

**EMBRACING RUBRICS FOR EVALUATING STUDENT
LEARNING OUTCOME ASSESSMENT: A STUDY CASE**

**Danjela BRAHO
Lorena MARGO ZEKO
Katerina SOTIR**

Department of Mathematics and Physics
Faculty of Natural Science and Human Science
“F. S. Noli” University

Abstract

Assessment is a very powerful tool in schools but the problem is how it can be used to extract as much detailed information about student work as possible. For this reason, in many countries of the world, the idea of using rubrics in assessment has been put forward, where the rubrics are schemes with several skills and levels, where for each skill and level there is a certain criterion to evaluate a student. In many countries of the world, this assessment method is very efficient and the question that arises is how effective it will be in the conditions of Albanian education.

Our research is focused on a concrete study of how the use of rubrics can work in Albanian schools and the differences that appear between the assessment used for years in our schools and the assessment using rubrics, with the aim of drawing a conclusion as to which this type of assessment is more suitable to be used in the conditions of Albanian education. Results revealed in our case that the evaluation with rubrics is better than the traditional method.

Key words: Rubrics evaluation, math education

Abstrakt

Vlerësimi është një mjet shumë i fuqishëm në shkollë, por problemi qëndron te mënyra se si mund të përdoret për të nxjerrë sa më shumë informacione të detajuara rreth punës së nxënësve. Për këtë arsye, në shumë vende të botës, është propozuar përdorimi i rubrikave në vlerësim, ku rubrikat janë skema me disa aftësi dhe nivele, ku për secilën aftësi dhe nivel ekziston një kriter i caktuar për të vlerësuar një nxënës. Në shumë vende të botës, kjo metodë vlerësimi është shumë efektive dhe pyetja që lind është se sa efektive do të jetë në kushtet e arsimit shqiptar.

Hulumtimi ynë është përqendruar në një studim konkret se si përdorimi i rubrikave mund të funksionojë në shkollat shqiptare dhe dallimet që shfaqen midis vlerësimit

të përdorur prej vitesh në shkollat tona dhe vlerësimit duke përdorur rubrika, me qëllim të nxjerrjes së një përfundimi se cili lloj vlerësimi është më i përshtatshëm për t'u përdorur në kushtet e arsimit shqiptar. Rezultatet në rastin tonë zbuluan se vlerësimi me rubrika është më i mirë se metoda tradicionale.

Fjalë kyçe: *Vlerësimi me rubrika, arsimi i matematikës*

Introduction

Albanian schools, same as those all over the world, are all the time looking for paths of continuous improvement, in order to ensure always a greater quality as well as to respond to the rapid technological development of nowadays, to prepare skilled professionals to adapt to market demands. The key to the success of any student in a given field is his foundation from previous education, and it is precisely for this reason that steps must be taken to improve education. One aspect where changes can be made is that of evaluation, as long as evaluation is that powerful tool that provides information both about the work of teachers and students. For this reason a large number of schools around the world have turned their attention to the use of rubrics in assessment as a way to bring innovation to education and a significant improvement in it.

This study is of great interest for a number of important reasons. First of all, there are several studies that show the benefits of using rubrics in assessment and the positive results that they give after their use in schools, which according to them is worth replacing the type of assessment that has been used for years. But, no change can be realized in the blink of an eye, without first doing tests and studies which will show that such a change is really worth it.

Secondly, we live in an era where changes are fast and diverse and thus we all have to adapt to these changes by trying to keep up with the times. Students are the future of the country and for this reason, special care must be taken with the policies that are applied in order to be able to adapt to the ever more rapid development.

Thirdly, European and wider schools with a very high quality of education have started applying these rubrics and studies have shown that the policies they have undertaken in most cases have given positive results and therefore there is a need to see that how these sections will work in our country.

Fourthly, the way of constructing rubrics with the help of specific criteria at first sight show that such kind of assessment is possible to provide detailed information about the individual work of each student. On the other hand, it

is impossible to come to a concrete conclusion without doing the necessary studies that will convince us of the effectiveness of using rubrics in the assessment.

Because as Jandris, 2001 said "The heart of assessment is a continuous process, in which the teacher, in collaboration with the student, uses the information to direct the next step of learning.". Educators use assessment results to help improve teaching and learning and to evaluate instructional practices, programs, and schools. Assessments are also used to measure student achievement and evaluate individual students' mastery of skills so that students, parents, administrators, and teachers know how students are progressing toward agreed-upon standards. Assessments are used to rank students, classrooms, and schools, to generate comparative information, and to create accountability data. They are a major-force in shaping public perception about the quality of our schools. Assessments generate the data on which instructional and policy decisions are made.

An abundant literature shows that the rubric is a suitable tool to evaluate competences (Cebrián, 2012; Pérez and Zambrano, 2012; Vidales and Recalde, 2012). A good assessment is part of the learning process, not just a means of documenting or judging that learning (Moore 2010). As Moore says "assessment is not something to be despised or ignored as irrelevant or impossible but embraced as an essential aspect of learning and the best way to truly understand and improve that learning" (Moore, 2010, p.1).

According to Panadero & Jonsson, (2013) rubrics have been designed and researched for their efficacy in promoting thinking and learning as well as making the assessment criteria required transparent to students. Instructional rubrics have also been found to be useful for teachers. Andrade (2000) for example advocates their use as they help teachers teach, make assessing student work quick and efficient, and help teachers justify to parents and others the grades that they assign to students. It is important to mention the concrete studies carried out in relation to the rubrics regarding the topic of mathematical reasoning presented in the article "Developing a Rubric for Assessing Mathematical Reasoning: A Design-Based Research Study in Primary Classrooms" by the authors (Loong et al, 2018). These authors describe a concrete study carried out, where they created a rubric for the assessment of mathematical reasoning and at the end of all this what is important to mention is that the result they got was positive and they recommended the use of rubrics in everyday school life. This article was an impetus for as to a concrete study of creating a rubric and to compare traditional evaluation with rubric evaluation in Albanian schools.

Methodology

Given that the purpose of this entire paper is to test how effective student assessment is by using rubrics, then the next step is to find a rubric including abilities and levels, and for each of them to formulate specific criteria. In this case we gone use a rubric created by K. Sotir as part of her Master Thesis. In this way, the half of the completed tests will be evaluated considering the completed rubric.

Naturally, the structure of the rubric will match the test questions as well as the knowledge of the chapter to be tested. In this way, this rubric contains three main skills and 5 levels, where for each level and skill there is a specific criterion that includes all possible performances of a student from poor to excellent level. The rubric that was used in the evaluation of half of the performed tests is presented in table 1.

Table 1: Rubrics for evaluation of subject “Ratio and Proportion”

Levels Skills	Poor achieveme nts	Achievem ents that need improvem ent	Average achieveme nt	Satisfying achieveme nts	Excellent achieveme nts
Analysis of problem containi ng ratios and proporti ons	Does not analyze any of the concepts applied in exercises such as ratio, equal ratios, proportion s, direct proportion s and the relation between ratio and proportion.	Makes incomplete and inaccurate analyzes of the concepts applied in the exercises, such as ratio, equal ratios, proportion, direct proportion as well as the	Makes incomplete but accurate analyzes of the concepts applied in the exercises, such as ratio, equal ratios, proportion, direct proportion and the relationshi	Make complete or partial but accurate and logical analyses, of the concepts applied in the exercises, such as ratio, equal ratios, proportion, direct proportion	Analyzes in a correct and original way the concept of ratio, equal ratios, proportion and direct proportion as well as the relationshi p that exists between ratio and

		relationship that exists between ratio and proportion.	proportion that exists between ratio and proportion.	relationship that exists between ratio and proportion.	proportion.
Solving the problem	Does not solve the problem and cannot make the connection with the previous knowledge applied in the exercises, such as multiplication, division, simplification, properties and perimeter of the rectangle, properties of the triangle, properties of the ratio in the proportion	Solves the problem incompletely and inaccurately and makes incomplete connections with previous knowledge applied to the problem, such as multiplication, division, simplification, perimeter of a rectangle, properties of a triangle, properties of ratio in proportion	Solves the problem incompletely but correctly, as well as makes the connection with the previous knowledge applied in the exercise, such as multiplication, division, simplification, perimeter of the rectangle, properties of the triangle, properties of the ratio in the	Solves the problem completely but makes non-essential mistakes, as well as makes the connection with the previous knowledge applied in the exercise such as multiplication, division, simplification, perimeter of the rectangle, properties of the triangle, properties of ratio to	Solves the problem in an accurate and original way and makes the connection with the previous knowledge applied in the exercises, such as simplification, the perimeter of the rectangle, the properties of the triangle, multiplication, division, properties of the ratio in the

	exercises.	exercises.	proportion.	proportion.	proportion.
Symbolics and logical arguments	Does not give mathematical arguments or does not use appropriate mathematical symbols.	Gives incomplete and incorrect mathematical arguments or does not use appropriate mathematical symbols.	Gives incomplete but correct mathematical arguments and uses mathematical symbolism in part or in full.	Gives complete mathematical arguments and uses mathematical symbols.	Gives correct and original mathematical arguments and uses appropriate mathematical symbols.

It is necessary to point out the type of rubric created, where precisely because of its very structure, it is analytical due to the fact that it also focuses on detail and is used to evaluate the individual work of each student. The criteria used aim to provide details regarding the students' work.

The skills included in this section are:

- Problem analysis containing ratios and proportions.
- Problem solving.
- Symbolism and logical arguments

The participants in this study were the students of the two different schools in Albania, 11 students from the first school and 18 students from the second school. They were contacted purposely and were given the rubric and the test. We created two groups by dividing in half the number of students by each school. So we created one group with 14 students and one with 15 students. Randomly we choose that the tests of the group with 14 students to be evaluated in traditional way and the group with 15 students with rubrics. The test contains 6 exercises related to ratios and proportions. For the evaluation with rubrics we proceeded in this way: Each exercise carries 6 points, where according to the rubric this means that each skill involved, such as problem analysis, problem solving, and logical arguments and symbolism, carries two points each. How many points each student will earn in each exercise for each skill will depend on the level of achievement

they demonstrate. In this way, our rubric contains 5 levels, where in each step from one level to another in an ascending scale, the student earns 0.5 more points. This means that if for exercise 1, the student for problem analysis ability meets the criterion of "weak achievements", out of two available points he will receive 0 points, but if he meets the criterion of "achievements that need improvement", it will gain 0.5 points out of two maximum points, and in turn if it meets the criterion of "average achievements" it will gain 1 point, in terms of "satisfactory achievements" it will gain 1.5 points and if it reaches the level of "excellent achievements" will earn the maximum points available which are 2.

Results

In this section, we provide an overview of the findings from these survey and try to understand by our results which type of evaluation is better in our case.

First we controlled if our data are normally distributed for both groups for this reason we applied Shapiro-Wilk test of normality and the results are presented in tables 2 and 3 (p-value in each case is greater than 0.05 meaning that we have normal distribution for both groups).

Table 2. Tests of Normality^a for group evaluated with traditional manner

	Kolmogorov-Smirnov ^b			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Total	.112	14	.200*	.987	14	.997

*. This is a lower bound of the true significance.

a. Type of evaluation = Traditional

b. Lilliefors Significance Correction

Table 3. Tests of Normality^a for group evaluated with rubric

	Kolmogorov-Smirnov ^b			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Total	.119	15	.200*	.930	15	.276

*. This is a lower bound of the true significance.

- a. Type of evaluation = Rubrics
- b. Lilliefors Significance Correction

In table 4 are presented descriptive statistics for both types of evaluations. In order to see if we have statistical significant difference between two types of evaluations since the data are normally distributed we performed Independent T test, results are presented in table 5.

Table 4. Descriptive Group Statistics

	Type of evaluation	N	Mean	Std. Deviation	Std. Error Mean
Total	Traditional	14	10.429	5.6120	1.4999
	Rubric	15	17.533	9.5421	2.4638

Using Independent Samples T test resulted a $p\text{-value}=0.022<0.05$ so we have statistical significant difference between two types of evaluations, where evaluation with rubrics results to be the best type of evaluation.

Table 5. Independent T test results

		Levene's Test for Equality of Variances		t-test for Equality of Means		
		F	Sig.	t	df	Sig. (2-tailed)
Total	Equal variances assumed	4.838	.037	-2.421	27	.022
	Equal variances not assumed			-2.463	22.911	.022

Also we performed an analysis for each exercise in order to understand in which type of exercises we have significant difference between two types of evaluations. Mann Whitney U test is used in this case due to the distribution of the variables (Table 6).

Table 6. Mean Ranks for each exercise

	Type of evaluation	N	Mean Rank	Sum of Ranks
Exercise 1	Traditional	14	13.96	195.50
	Rubrics	15	15.97	239.50
	Total	29		
Exercise 2	Traditional	14	12.64	177.00
	Rubrics	15	17.20	258.00
	Total	29		
Exercise 3	Traditional	14	12.54	175.50
	Rubrics	15	17.30	259.50
	Total	29		
Exercise 4	Traditional	14	14.18	198.50
	Rubrics	15	15.77	236.50
	Total	29		
Exercise 5	Traditional	14	11.79	165.00
	Rubrics	15	18.00	270.00
	Total	29		
Exercise 6	Traditional	14	9.43	132.00
	Rubrics	15	20.20	303.00
	Total	29		

Table 7 Test Statistics^a

	Exercise 1	Exercise 2	Exercise 3	Exercise 4	Exercise 5	Exercise 6
Mann-Whitney U	90.500	72.000	70.500	93.500	60.000	27.000
Wilcoxon W	195.500	177.000	175.500	198.500	165.000	132.000
Z	-.641	-1.453	-1.536	-.507	-1.988	-3.736
Asymp. Sig. (2-tailed)	.522	.146	.124	.612	.047	.000

a. Grouping Variable: Type of evaluation

b. Not corrected for ties.

As seen in table 7 we have statistical significant difference between types of evaluation in exercises 5 and 6 ($p\text{-value}=0.047<0.05$ in Exercise 5 and $p\text{-value}=0<0.05$ in Exercise 6).

Conclusions and discussions

As stated before in previous section we have statistical significant difference between evaluation with rubric and the traditional method. Results of Independent T test applied in this case, since our data is normally distributed, affirm that evaluating with rubrics is more effective. Moreover findings revealed that we have statistical significant difference in exercises 5 and 6 based on type of evaluation.

In general rubrics are seen as valuable tools for evaluating student work in mathematics. They provide clear criteria and levels of achievement for assessing various aspects of mathematical performance. They offer several positive benefits when used effectively in educational settings. Some of these advantages are:

Rubrics explicitly communicate what instructors value and what they'll assess. Students who understand these expectations are motivated to meet the identified standards.

Rubrics ensure consistent and objective assessment. Every assignment is evaluated using the same criteria, leading to more accurate and unbiased scoring.

Finding the right language to express performance expectations requires careful thought and often need revisions before becoming practical and user-friendly.

We are also aware that rubrics have as well their disadvantages such as: While they provide structure, they may not account for unique qualities or strengths that individual students bring to a task.

In summary, while rubrics offer valuable benefits, educators should be mindful of their limitations and adapt them appropriately to enhance student learning and assessment processes and available resources.

References

- Andrade, H. G. (2000). Using Rubrics to Promote Thinking and Learning. *Educational Leadership*, 57(5), 13-18.
- Cebrián, M. (2012). Desarrollo conjunto de herramientas eRúbricas federadas para la evaluación por competencias. II Congreso Internacional sobre Evaluación por competencias mediante eRúbricas.
- Jandris, T.P. (2001). *Essentials for principals: Data-based decision-making*. Alexandria, VA: National Association of Elementary School Principals and Arlington, VA: Educational Research Service.
- Loong, E., Vale, C., Widjaja, W., Herbert, S., Bragg, L., Davidson, A., (2018) "Developing a Rubric for Assessing Mathematical Reasoning: A Design-Based Research Study in Primary Classrooms"
- Moore, B. (2010). Using Assessment to Improve Instruction. Online Video Modules. Retrieved from <http://pandora.cii.wvu.edu/cii/resources/modules/assessment/>
- Panadero, E. & Jonsson, A. (2013) The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, 9, 129–144
- Pérez, L., & Zambrano, D. (2012). Aprendizaje Basado En Competencias Utilizando Una Rúbrica Analítica. II Congreso Internacional sobre Evaluación por Competencias mediante e rubricas. Málaga, octubre 2012.
- Toalongo, X., Trelles, C., Alsina, Á. "Design, Construction and Validation of a Rubric to Evaluate Mathematical Modelling in School Education". *Mathematics* 2022, 10, 4662. <https://doi.org/10.3390/math10244662> .
- Vidales, K., & Recalde, I. (2012). El seguimiento y la evaluación de las Competencias Transversales a través de las rúbricas. II Congreso Internacional sobre Evaluación por Competencias mediante e rubricas. Málaga, octubre, 2012.