bad checklists. Bad checklists are vague and imprecise, too long, difficult to use and impractical; they treat their users as ignorant, Gawande notes. Good checklists are accurate, efficient, well-targeted and easy to use in all situations, provide clear clues on the most critical and important aspects and are, above all, practical. Checklists help people, including experts, to manage a complex process to establish priorities clearly, to work better in teams. Good checklists really work.

Before starting to build a checklist, we need to clearly define its purpose and to bear in mind that it should be as short as possible. The ideal is to have between five and ten items, that is, within the limitations of our working memory. When it is longer, it is likely to induce distraction. The wording must be simple, accurate and must belong to the common lexicon of context. The layout must have a sober typography, with appropriate use of case, be easy to read, ideally occupying no more than one page. It should be of quick application.





## B - Specific guidelines for collaborative learning

The guidelines proposed below follow the general advice given in Section A of this document, to demonstrate to you how this can be used to construct tools to assess collaborative learning. The example [CO-LAB Scenario 'Collaborative work – Towards a Healthy City \(CCL, Europe\)'](#) has been used as a reference to provide the learning context, and should be consulted in tandem with the assessment guidelines proposed below. This learning scenario has been analysed in terms of the appropriate formative, self- and peer-assessments that should accompany it, and a specific rubric and two checklists are proposed for this purpose. Please note that these tools have been designed to be easily adaptable, so that you can also use them in relation to other [CO-LAB Scenarios](#) as well as any new collaborative learning scenario you design yourself. This means that you can tailor the tools to suit your own specific teaching and learning context.

### B1 - Purpose of assessment: Formative Assessment

What to assess	How to assess
Competences to do teamwork by considering playing different roles within the group.	Through feedback from teachers, other observers (e.g. colleagues), or self-assess. By using checklists or rubrics.
Ability to synthesize the results of the project or task.	Through infographics. By presenting results using concept maps or mind maps.
Expertise in "mentoring" colleagues in the group.	Through feedback from other group members. Through feedback (comments) from teachers or other observers.
The ability to critically analyse the decisions taken by the group.	Through group reports and progress reports. Using checklists.

### B2 - Purpose of assessment: Self-assessment and peer assessment

What to assess	How to assess
Competences to do teamwork by considering playing different roles within the group.	By using checklists.
Ability to synthesize the results of the project or task.	By using rubrics.
Expertise in "mentoring" colleagues in the group.	By using checklists.
The ability to critically analyse the decisions taken by the group.	By using rubrics.



## B3 - Sample tools to assess students' collaborative work

### B3.1 - Tool: Rubric

This rubric aims to assess collaborative work skills and may be considered as an example. No knowledge or results are assessed, but the model can be adapted for that purpose.

<b>RUBRIC FOR ASSESSING STUDENTS' COLLABORATIVE SKILLS</b>					
Applies to: Group and/or group members individually					
Criteria	Quality descriptors				Score
	(4 points)	(3 points)	(2 points)	(1 point)	
<b>Focus on tasks</b>	The focus of the work was constant and all tasks were completed properly.	The focus was mostly directed to what should be done and most of the tasks have been completed properly.	The focus was often directed to what should be done, but some tasks were not completed properly.	Rarely the focus of the work was directed to what should be done and many tasks were not completed properly.	
<b>Mutual help</b>	Whenever there was difficulties, these were faced and help was provided.	Most of the difficulties were faced and nearly always help was provided.	Many of the difficulties were faced and often help was provided.	Difficulties were rarely faced and help was rarely provided.	
<b>Responsibility and reliability</b>	Systematically punctual for group meetings and does not depend on others to do the assigned work.	Usually punctual for group meetings and most of time does not depend on others to do the assigned work.	Often punctual for group meetings and most often does not depend on others to do the assigned work.	Rarely punctual for group meetings and most often depends on others to do the assigned work.	
<b>Team Leader's performance</b>	The entire work plan was duly presented to team members and completion of tasks of each team member was monitored until complete.	The work plan was mostly presented to team members and the majority of tasks were monitored until completion.	The work plan was presented to almost all team members and a large number of tasks were monitored until completion.	The work plan was presented to some of team members but only a few tasks were monitored until completion.	
<b>Team Reporter's performance</b>	All progress updates explain clearly what each of the team members are doing.	Most of the progress updates explain clearly what each of the	Many progress updates explain clearly what each of the team members are doing.	Only a few progress updates explain what each of the team	







## RUBRIC FOR ASSESSING STUDENTS' COLLABORATIVE SKILLS

Applies to: Group and/or group members individually

Criteria	Quality descriptors				Score
	(4 points)	(3 points)	(2 points)	(1 point)	
<b>Organiser's performance</b>	All team members received their logins on time to access online tools, blogs or websites.	Most of the team members received their logins on time to access online tools, blogs or websites.	Many of the team members received their logins on time to access online tools, blogs or websites.	Only a few members received their logins on time to access online tools, blogs or websites.	
<b>Lead Researcher's performance</b>	All research questions, sources and tools were previously validated.	The majority of research questions, sources and tools were previously validated.	Many of the research questions, sources and tools were previously validated.	Just a few research questions, sources or tools were previously validated.	
<b>TOTAL</b>					

### Meaning of the results

Our goal is for students to achieve the highest level (4 points = 100%) for each criterion. If assessment focuses only on one activity or a specific period of time in which a student has only played one role within the group, in this case we take into account only the criteria C1, C2 and C3 in addition to one of the criteria C4, C5, C6 or C7 (see table below). However, the assessment should apply to activities in which a student has played all the roles. It is up to teachers to decide which criteria to apply. It is also possible to apply the group performance criteria separately. We can consider an example in the rubric grid below.

(C <sub>n</sub> ) Criteria	4 Points	3 Points	2 Points	1 Points	SUM
C <sub>1</sub> Focus on tasks	100%	75%	50%	25%	X
C <sub>2</sub> Mutual-help	100%	75%	50%	25%	X
C <sub>3</sub> Responsibility and reliability	100%	75%	50%	25%	X
C <sub>4</sub> Team Leader role performance	100%	75%	50%	25%	X(*)
C <sub>5</sub> Team Reporter role performance	100%	75%	50%	25%	
C <sub>6</sub> Organiser role performance	100%	75%	50%	25%	
C <sub>7</sub> Lead Researcher role performance	100%	75%	50%	25%	

(\*) Choose one

Once you have calculated the student's performance using this rubric grid, it will then be necessary to assign this score a certain weight, so that it can be used as part of the student's summative assessment.

### B3.2 Tool: Checklist for self-assessment

The purpose of this checklist is to give the students themselves and other stakeholders (peers and teachers) a clear idea about their ability to work in a team, to play their role and to know and understand the roles of others. This can be achieved by the student identifying themselves with the goals of the group and the commitment to the work of the group, and by being able to assess their contribution and that of peers.





## CHECKLIST FOR SELF-ASSESSMENT OF STUDENTS' COLLABORATIVE SKILLS

Applies to: the student individually

Yes/No	Statements
	I know exactly what my tasks in this group are.
	When I faced difficulties, I looked for help both inside and outside of the group.
	I presented to the group sources or documents relating to the subject (e.g. books, texts, web sites, videos).
	I have put forward to the group the issues which are relevant to our work.
	I helped my colleagues when they asked me for help.
	I helped my colleagues when I realized that they had problems, even without them asking me for help.
	I know perfectly well what are the roles and tasks of each colleague in my group.
	Usually I express my views and opinions clearly to my group.
	I feel able to assess the contributions of my peers' in the work done by our group.
	I believe I would not be able to do a better job myself than what was achieved by my group.

### Meaning of the results

Our goal here is for students to respond 'yes' to all items. Although we cannot establish student's performance only by summing up the items they responded yes to, this checklist does help give us a holistic idea of their individual commitment to the goals of the group.

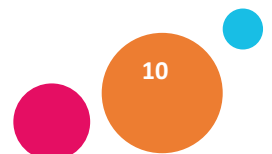
In this case, since there are 10 items in the checklist, each item labelled 'yes' can correspond to 10% of the maximum score. This way, we can set levels similar to that adopted in the previous tool. Where the total amounts to below 50% this shows that the student has difficulties in working in groups.

As before, the relative weight given to the student's checklist score will then need to be defined in the summative assessment.

### B3.3 Tool: Checklist for peer-assessment

The purpose of this checklist is to help students recognize others' ability to work in teams, using it as a learning experience to improve their own collaborative learning skills, and identifying key competences necessary for collaborative work. Each group chooses another group to evaluate using the checklist.

Please note that not all statements in the checklist below will apply to a specific activity. As a first task the class should identify the statements that should be applied and use only those for evaluation. In addition, it might be necessary to add some statements.





## CHECKLIST FOR PEER ASSESSMENT OF STUDENTS' COLLABORATIVE SKILLS

Applies to: the team/group

Yes/No	Statements
	All group members actively contributed to the final product.
	Group members gave each other support and feedback.
	When the group was having trouble, other groups spontaneously helped.
	When the group asked for help, other groups helped them immediately.
	In the end, everyone seemed satisfied with their group's work.
	Group members exchanged and negotiated between them their ideas, strategies, tools and/or resources to carry out the activity.
	The group defined the tasks and the role of each member early on.
	The group asked the opinion and suggestions of other groups.
	The group assisted and gave advice to other groups.
	The group requested comments to other groups before finishing the activity.
	The group accepted critical comments from other groups.

### Meaning of the results

Our goal here is for students to respond 'yes' to all items. Some items refer to active attitudes and others refer to proactive attitudes that tend to strengthen team consciousness. On that basis, all items must be considered and every "Yes" should be assigned 1/n%. Once again, in order to reflect this checklist score in the summative assessment it will be necessary to establish its relative weight.

### References

- Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York, NY: Longman.
- Anderson, R. S. (1998). Why talk about different ways to grade? The shift from traditional assessment to alternative assessment. *New Directions for Teaching and Learning*, 74(Summer), 5-16.
- Brookhart, S. M. (2013). *How to create and use rubrics for formative assessment and grading*. Alexandria, VA: ASCD.
- Evans, J. A., & Foster, J. G. (2011). Metaknowledge. *Science*, 331(11), 721-725. doi:10.1126/science.1201765
- Gawande, A. (2010). *The Checklist Manifesto: How to Get Things Right*. London: Profile Books.
- Goodrich, H. (1996). Understanding rubrics. *Educational Leadership*, 54(4), 14-17.

*The work presented in this document is supported by the European Commission's Erasmus+ programme – project Collaborative Education Lab (Grant Agreement 2015 – 3068 /001 - 001). The content of this document is the sole responsibility of the author and it does not represent the opinion of the European Commission and the Commission is not responsible for any use that might be made of information contained herein.*

