THE DESIGN OF A RUBRIC FOR DEFINING AND ASSESSING DIGITAL EDUCATION SKILLS OF HIGHER EDUCATION STUDENTS

Herve Platteaux, Emanuelle Salietti, Centre NTE DIT – University of Fribourg, Laura Molteni, University of Fribourg, Switzerland

Summary

The usage of rubrics is nowadays a developing trend in the world of Higher Education. One can think about two main reasons for this. First, even if there is no definitive proof, rubrics seem to be adequate for supporting the learning of complex skills, in particular for formative assessment. Rubrics are then finding a natural place in HE institutions in the context of the 21ˢᵗ Century where digital education skills become more and more important and need to be well defined and assessed. Secondly, rubrics are based on very easy principles and this simplicity may contribute to the trend noted.

However, our experience of doing rubrics for defining and assessing the students' digital education skills revealed us that the design of rubrics needs its basic principles but also additional rules in order to make a rubric that can be used as an efficient assessment tool. In this perspective, we decided to compile and explain in this article the rules that we have applied during our rubric design work. Some rules were found in the literature; other ones were elaborated during our work progression. With this compilation, we want to bring the reader concrete guiding elements and steps for the design of rubrics. A general rule seems to emerge from our work: a rubric maker should always try to distinguish between all the aspects of the competences needed to perform a task and all the aspects of all the different levels that can be seen in the competences of a person who is performing the task.

Introduction

When joining a university cursus, students usually have a previous experience with a few computer tools such as Word, PowerPoint and entertainment social media. But they have to further develop their digital skills in order to learn their student job. The 21ˢᵗ century context makes digital education always and everywhere present in Higher Education. Then the Personal Learning Environment (PLE) of students is changing radically over the first semesters of their life at University. The pedagogical and technological aspects of their PLE change because of all digital skills that students have to develop for their learning tasks and for a good use of computer tools that can help to perform these tasks.

Concretely, students need to learn how various learning tasks, linked to their course activities, must be performed in the academic world (research of thematic information, academic

213

document writing and team collaborative working, etc.). For example, in order to do a good thematic research of information, they have to learn and use concepts like "information validity". Students also need to learn how to complete their computer and network know-how by developing usages of new features and new tools, usually unknown from young students. For example, they have to learn how to manage in Learning Management Systems, in scientific online databases and with bibliographical managers.

To help the students to learn about these multiple facets of digital education, we developed a competence framework that is structured on the description of the student learning tasks, with their usual steps and finalities, and on the proposition of the computer tools that can help students performing these tasks. The delivery of these description of tasks and propositions of tools, taking the form of an online guide (myple.ch), was the first achievement of a student support project at the University of Fribourg. A second achievement was the creation of a competence framework, written as a series of rubrics. With these rubrics, we want to propose, for all the tasks documented in myple.ch, a detailed competence description and a formative assessment tool (Platteaux, Sieber, & Lambert, 2017). With these resources, we aim to help guidance of individual students for developing their digital competences and for identifying their individual strengths and weaknesses.

This first version of our rubric leaded our team to the identification of a few defaults that can be overpassed when writing the rubric in order to make it an efficient assessment tool. The aim of this article is to present the analysis of these defaults and the solutions that were found to counterbalance them. Through this analysis, our article wants to be very pragmatic, to show principles and examples about rubrics taken from the scientific literature, and from our own work, and to bring guidelines that could bridge theoretical principles and concrete work results for the doing of a rubric. We are placing thus our work in the perspective of a recent claim:

> *"One of the major causes for problems in rubric design is that there is very little research on how to formulate clear, meaningful, unidimensional and differentiating dimensions that are used to describe skills' mastery levels"* (Rusman & Dirkx, 2017; p.4).

Analysis of rubrics' design, between a free simple shape and constraining rules

Recent works showed that the term rubric is used with many meanings (Dawson, 2015). Then, to introduce our analysis with a clear notion, we refer to the following two simple definitions of rubrics:

Global definition:

> *"A rubric has three essential features: evaluative criteria, quality definitions and a scoring strategy"* (Popham, 1997; p.72);

Operational definition:

> *"A rubric is a matrix containing the various factors of an assignment along one dimension (rows) and descriptors of the qualitative levels of accomplishment along the other dimension (columns)" (Anderson & Mohrweis, 2008; p.85).*

With such definitions, one can understand that the free basic shape of rubrics is advantageous for the description and the assessment of skills that are needed to perform a task:

- Natural basic structure: the two dimensions of the rubric structure favours a natural construction where lines show the aspects of a task (the evaluative criteria) and columns show the levels of accomplishment (the quality definitions);
- Convenient for substructures: different substructures can be easily built up and presented, by separating or regrouping series of lines and/or columns;
- As many lines as wished: the different activities of the described task, and/or their associated skills, can be explained with the desired number of details;
- As many columns as wished: the number of accomplishment levels can be chosen freely, ie. the desired precision of the assessment scale.

This freedom of design and this simplicity of rubric attract many educators who are looking for assessment tools. In particular rubrics are more and more present in educational institutions, in the context of the 21st century skills' development, because "rubrics are used as an instrument to support learning complex skills in schools" (Rusman & Dirkx, 2017; p.1). More precisely, authors like Lowe and her colleagues, working about the Information Literacy skills, assert the power of rubrics because they allow for objective, authentic assessment of student work: "Rubrics make clear to students the expectations of their instructors and provide consistent and transparent performance criteria" (Lowe, Booth, Stone, & Tagge, 2015; p.492). Other authors emphasize the need for more research to better understand how much the formative use of rubrics impacts learning (Greenberg, 2015).

Then the problematic of the rubric use is in the educational world: how can rubrics help the learning of complex competences and its assessment? With this perspective, the main aspect to be understood in the rubric design is perhaps not its matrix shape: "It may turn out that it is not rubrics per se (that is, rubrics as an assessment tool in a particular form), but the provision of focused learning goals, criteria, and performance descriptions in whatever form that supports learning and motivational outcomes for students" (Brookhart & Chen, 2015; p.364). With this in mind, we want here to review, in the literature and in our experience, the rules that can be listed for the writing of rubrics' evaluative criteria and for the definition of rubrics' quality scales.

Here below are the rules that we have found about the skills' evaluative criteria:

- Popham underlines how the task description writing is central: "the rubric's most important component is the set of evaluative criteria to be used when judging students' performances" (1997; p.73).
- For the writing of skills, it is very helpful to use, as a basis, the description of the associated task and, in addition, to establish a clear hierarchical cascading structure of the task, of its activities and of their actions (with the corresponding competence, skills and sub-skills). This structure simplifies much the work of design (our experience).
- The wording of skills is very important: "a balance between generalized wording, which increases usability, and detailed descriptions, which ensure greater reliability, must be achieved" (Rusman & Dirkx, 2017; p.3).
- "The purpose of internal validity evidence is to demonstrate the relationships among the criteria within a rubric." (Brookhart & Chen, 2015; p.355).

Here below are the rules that we have found about the skills' evaluation scales:

- The scoring strategy can be holistic or analytic (Popham, 1997; p.72):
    - "Using a holistic strategy, the scorer takes all of the evaluative criteria into consideration but aggregates them to make a single, overall quality judgment."
    - "An analytic strategy requires the scorer to render criterion-by-criterion scores that may or may not ultimately be aggregated into an overall score."
- "Rubrics can contain detailed grading logic, with numbers and even formulae; alternatively they can have no numbers, and be suggestive of broad quality levels" (Dawson, 2015; p.3).
- "For the scale to be generic enough to be applied in a variety of university courses, the descriptors need to refer to a spread of performances at each level. On the other hand, there is a risk that these statements may be too general and thus lead to inconsistent interpretation of the data" (Simon & Forgette-Giroux, 2001; p.105).
- "The levels in a rubric should be parallel in content, which means that if an indicator of quality is discussed in one level, it is discussed in all levels" (Rusman & Dirkx, 2017; p.3).
- "The consistency between the dimensions used within the performance indicator descriptions within and across rubrics was quite poor. Most rubrics used words signaling a mastery level only in one or two performance indicators per constituent sub-skills. Also across rubrics, many different verbal qualifiers or signaling words were used to describe the same scales" (Rusman & Dirkx, 2017; p.7).
- The number of levels in evaluation scales, who is right?
    - If the rubrics have few levels, there is a decrease of the assessment quality because: "People usually avoid extreme positions so a scale with only a few steps may, in actual use, shrink to one or two points" (Bandura, 2006; p.311);
    - A few assessment levels and a good differentiation between the levels maintain the assessment reliability (Berthiaume & Rege Colet, 2013).
- Rules and existing data can help much the choice of the adequate verbal qualifiers for the definition of a good evaluation scale (Rohrmann, 2007).

Application to the creation of a rubric for digital education skills of higher education students

At the beginning of the work presented in this article, we had a first version of a competence framework written as a rubric designed as a table and for a use that Dawson calls "Quality definition", where each cell defines "a particular evaluative criterion at a particular quality level" (2015; p.8). Also, in this case the evaluation strategy is holistic, according to Popham (1997). Figure 1 shows how our rubric was structured. One can see that the quality level was based on a three level scale. The competences needed to perform a complete learning task (for example: Writing an academic paper) was described by a series of tables, each table dedicated to a competence needed to perform an activity of the task (for example: Structuring a written document with a table of contents). Finally, every single line of a table was focused on a skill that can be mobilized to perform an activity (for example: Making an automatic table of contents in Word).

Activities	Actions	Level 1 "discover"	Level 2 "try and choose"	Level 3 "deepen"
Competence 1 for Activity 1	Indicator 1.1 for Action 1 of Activity 1	Description of every action at every level of competence C1 written with 2 types of evaluation criteria: • management of information • use of device		
	Indicator 1.2 for Action 2 of Activity 1			
	Indicator 1.3 for Action 3 of Activity 1			

Figure 1. Original rubric structure of our competence framework (Platteaux et al., 2017; p.3065)

In this first version, we identified three main defaults in the rubric cells: (a) not systematic description of the skills, (b) mixing of information elements that are descriptive and evaluative; (c) the three levels' scale does not allow the efficient assessment of the skills associated to the cells.

The article considers in the following lines how we improved the first version of our rubric. Doing this, we would also to link design principles of rubrics with concrete examples of rubrics illustrating how principles can be applied.

Let's see first how we improved the systematic description of the skills. Our main problem in the version 1 was a lack of consistency in the description of the different skills of each activity. More precisely, we noticed that is was very difficult to write different accomplishment levels of the same skill. Very concretely, there are two problems: (a) it is difficult to write the absence of a skill (for the low level cells) and (b) it is difficult to initiate and maintain consistency of the evaluation levels when the described skills have different sub-skills that are logically linked to different evaluation scales.

Then we decided to rewrite every skill of an activity in three steps: (a) regrouping all the description elements of the skill, (the elements that were dispersed in the version 1 through the different level cells); (b) listing of all the sub-skills of the skill (two types: knowledge and know-how) and (c) writing of the *upper level* of the skill.

We insist here on how much the second step helped us to progress towards a complete and clear description of all needed skills. Indeed, educators are used to think with actions of learners, linked to knowledge and/or know-how. Knowledge and know-how are then good helps for finding all the sub-skills of a skill. Furthermore, for a competence framework about digital skills, it is also really helpful to think about knowledge and know-how that are linked either to the computer tools, and their features, and either to the learning task itself.

We do advice also any rubric designer on the importance of the step 3, it means to target the achievement of an *upper level* skill description. With *upper level*, we mean that the rubric designer should first assemble all the skills that are mobilized by a student who is perfectly performing the activity. If you have such a result in front of your eyes, you can think about evaluation scales and skill levels. On the contrary, if you try to write down all the different levels of achievement of a skill that is not well defined, you are in trouble (methodological trap).

We see this trap like the cause of the second main default of our rubric version 1. It consists mainly in the mixing of two types of information: (a) description elements of the skills needed to perform a task activity and (b) definition elements of the assessment scale(s) for these skills. For our competence framework, our experience revealed that a clear way to distinguish the two types of information is to think to them as follows:

- If information refers to the task, for example a particular tool that can be used to perform this activity, this informative element is a description element of the task. It is then turned into a skill for a competence framework like in the following example:
    - The student can use the computer text treatment tools;

- If information refers to the way an individual person is able to perform a particular step of the task or to use a tool that can help to perform this step, this informative element must be taken into account in order to set the evaluation scales and the skill accomplishment levels. Usually such information is qualitative of quantitative, such as in the following examples:
    - The student can use four computer text treatment tools;
    - The student can use very well the computer text treatment tools;
    - The student can use all the principal features of the computer text treatment tools.

Concerning the third default of our rubric version 1, the solution that we have built can be said shortly: we turned to an analytic scoring strategy. Indeed, our analysis of the third default was enhancing different points:

- Many students using our rubrics were never choosing the lowest level, in any skill. By having 3 levels only, our assessment tool precision was decreasing much.
- Many times, students were telling us: "I feel to be between two levels" or, more precisely, "For this aspect of the skill, I feel to be at level 2. But, for this other aspect, at level 3." We

then thought that different evaluation scales for different evaluation criteria was a better scoring strategy.

- We analysed the time spent by students to answer all the questions of a rubric. This was significantly less than what was necessary to read the descriptions of the skill levels. It was obvious that our students were not reading all the descriptions, but perhaps just the title of the activity competence, before to evaluate themselves. Then we were doubtful about the evaluation quality obtained with such a rubric.

On this basis, concerning the scaling strategy, different transformations were operated on our rubric version 1. Many of them were inspired by the idea of dimension as expressed by Rusman and Dirkx (2017) who emphasize that three performance indicators are commonly used: amount, frequency and intensity.

Let us also underline that, at this stage of the rubric design, there can be a few roundtrips between the writing of the skills and the set-up of the evaluation scales. The roundtrips done must increase the coherence and the operational aspect of the skills and of the evaluation scales. Then this stage of design has implications and/or depends on a specific skill but also on the whole rubric. The choices to make are different if the scaling strategy of the rubric is supporting a more qualitative, or quantitative, evaluation approach. The example below, extracted from our work where we have chosen mainly a writing 2 type, will clarify what we are discussing now:

- Writing 1 of a know-how skill: To know how to use various selection criteria in order to decide if a document, found with a request in a bibliographical database, meets the needs of an information retrieval task;

    - Evaluation scale 1: The rubric can propose the student to evaluate what type of action he is usually doing during this activity, for example by using the revised Bloom's taxonomy (Krathwohl, 2002): to apply, to analyse, to evaluate, to create. We have then set a qualitative evaluation for this criterion.

- Writing 2 of a know-how skill (more operational): To know how to analyse a document, found with a request in a bibliographical database, according to various selection criteria in order to decide if the document meets the needs of an information retrieval task;

    - Evaluation scale 1: The rubric can propose the student to evaluate how much help he needs to perform the analysis described in the skill or how often he is performing such an analysis when adopting or rejecting a document. We have then set a quantitative evaluation for this criterion.

After this phase, aiming at choosing the exact skill for an evaluation criterion and at defining the associated type of scale, we used the considerations that are validated for defining *verbal qualifiers* (Rohrman, 2007; p.11):

> "*(1) appropriate position on the dimension to be measured; (2) low ambiguity*
> *(i.e., low standard deviation in the scaling results); (3) linguistic compatibility*

with the other VSPLs chosen for designing a scale; (4) sufficient familiarity of the expression; (5) reasonable likelihood of utilization when used in substantive research."

Then, by using the quantitative analysis done by the same author, we obtained for our scales:

- Intensity qualifiers for a 6-point scale: not at all, a little, partly, quite, very, extremely;
- Frequency qualifiers for a 6-point scale: never, rarely, sometimes, fairly often, very often, always.

Conclusions and perspectives

We focused this article about concrete facets of the rubrics' design. To summarize our experience, we can recommend a design of such assessment tools that follows the steps presented in the Table 1.

Table 1: Steps for the design of a rubric

Design steps	Work to be done	Expected results
Task writing	To write down the task, with its different usual chronological activities (steps) and for each of them: their finalities, main learning concepts and main computer and methodological tools to be used when performing the activity, tricks and traps.	A comprehensive overall view of the whole task. A comprehensive view of the diverse activities composing the whole task. A good understanding of the resources needed to perform the task (concepts, tools, etc.).
Skill listing	To define the main competence associated to each activity and to create the list of the skills that are "the constituents" of the competence. To do a listing, for each skill, of its sub-skills by using the two types: knowledge and know-how.	All the aspects of the task transformed in a list of skills and sub-skills that are mobilized when somebody performs the task. All the information elements that make it possible to set the evaluation scales and the skill levels.
Scaling strategy	To choose between a more or less qualitative or quantitative scaling strategy and more or less operational skills.	All the evaluation information transformed in the various necessary evaluation scales and skill levels.

As a conclusion, it seems to be that rubrics were mostly used with their basic matrix shape where a single skill evaluation scale is used and based only on a few accomplishment levels. The simplicity of a few rubric principles perhaps damaged the educational potential of rubrics, deeply related to assessment. However, it is possible to take advantage of different rules to make the rubrics less simple but more efficient. We have found a few of them but we were surprised about the fact that the literature about rubrics seems to be poor in the rules about skill description. More information can be found about the scaling strategy. This is an axis for future research about rubrics because the skill description and the scaling strategy are

deeply linked, as we showed in this article. Other authors concluded: "a literature that is beyond its infancy but not yet mature." (Brookhart & Chen, 2015; p.362).

References

1. Anderson, J. S., & Mohrweis, L. C. (2008). Using Rubrics to Assess Accounting Students' Writing, Oral Presentations, and Ethics Skills. *Journal of Business Education, 1*(2), 85–94.

2. Bandura, A. (2006). Guide for constructing self-efficacy scales. In T. Urdan & F. Pajares (Eds.), *SelfEfficacy Beliefs of Adolescents.* IAP.

3. Berthiaume, D., & Rege Colet, N. (2013). Comment développer une grille d'évaluation des apprentissages? In D. Berthiaume & N. Rege-Colet (Eds.), *La pédagogie de l'enseignement supérieur: repères théoriques et applications pratiques* (pp. 269–283). Peter Lang.

4. Brookhart, S. M., & Chen, F. (2015). The quality and effectiveness of descriptive rubrics. *Educational Review, 67*(3), 343–368. https://doi.org/10.1080/00131911.2014.929565

5. Dawson, P. (2015). Assessment rubrics: towards clearer and more replicable design, research and practice. *Assessment & Evaluation in Higher Education, 42*(3), 347-360. https://doi.org/10.1080/02602938.2015.1111294

6. Greenberg, K. P. (2015). Rubric Use in Formative Assessment: A Detailed Behavioral Rubric Helps Students Improve Their Scientific Writing Skills. *Teaching of Psychology, 42*(3), 211–217. https://doi.org/10.1177/0098628315587618

7. Krathwohl, D., R. (2002). A Revision of Bloom's Taxonomy: An Overview. *Theory into Practice, 41*(4), 212–218.

8. Lowe, M. S., Booth, C., Stone, S., & Tagge, N. (2015). Impacting Information Literacy Learning in First-Year Seminars: A Rubric-Based Evaluation. *Portal-Libraries and the Academy, 15*(3), 489–512. https://doi.org/10.1353/pla.2015.0030

9. Platteaux, H., Sieber, M., & Lambert, M. (2017). A competence framework of the student's job: helping higher education students to develop their ICT skills. *Proceedings of INTED2017*, 3064–3071. Valencia Spain – 6-8 March, 2017. IATED. Retrieved from https://library.iated.org/publications/INTED2017

10. Popham, W. J. (1997). What's Wrong – and What's Right – with Rubrics. *Educational Leadership, 55*(2), 72–75.

11. Rohrmann, B. (2007). *Verbal qualifiers for rating scales: Sociolinguistic considerations and psychometric data.* University of Melbourne. Retrieved from http://www.rohrmannresearch.net/pdfs/rohrmann-vqs-report.pdf

12. Rusman, E., & Dirkx, K. (2017). Developing Rubrics to Assess Complex (Generic) Skills in the Classroom: How to Distinguish Skills' Mastery Levels? *Practical Assessment, Research & Evaluation, 22*(12). Retrieved from http://pareonline.net/getvn.asp?v=22&n=12

13. Simon, M., & Forgette-Giroux, R. (2001). A rubric for scoring postsecondary academic skills. *Practical Assessment, Research & Evaluation, 7*(18), 103–121.