

# Teachers' Perception for Evaluation of the Primary School Students With Final Grade

Bashkim Rakaj

International University of Novi Pazar, Novi Pazar, Serbia

The most difficult and delicate part of the learning process is assessment. Assessment is difficult because during his/her accomplishment the teacher has to consider not only the acquisition of knowledge from students, but also the overall degree of development of their competencies. It is delicate, because through assessment we can influence the emotional side of students and their "willingness" to learn. Despite these facts, teachers need to evaluate students during the learning process. The purpose of the research is to find out what is the perception of teachers about the level of evaluation of students with final grade. The overall results showed that 89% of teachers agree, 9% of them have a neutral attitude, and 2% do not agree that the evaluation of students with a final grade is done taking into account many aspects and using many methods, the overall average of the results,  $M = 4.36$ . Based on the empirical results, it was found that teachers who have completed assessment training have a more positive approach to student assessment, as the average score is higher than teachers who have not completed assessment training. From the qualitative results it was understood that the Teachers did not encounter any difficulties during the assessment of the students.

*Keywords:* teacher, student, final grade, perception, evaluation

## Introduction

Today, to implement the Kosovo Curriculum framework, teachers are required to focus on developing the competencies of all students. To achieve this, teachers should be careful to include in the daily plan different ways of assessment, which will help to assess not only the knowledge, but also the skills and habits of students.

While knowledge assessment is easier, and can be accomplished by asking questions, engaging to solve the exercise, holding tests, and checking exercises, remains one more difficult task for teachers to find appropriate ways to assess students' skills, for example the ability of students to think critically, to solve exercises, research information, use information technology, as well as different skills performing individual exercises, in pairs and in groups. Assessing these elements of the student will help the teacher to judge the degree of development of students' competencies and orientate his work towards their advancement.

Teachers need to be creative to find different forms of assessment within the classroom lessons. They should use different assessment tools, as well as involve students in the assessment process.

In order to assess students as objectively as possible, teachers should use different assessment instruments. In addition to the standard instruments that are tests and control tasks, teachers are good to practice the evaluation

---

Bashkim Rakaj, PhD candidate, International University of Novi Pazar, Novi Pazar, Serbia; master of Social Policy Analysis, University of Tirana, Tirana, Albania.

Correspondence concerning this article should be addressed to Bashkim Rakaj, Center for social work, Prizren, Kosovo.

of the file prepared by the student, the accomplishment of projects and their presentations, performing homework, etc. (Mula, 2011).

In addition to the standard forms of assessment mentioned above, teachers can also use the following forms: engaging students in creating models (models), researching information on the Internet, identifying and formulating examples from practical life, presenting the solution of those assignments, participation in discussions, engaging students to help friends who have difficulties in learning difficulties, etc.

### **Reviewing Bibliography**

Assessing students with a final grade in nine-year schools is an ongoing process carried out by the teacher. Assessment is a successful process, which takes place during the daily teaching work and with which information is collected from the measurement process (Zylfiu, 2005), is a conscious, comparable activity that shows the results achieved by students against learning objectives and draws conclusions from this comparison (Limanaqi, 2003). Summative assessment is comprehensive and assesses students who learn at the end of a course or program. It is used to determine if students have accomplished the standards or the goal (Adams & Ray, 2016), identify the standard and achieve the objectives at the right moment, normally at the end of another period, e.g. in the end of one or several Chapters, semester or in the end of the school year (Musai, 2003).

Teachers should be able to combine forms and methods of teaching, which would take into account the psychophysical abilities of students, eventual difficulties and generally create a suitable environment for learning. Teachers should foster communication and interpersonal relationships in the classroom, providing a place for students' active participation in building their learning. After a student-teacher interaction, for a semester or for an entire year, it should bring the student's final grade. Others request the work in groups and folders, a small number of teachers do assessment with a grade not paying attention to their final achievements than considering their progress and efforts (Woolfolk, 2011).

Classroom assessments can take a variety of forms, unit tests, portfolio, projects, performances, oral presentations, etc.

Student assessment is an integral part of teaching and learning, through which we understand how much students have acquired knowledge, skills, habits, and abilities (Devetaku & Mehmeti, 2010). Genuine assessments require students to apply acquired skills and habits, as they would in real life.

In order to assess objectively, the role of the teacher, his professional training is important. The active attitude of the teacher in the learning process, attention to all students, his attention in assessment, principle, assessment, his tact in assessment (Kraja, 2002), continuous assessment of students and taking as a basis for action and interaction of students. Assessment should not be an isolated formality, but a process of systematic collection, analysis, and interpretation of information in order to determine to what extent the student has mastered the instructional objectives.

In addition to aiming to assess the student, the teacher also assesses his/her work, commitments, and learning efforts. Therefore, students' grades can be used by teachers as a tool for corrective demands of his work (Zylfiu, 2005).

Ongoing assessment is, in fact, the assessment made by both teachers and students of the work done throughout the course. The grade obtained at the end of the semester or school year is a direct reflection of all student achievement. When we mention the fact, that this assessment is really an assessment made by students as well, to some extent, here, we will also include self-assessment.

The evaluation includes the quantitative and qualitative description of the work done, based on the most objective measurement (Zylfiu, 2005). The main purpose of evaluation is to make decisions that aim to improve learning outcomes and the learning process itself. Decision making requires the continuous collection of information to measure and evaluate all aspects of the process. Assessment practices have a crucial impact on the learning process.

Assessment is one of the most powerful educational tools for promoting learning effectiveness. But, it must be used in the right manner, it must be in the use of teaching, learning, in order to raise the school achievement of students (Black & William, 1998).

A topic like assessment is very important in education. For this reason, many teachers have shown interest in research related to educational evaluation. Over the past 35 years, assessment has been extensively studied around the world, in the use of assessment and improving learning.

Researchers in the field of assessment provide different definitions for assessment:

- Assessment is a process that determines the degree to which learning objectives are achieved (Tyler, 1950);
- Evaluation serves to determine the degree of fulfillment of responsibility (Scriven, 1967);
- A systematic review of the value and merit of an object (Joint Committee on Standards for Educational Evaluation, 1981);
- The act of systematic collection of information, against the nature and quality of educational facilities (Nevo, 1995).

The assessment goes through several stages:

- The first phase determines the purpose and exactness of what will be assessed;
- The second phase defines the methodology of evaluation, how it will be assessed, which techniques and instruments of evaluation will be applied;
- The third phase is the implementation of the assessment;
- The fourth phase is informing with the achieved results;
- The fifth phase is the comparison of results with evaluation objectives.

Measurement in the learning process is also a very important process, although they are not the same as assessment, they have similarities. Measurement is a comparison process through which learning objects are identified and described (Petrela, 2001).

Measurement cannot be possible without a clearly defined and evaluated standard which can be used as a basis for comparison (Mita, 2003).

## **Methodology and Methods**

### **The Purpose of the Research**

The purpose of the research is to find out what is the perception of teachers about the level of evaluation of students with final grade. In the framework of the general goal, comparisons will be made of teachers with different characteristics for the level of assessment they give to students in the ninth-year schools.

### **Research Questions and Hypotheses**

#### **Research questions.**

1. What is the correlation between the age of teachers and the evaluation of students with a final grade?

2. Is there a correlation between teachers' education and students' final grade assessment?
3. What is the correlation between teacher training for student assessment and their final grade assessment?
4. Is there a correlation between teachers' experience and students' evaluation of the final grade?
5. What is the correlation between each other's evaluation and students' self-assessment?

### **Research hypotheses.**

Hypothesis 1: There are statistically significant correlations between teachers of different ages in the assessment of students with final grade.

Hypothesis 2: There is a statistically significant correlation between teachers' education and students' final grade.

Hypothesis 3: Teachers who have carried out students' assessment training have a more positive approach to student assessment.

Hypothesis 4: There is a statistically significant correlation between teachers' experience and students' assessment through experimental work, projects, hand-made works.

Hypothesis 5: Teachers who apply each other's assessment system also practice student self-assessment.

### **Participants**

The research involved 147 teachers, where 140 were involved in quantitative research, while seven teachers were involved in qualitative research. In quantitative research, out of 140 teachers, 52 of them work in the village, while 88 work in the town schools. Out of these, 65 work in grades 1-5 and 75 work in classrooms 6-9. From them, 87 are female and 53 are male. In terms of age, 17 teachers are aged up to 30 years, 39 aged 31-40 years, 36 aged 41-50 years, 39 aged 51-60 years, and nine aged over 60 years. Regarding educational preparation, nine of them have completed Teacher's Training Education, 107 with Faculty, 24 teachers have a Master's Degree in Science. Regarding the experience in education, seven teachers said that they have up to five years of experience, 44 of them have 6-15 years of experience, 35 have 16-25 years of experience, 33 teachers have 26-35 years of experience, and 11 teachers have experience over 35 years. After starting work as teachers, 120 of them have completed training for student assessment, while 20 of them have not completed any training of this nature.

### **Instruments of Research**

To accomplish the research, we applied two measuring instruments. Quantitative data were provided through the questionnaire, while qualitative data are provided through the interview.

The questionnaire is an instrument specially prepared for the needs of the study and consists of two parts. In the first part, the demographic data of teachers are presented. In the second part, 12 questions of the Likert scale, with five scores, are asked, related to the assessment of students with a final grade. After carrying out the research, for the needs of the study, from five scales (1: I do not agree at all, 2: I do not agree, 3: Neutral, 4: I agree, 5: I completely agree), the recoding was done in three scales (1: I do not agree, 2: Neutral, 3: Agree).

The interview is an instrument specially prepared for this research, and it asks five open-ended questions.

Also, to test the hypotheses, the variables were re-coded and transformed into a set of variables, named: assessment of students with a final grade.

**Questionnaire reliability.** In the first part of the results of the reliability analysis, descriptive statistics are given, which consist of: means, variances, and variance-covariance matrices, which show the relationship between variance and questions.

Average meter,  $M = 52.33$  and standard deviation,  $DS = 4.406$ . The total average is 4.36, the average variance is 0.494, the average interval is 0.892, while the variance change interval is 0.510. The total average of the inter-item correlations is 0.226, the minimum correlation is -0.118, while the maximum correlation is 0.727.

In the item-total statistics section, when a question is extracted from the meter it is calculated the mean of the gauge and the variance from the remaining questions (Scale Mean if Item Deleted and Scale Variance if Item Deleted), as well as the correlation between the question, derived from the meter and the total of other gauge questions (Corrected Item-Total Correlation). They have also been reported as Squared Multiple Correlation coefficients and the value of reliability in relation to the remaining questions (Cronbach's Alpha if Item Deleted).

From the results, it is noticed that the question-total correlations are between 0.210 and 0.589.

Cronbach's Alpha total meter reliability coefficient is 0.758, which is high value and indicates that the meter is reliable.

By value obtained (Cronbach's Alpha if Item Deleted), ranking of questions by highest value from smallest to largest is: P12, P9, P2, P3, P8, P4, P1, P6, P7, P5, P11, P10.

Based on the analysis of variance, we can say that the difference between the measurements,  $p = 0.000$ , is statistically significant. Hotelling's  $T^2$  test, which tests the equivalence of question averages is calculated as  $p = 0.000$ . From this result it is understood that there is an important difference statistically at least between the two question averages.

Table 1

*Reliability Results According to Alfa Cronbach's Model*

| Item Statistics                                                                                                  | <i>M</i> | <i>SD</i> | <i>N</i> |
|------------------------------------------------------------------------------------------------------------------|----------|-----------|----------|
| I make the assessment based on the student's commitment for a long period.                                       | 4.53     | 0.593     | 139      |
| When assessing students with a final grade, I take into account their commitments.                               | 4.65     | 0.600     | 139      |
| To evaluate students I use several evaluation methods.                                                           | 4.55     | 0.650     | 139      |
| To evaluate students, I take into account their interaction, homework engagements, tests, and other methods.     | 4.62     | 0.607     | 139      |
| Students who are active, give ideas, are creative in solving tasks, I rate them with the highest grade.          | 4.60     | 0.560     | 139      |
| Assessing students I use objective tests (e.g. questions with alternative answers, YES/NO, short answers, etc.). | 4.35     | 0.750     | 139      |
| Students are also assessed on the basis of experimental work, projects, handmade works.                          | 4.22     | 0.759     | 139      |
| For the assessment of students I take as a basis their systematic monitoring (regular observation).              | 4.40     | 0.738     | 139      |
| For the assessment of students, I take into account the oral answers and written answers (drafts, exams, etc.).  | 4.50     | 0.695     | 139      |
| Students assess each other.                                                                                      | 3.76     | 0.908     | 139      |
| Students do self-assessment.                                                                                     | 3.82     | 0.870     | 139      |
| Students are satisfied with my final grade.                                                                      | 4.34     | 0.597     | 139      |

Table 2

*Inter-Item Correlation Matrix*

|    | P1    | P2    | P3    | P4    | P5    | P6    | P7    | P8    | P9    | P10    | P11    | P12   |
|----|-------|-------|-------|-------|-------|-------|-------|-------|-------|--------|--------|-------|
| P1 | 1.000 | 0.571 | 0.507 | 0.286 | 0.268 | 0.210 | 0.145 | 0.243 | 0.128 | -0.026 | 0.089  | 0.184 |
| P2 | 0.571 | 1.000 | 0.727 | 0.504 | 0.250 | 0.246 | 0.120 | 0.350 | 0.283 | -0.080 | -0.067 | 0.295 |
| P3 | 0.507 | 0.727 | 1.000 | 0.594 | 0.308 | 0.132 | 0.064 | 0.189 | 0.301 | -0.063 | -0.002 | 0.336 |
| P4 | 0.286 | 0.504 | 0.594 | 1.000 | 0.385 | 0.281 | 0.007 | 0.194 | 0.332 | -0.118 | 0.020  | 0.359 |

Table 2 to be continued

|     |        |        |        |        |        |       |       |       |       |        |        |       |
|-----|--------|--------|--------|--------|--------|-------|-------|-------|-------|--------|--------|-------|
| P5  | 0.268  | 0.250  | 0.308  | 0.385  | 1.000  | 0.248 | 0.134 | 0.224 | 0.322 | -0.106 | -0.088 | 0.252 |
| P6  | 0.210  | 0.246  | 0.132  | 0.281  | 0.248  | 1.000 | 0.311 | 0.270 | 0.162 | 0.117  | 0.042  | 0.331 |
| P7  | 0.145  | 0.120  | 0.064  | 0.007  | 0.134  | 0.311 | 1.000 | 0.402 | 0.290 | 0.235  | 0.180  | 0.254 |
| P8  | 0.243  | 0.350  | 0.189  | 0.194  | 0.224  | 0.270 | 0.402 | 1.000 | 0.405 | 0.027  | 0.078  | 0.319 |
| P9  | 0.128  | 0.283  | 0.301  | 0.332  | 0.322  | 0.162 | 0.290 | 0.405 | 1.000 | 0.148  | 0.113  | 0.518 |
| P10 | -0.026 | -0.080 | -0.063 | -0.118 | -0.106 | 0.117 | 0.235 | 0.027 | 0.148 | 1.000  | 0.650  | 0.248 |
| P11 | 0.089  | -0.067 | -0.002 | 0.020  | -0.088 | 0.042 | 0.180 | 0.078 | 0.113 | 0.650  | 1.000  | 0.286 |
| P12 | 0.184  | 0.295  | 0.336  | 0.359  | 0.252  | 0.331 | 0.254 | 0.319 | 0.518 | 0.248  | 0.286  | 1.000 |

Table 3

*Inter-Item Covariance Matrix*

|     | P1     | P2     | P3     | P4     | P5     | P6    | P7    | P8    | P9    | P10    | P11    | P12   |
|-----|--------|--------|--------|--------|--------|-------|-------|-------|-------|--------|--------|-------|
| P1  | 0.352  | 0.204  | 0.196  | 0.103  | 0.089  | 0.094 | 0.065 | 0.107 | 0.053 | -0.014 | 0.046  | 0.065 |
| P2  | 0.204  | 0.360  | 0.284  | 0.183  | 0.084  | 0.111 | 0.055 | 0.155 | 0.118 | -0.043 | -0.035 | 0.106 |
| P3  | 0.196  | 0.284  | 0.423  | 0.234  | 0.112  | 0.064 | 0.032 | 0.091 | 0.136 | -0.037 | -0.001 | 0.130 |
| P4  | 0.103  | 0.183  | 0.234  | 0.368  | 0.131  | 0.128 | 0.003 | 0.087 | 0.140 | -0.065 | 0.011  | 0.130 |
| P5  | 0.089  | 0.084  | 0.112  | 0.131  | 0.313  | 0.104 | 0.057 | 0.092 | 0.125 | -0.054 | -0.043 | 0.084 |
| P6  | 0.094  | 0.111  | 0.064  | 0.128  | 0.104  | 0.563 | 0.177 | 0.149 | 0.085 | 0.080  | 0.028  | 0.148 |
| P7  | 0.065  | 0.055  | 0.032  | 0.003  | 0.057  | 0.177 | 0.576 | 0.226 | 0.153 | 0.162  | 0.119  | 0.115 |
| P8  | 0.107  | 0.155  | 0.091  | 0.087  | 0.092  | 0.149 | 0.226 | 0.545 | 0.208 | 0.018  | 0.050  | 0.141 |
| P9  | 0.053  | 0.118  | 0.136  | 0.140  | 0.125  | 0.085 | 0.153 | 0.208 | 0.484 | 0.093  | 0.068  | 0.215 |
| P10 | -0.014 | -0.043 | -0.037 | -0.065 | -0.054 | 0.080 | 0.162 | 0.018 | 0.093 | 0.824  | 0.514  | 0.134 |
| P11 | 0.046  | -0.035 | -0.001 | 0.011  | -0.043 | 0.028 | 0.119 | 0.050 | 0.068 | 0.514  | 0.757  | 0.148 |
| P12 | 0.065  | 0.106  | 0.130  | 0.130  | 0.084  | 0.148 | 0.115 | 0.141 | 0.215 | 0.134  | 0.148  | 0.356 |

Table 4

*Case Processing Summary*

|       |                       | N   | %     |
|-------|-----------------------|-----|-------|
| Cases | Valid                 | 139 | 99.3  |
|       | Excluded <sup>a</sup> | 1   | 0.7   |
|       | Total                 | 140 | 100.0 |

a. Listwise deletion based on all variables in the procedure.

Table 5

*Scale Statistics*

| Mean  | Variance | Std. deviation | N of items |
|-------|----------|----------------|------------|
| 52.33 | 19.411   | 4.406          | 12         |

Table 6

*Summary Item Statistics*

|                         | Mean  | Minimum | Maximum | Range | Maximum/minimum | Variance | N of items |
|-------------------------|-------|---------|---------|-------|-----------------|----------|------------|
| Item means              | 4.361 | 3.755   | 4.647   | 0.892 | 1.238           | 0.089    | 12         |
| Item variances          | 0.494 | 0.313   | 0.824   | 0.510 | 2.629           | 0.028    | 12         |
| Inter-item covariances  | 0.102 | -0.065  | 0.514   | 0.579 | -7.908          | 0.008    | 12         |
| Inter-item correlations | 0.226 | -0.118  | 0.727   | 0.845 | -6.160          | 0.033    | 12         |

Table 7

*Item-Total Statistics*

|     | Scale mean if item deleted | Scale variance if item deleted | Corrected item-total correlation | Squared multiple correlation | Cronbach's alpha if item deleted |
|-----|----------------------------|--------------------------------|----------------------------------|------------------------------|----------------------------------|
| P1  | 47.80                      | 17.046                         | 0.411                            | 0.407                        | 0.741                            |
| P2  | 47.68                      | 16.609                         | 0.499                            | 0.642                        | 0.732                            |
| P2  | 47.78                      | 16.508                         | 0.470                            | 0.641                        | 0.734                            |
| P4  | 47.71                      | 16.873                         | 0.435                            | 0.490                        | 0.738                            |
| P5  | 47.73                      | 17.533                         | 0.334                            | 0.262                        | 0.748                            |
| P6  | 47.98                      | 16.514                         | 0.383                            | 0.283                        | 0.743                            |
| P7  | 48.12                      | 16.508                         | 0.377                            | 0.280                        | 0.744                            |
| P8  | 47.94                      | 16.220                         | 0.445                            | 0.345                        | 0.735                            |
| P9  | 47.83                      | 16.139                         | 0.499                            | 0.406                        | 0.729                            |
| P10 | 48.58                      | 17.014                         | 0.210                            | 0.493                        | 0.770                            |
| P11 | 48.51                      | 16.846                         | 0.253                            | 0.492                        | 0.763                            |
| P12 | 47.99                      | 16.225                         | 0.589                            | 0.422                        | 0.723                            |

Table 8

*ANOVA*

| ANOVA             |               |                |      |             |          |       |
|-------------------|---------------|----------------|------|-------------|----------|-------|
|                   |               | Sum of squares | df   | Mean square | <i>F</i> | Sig.  |
| Between people    |               | 223.231        | 138  | 1.618       |          |       |
| Within people     | Between items | 135.494        | 11   | 12.318      | 31.478   | 0.000 |
|                   | Residual      | 594.006        | 1518 | 0.391       |          |       |
|                   | Total         | 729.500        | 1529 | 0.477       |          |       |
| Total             |               | 952.731        | 1667 | 0.572       |          |       |
| Grand mean = 4.36 |               |                |      |             |          |       |

Table 9

*Hotelling's T-Squared Test*

| Hotelling's T-Squared | F      | df1 | df2 | Sig   |
|-----------------------|--------|-----|-----|-------|
| 119.709               | 10.094 | 11  | 128 | 0.000 |

Table 10

*Reliability Statistics*

| Cronbach's Alpha | Cronbach's alpha based on standardized items | N of items |
|------------------|----------------------------------------------|------------|
| 0.758            | 0.778                                        | 12         |

Based on the reliability score gained from Cronbach's Alpha model, it can be concluded that the reliability of the instrument is high.

**Data Collection Procedure**

The application of the research instruments took place during November 2020. The questionnaires and interviews were distributed to the teachers, but before starting to complete the questionnaires and interviews, the teachers were informed and received the necessary instructions on how to complete these instruments. Applying

field research also has its difficulties, as in many cases, participants do not want to cooperate, they are not honest in their answers, they complete the instruments just to complete them, etc. In order to overcome this problem, we tried to ensure that the participants were instructed in writing (in instruments), but also orally on how to complete, to be asked orally to give the most honest answers and to be careful in giving answers. The questionnaires were completed at some time after the lesson in the presence of the problem researcher. If during the completion of the questionnaires, the participants needed any clarification, the problem researcher explained it to each of them.

### **Methods of Statistical Data Analysis**

Quantitative research data were analyzed through the statistical package for social sciences (SPSS), while qualitative research data were presented in a short form of participants' attitudes. To test the internal consistency of the instrument, Alpha Cronbach was used, taking the value above 0.7 as a value that proves whether or not the questionnaire has internal consistency. The level of statistical significance 0.05 was set for the evaluation of the results of all statistical tests.

The ANOVA One-Way Test was used to compare the averages of teachers' age, schooling, and experience, and student assessment. *T*-test, is used to test the interconnectedness of training for assessment and assessment of students with final grade. To estimate the correlation between each other's evaluation and self-evaluation, Pearson Correlation Coefficient was used.

Correlation coefficient:

$r = 0.00-0.25$ : Very weak connection;

$r = 0.26-0.49$ : Weak connection;

$r = 0.50-0.69$ : Medium connection;

$r = 0.70-0.89$ : High connection;

$r = 0.89-1.00$ : Very high.

## **Research Results**

### **Descriptive Results**

Based on the empirical results, it was found that about 95% of teachers claim that they assess students based on their commitment over a long period of time, while 5% of teachers have a neutral attitude towards this statement.

Around 95% of teachers state that when assessing students with a final grade, they take their commitments as a basis, while 5% of teachers have a neutral attitude.

Around 91.4% of teachers claim to use some assessment methods, while 8% of teachers have a neutral attitude.

Around 95% of teachers said that in order to assess students, they take into account their interaction, homework assignments, tests, and other methods, 4.3% of teachers have a neutral attitude while 0.7% of them do not agree with this statement.

Around 96.4% of teachers claim that students who are active, give ideas, are creative in solving exercises, I rate them with the highest grade, while 3.6% of teachers have a neutral attitude.

Around 90.8% of teachers state that they also use objective tests for students' assessment (e.g. questions with alternative answers, YES/NO, short answers, etc.), 5.6% of teachers have a neutral attitude, while 3.6% of them do not agree with this statement.

Around 87.2% of teachers state that students are assessed on the basis of experimental work, projects, workshops, 11.4% of teachers have a neutral attitude towards this statement, while 1.4% of them do not agree with this statement.

Around 87.9% of teachers state that, for the assessment of students I take into account their systematic monitoring (regular observation), 10.4% of them have a neutral attitude, while 1.4% of them do not agree with this statement.

Around 89.3% of teachers state that, for the assessment of students, they also take oral answers and written answers (drafts, exams, etc.), 10% of them have a neutral attitude, while 0.7% of them do not agree with this statement.

Around 68.6% of teachers state that students assess each other, 20% of them have a neutral attitude, while 11.4% of teachers do not agree with this statement.

Around 73.6% of teachers claim that students do self-assessment, 17.9% of them have a neutral attitude, while 8.5% of teachers do not agree with this statement.

Around 95% of teachers agree with the statement that students are satisfied with my evaluation with a final grade, 4.3% of teachers have a neutral attitude, while 0.7% of them do not agree with this statement.

The overall results showed that 89% of teachers agree, 9% of them have a neutral attitude, and 2% do not agree that the assessment of students with a final grade is done taking into account many aspects and using many methods, the overall average of the results, is  $M = 4.36$ .

Table 11

*Descriptive Results*

| Item statistics                                                                                                      | Agree | Neutral | Do not agree | M    | SD    |
|----------------------------------------------------------------------------------------------------------------------|-------|---------|--------------|------|-------|
| I make the evaluation based on the student's commitment for a long period.                                           |       | 5%      | 0%           | 4.53 | 0.593 |
| When assessing students with a final grade, I take into account their commitments.                                   | 95%   | 4.3%    | 0.7%         | 4.65 | 0.600 |
| To evaluate students I use several assessment methods.                                                               | 91.4% | 8.6%    | 0%           | 4.55 | 0.650 |
| To evaluate students, I take into account their interaction, homework engagements, tests, and other methods.         | 95%   | 4.3%    | 0.7%         | 4.62 | 0.607 |
| Students who are active, give ideas, are creative in solving tasks, I rate them with the highest grade.              | 96.4% | 3.6%    | 0%           | 4.60 | 0.560 |
| For assessing students I use objective tests (e.g. questions with alternative answers, YES/NO, short answers, etc.). | 90.8% | 5.6%    | 3.6%         | 4.35 | 0.750 |
| Students are also assessed on the basis of experimental work, projects, handmade works.                              | 87.2% | 11.4%   | 1.4%         | 4.22 | 0.759 |
| For the assessment of students I take as a basis their systematic monitoring (regular observation).                  | 87.9% | 10.7%   | 1.4%         | 4.40 | 0.738 |
| For the assessment of students, I take into account the oral answers and written answers (drafts, exams, etc.).      | 89.3% | 10%     | 0.7%         | 4.50 | 0.695 |
| Students assess each other.                                                                                          | 68.6% | 20%     | 11.4%        | 3.76 | 0.908 |
| Students do self-assessment.                                                                                         | 73.6% | 17.9%   | 8.5%         | 3.82 | 0.870 |
| Students are satisfied with my final grade.                                                                          | 95%   | 4.3%    | 0.7%         | 4.34 | 0.597 |
| Overall results                                                                                                      | 89%   | 9%      | 2%           | 4.36 | /     |

**Correlation Results**

Based on the correlation results, high and statistically significant correlation exists between the second question and the third question,  $r = 0.727^{**}$ ,  $p = 0.000 < 0.05$ , the mean correlation exists between the second

question and the third question four,  $r = 0.505^{**}$ ,  $p = 0.000 < 0.05$ , the third question and the fourth question,  $r = 0.596^{**}$ ,  $p = 0.000 < 0.05$ , the tenth question and the eleventh question,  $r = 0.647^{**}$ ,  $p = 0.000 < 0.05$ . Correlations between other variables, although in most cases statistically significant, the correlation is weak.

Table 12

*Correlation Results*

| Correlations |                     | P2      | P3      | P4      | P5      | P6      | P7      | P8      | P9      | P10     | P11     | P12     |
|--------------|---------------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| P1           | Pearson Correlation | 0.565** | 0.500** | 0.281** | 0.262** | 0.212*  | 0.138   | 0.237** | 0.139   | -0.020  | 0.087   | 0.175*  |
|              | Sig. (2-tailed)     | 0.000   | 0.000   | 0.001   | 0.002   | 0.012   | 0.105   | 0.005   | 0.102   | 0.814   | 0.307   | 0.038   |
| P2           | Pearson Correlation | 1       | 0.727** | 0.505** | 0.253** | 0.243** | 0.124   | 0.352** | 0.269** | -0.083  | -0.066  | 0.298** |
|              | Sig. (2-tailed)     |         | 0.000   | 0.000   | 0.003   | 0.004   | 0.144   | 0.000   | 0.001   | 0.331   | 0.440   | 0.000   |
| P3           | Pearson Correlation | 0.727** | 1       | 0.596** | 0.310** | 0.129   | 0.069   | 0.192*  | 0.285** | -0.067  | -0.001  | 0.339** |
|              | Sig. (2-tailed)     | 0.000   |         | 0.000   | 0.000   | 0.129   | 0.418   | 0.023   | 0.001   | 0.430   | 0.991   | 0.000   |
| P4           | Pearson Correlation | 0.505** | 0.596** | 1       | 0.387** | 0.279** | 0.012   | 0.197*  | 0.316** | -0.121  | 0.021   | 0.362** |
|              | Sig. (2-tailed)     | 0.000   | 0.000   |         | 0.000   | 0.001   | 0.893   | 0.020   | 0.000   | 0.154   | 0.805   | 0.000   |
| P5           | Pearson Correlation | 0.253** | 0.310** | 0.387** | 1       | 0.245** | 0.139   | 0.227** | 0.306** | -0.110  | -0.086  | 0.256** |
|              | Sig. (2-tailed)     | 0.003   | 0.000   | 0.000   |         | 0.004   | 0.102   | 0.007   | 0.000   | 0.195   | 0.310   | 0.002   |
| P6           | Pearson Correlation | 0.243** | 0.129   | 0.279** | 0.245** | 1       | 0.306** | 0.266** | 0.166*  | 0.119   | 0.042   | 0.325** |
|              | Sig. (2-tailed)     | 0.004   | 0.129   | 0.001   | 0.004   |         | 0.000   | 0.001   | 0.049   | 0.160   | 0.626   | 0.000   |
| P7           | Pearson Correlation | 0.124   | 0.069   | 0.012   | 0.139   | 0.306** | 1       | 0.406** | 0.268** | 0.227** | 0.181*  | 0.260** |
|              | Sig. (2-tailed)     | 0.144   | 0.418   | 0.893   | 0.102   | 0.000   |         | 0.000   | 0.001   | 0.007   | 0.033   | 0.002   |
| P8           | Pearson Correlation | 0.352** | 0.192*  | 0.197*  | 0.227** | 0.266** | 0.406** | 1       | 0.385** | 0.022   | 0.079   | 0.324** |
|              | Sig. (2-tailed)     | 0.000   | 0.023   | 0.020   | 0.007   | 0.001   | 0.000   |         | 0.000   | 0.801   | 0.355   | 0.000   |
| P9           | Pearson Correlation | 0.269** | 0.285** | 0.316** | 0.306** | 0.166*  | 0.268** | 0.385** | 1       | 0.158   | 0.108   | 0.491** |
|              | Sig. (2-tailed)     | 0.001   | 0.001   | 0.000   | 0.000   | 0.049   | 0.001   | 0.000   |         | 0.063   | 0.205   | 0.000   |
| P10          | Pearson Correlation | -0.083  | -0.067  | -0.121  | -0.110  | 0.119   | 0.227** | 0.022   | 0.158   | 1       | 0.647** | 0.239** |
|              | Sig. (2-tailed)     | 0.331   | 0.430   | 0.154   | 0.195   | 0.160   | 0.007   | 0.801   | 0.063   |         | 0.000   | 0.004   |
| P11          | Pearson Correlation | -0.066  | -0.001  | 0.021   | -0.086  | 0.042   | 0.181*  | 0.079   | 0.108   | 0.647** | 1       | 0.286** |
|              | Sig. (2-tailed)     | 0.440   | 0.991   | 0.805   | 0.310   | 0.626   | 0.033   | 0.355   | 0.205   | 0.000   |         | 0.001   |
| P12          | Pearson Correlation | 0.298** | 0.339** | 0.362** | 0.256** | 0.325** | 0.260** | 0.324** | 0.491** | 0.239** | 0.286** | 1       |
|              | Sig. (2-tailed)     | 0.000   | 0.000   | 0.000   | 0.002   | 0.000   | 0.002   | 0.000   | 0.000   | 0.004   | 0.001   |         |

\*\* Correlation is significant at the 0.01 level (2-tailed).

\* Correlation is significant at the 0.05 level (2-tailed).

**Results Related to Hypothesis****Hypothesis 1:**

Alternative hypothesis ( $H_a$ ): Younger teachers are more concerned with students' engagements for their final grade assessment.

Hypothesis zero ( $H_0$ ): Younger teachers no longer take into account students' commitments for their final grade assessment.

Hypothesis testing was carried out through one-way ANOVA parametric test. The average scores of teachers of different ages, regarding the consideration of students' commitment to the final grade assessment, are approximate, the value of  $F$  is low, 1.420, while  $p = 0.231 > 0.05$ . From these results we can conclude that the alternative hypothesis is not supported.

Table 13

*Testing the First Hypothesis*

| Descriptives                              |     |       |                |            |                                  |             |
|-------------------------------------------|-----|-------|----------------|------------|----------------------------------|-------------|
| Assessment of students with a final grade |     |       |                |            |                                  |             |
|                                           | N   | Mean  | Std. deviation | Std. error | 95% confidence interval for mean |             |
|                                           |     |       |                |            | Lower bound                      | Upper bound |
| Up to 30                                  | 17  | 46.29 | 3.619          | 0.878      | 44.43                            | 48.15       |
| 31-40                                     | 39  | 48.74 | 4.031          | 0.645      | 47.44                            | 50.05       |
| 41-50                                     | 36  | 48.31 | 3.520          | 0.587      | 47.11                            | 49.50       |
| 51-60                                     | 39  | 48.00 | 3.656          | 0.585      | 46.81                            | 49.19       |
| Over 60                                   | 9   | 46.67 | 6.892          | 2.297      | 41.37                            | 51.96       |
| Total                                     | 140 | 47.99 | 4.013          | 0.339      | 47.32                            | 48.66       |

Table 14

*ANOVA of the First Hypothesis*

| Assessment of students with a final grade |                |     |             |       |       |
|-------------------------------------------|----------------|-----|-------------|-------|-------|
|                                           | Sum of squares | df  | Mean square | F     | Sig.  |
| Between groups                            | 90.389         | 4   | 22.597      | 1.420 | 0.231 |
| Within groups                             | 2,148.604      | 135 | 15.916      |       |       |
| Total                                     | 2,238.993      | 139 |             |       |       |

## Hypothesis 2:

Alternative hypothesis ( $H_a$ ): There is a statistically significant correlation between teacher's education and final grade student assessment.

Hypothesis zero ( $H_0$ ): There is no statistically significant correlation between teacher's education and students' final grade.

The results showed that teachers who have completed teacher's Training School have the highest average  $M = 49.33$ , then teachers who have completed College,  $M = 48.27$ , while teachers who have completed Master level have the lowest average,  $M = 46.25$ ,  $F = 3.114$ ,  $p = 0.048 < 0.05$ . Based on these results, we can conclude that teachers who have completed Teacher's Training School are better prepared for students' assessment.

Empirical results showed that there is a correlation between teachers' education and final grade students' assessment.

Table 15

*Testing of Second Hypothesis*

| Descriptive                               |     |       |                |            |                                  |             |
|-------------------------------------------|-----|-------|----------------|------------|----------------------------------|-------------|
| Assessment of students with a final grade |     |       |                |            |                                  |             |
|                                           | N   | Mean  | Std. deviation | Std. error | 95% confidence interval for mean |             |
|                                           |     |       |                |            | Lower bound                      | Upper bound |
| Teacher's training school                 | 9   | 49.33 | 2.398          | 0.799      | 47.49                            | 51.18       |
| Faculty                                   | 107 | 48.27 | 3.840          | 0.371      | 47.54                            | 49.01       |
| Master                                    | 24  | 46.25 | 4.812          | 0.982      | 44.22                            | 48.28       |
| Total                                     | 140 | 47.99 | 4.013          | 0.339      | 47.32                            | 48.66       |

Table 16

*ANOVA of Testing of Second Hypothesis*

| Assessment of students with a final grade |                |     |             |          |       |
|-------------------------------------------|----------------|-----|-------------|----------|-------|
|                                           | Sum of Squares | df  | Mean Square | <i>F</i> | Sig.  |
| Between groups                            | 97.353         | 2   | 48.676      | 3.114    | 0.048 |
| Within groups                             | 2,141.640      | 137 | 15.632      |          |       |
| Total                                     | 2,238.993      | 139 |             |          |       |

## Hypothesis 3:

Alternative hypothesis ( $H_a$ ): Teachers who have conducted student's assessment training have more positive approach to students' assessment.

Hypothesis zero ( $H_0$ ): There are no differences between teachers who have completed and those who have not completed training of student assessment.

According to the empirical results it was found that teachers who have completed training for assessment, they have an average of 48.4, while those who have not completed training,  $M = 45.55$ ,  $F = 0.119$ , while  $p < 0.05$ , as in the case of equal assumptions (equal variances assumed), as well as in equal variances not assumed.

The results showed that the alternative hypothesis was endorsed, as teachers who have carried out assessment training have more positive approach to students' assessment.

Table 17

*Testing of the Third Hypothesis*

| Group Statistics                          |                                             |          |       |                |  |
|-------------------------------------------|---------------------------------------------|----------|-------|----------------|--|
|                                           | Have you completed any assessment training? | <i>N</i> | Mean  | Std. deviation |  |
| Assessment of students with a final grade | Yes                                         | 120      | 48.40 | 3.844          |  |
|                                           | No                                          | 20       | 45.55 | 4.236          |  |

Table 18

*Independent Samples Test*

| Assessment of students with a final grade | Levene's test for equality of variances |       | t-test for equality of means |        |                 |                 |                       |                                           |       |
|-------------------------------------------|-----------------------------------------|-------|------------------------------|--------|-----------------|-----------------|-----------------------|-------------------------------------------|-------|
|                                           | <i>F</i>                                | Sig.  | <i>t</i>                     | df     | Sig. (2-tailed) | Mean difference | Std. error difference | 95% confidence interval of the difference |       |
|                                           |                                         |       |                              |        |                 |                 |                       | Lower                                     | Upper |
| Equal variances assumed                   | 0.119                                   | 0.730 | 3.025                        | 138    | 0.003           | 2.850           | 0.942                 | 0.987                                     | 4.713 |
| Equal variances not assumed               |                                         |       | 2.821                        | 24.501 | 0.009           | 2.850           | 1.010                 | 0.767                                     | 4.933 |

## Hypothesis 4:

Alternative Hypothesis ( $H_a$ ): There are statistically significant correlations between teachers' experience and students' assessment through experimental work, projects, handmade works.

Hypothesis zero ( $H_0$ ): There is no statistically significant correlation between teachers' experience and students' assessment through experimental work, projects, handmade works.

According to the results, teachers with experience of 6-15 years have the highest averages,  $M = 4.34$ , then those with experience 16-25,  $M = 4.31$ , then experienced teachers 26-35,  $M = 4.3$ , experienced teachers up to five years,  $M = 3.88$ , while teachers with experience over 35 years have the lowest average,  $M = 3.73$ ,  $F = 2.629$ ,  $p = 0.037 < 0.05$ .

The results showed that there are differences between teachers with different experience in assessing students with final grades. Most of them who take experimental work, projects, and handmade works as a basis are teachers who have 6-15 years of experience. From these results we can conclude that the alternative hypothesis is endorsed.

Table 19

*Testing of the Fourth Hypothesis*

| Descriptives                                                                           |     |      |       |            |                                  |             |
|----------------------------------------------------------------------------------------|-----|------|-------|------------|----------------------------------|-------------|
| Students are also assessed on the basis of experimental work, projects, handmade works |     |      |       |            |                                  |             |
|                                                                                        | N   | Mean | SD    | Std. error | 95% confidence interval for mean |             |
|                                                                                        |     |      |       |            | Lower bound                      | Upper bound |
| Up five years                                                                          | 17  | 3.88 | 0.993 | 0.241      | 3.37                             | 4.39        |
| 6-15 years                                                                             | 44  | 4.34 | 0.645 | 0.097      | 4.14                             | 4.54        |
| 16-25 years                                                                            | 35  | 4.31 | 0.631 | 0.107      | 4.10                             | 4.53        |
| 26-35 years                                                                            | 33  | 4.30 | 0.637 | 0.111      | 4.08                             | 4.53        |
| Over 35 years                                                                          | 11  | 3.73 | 1.191 | 0.359      | 2.93                             | 4.53        |
| Total                                                                                  | 140 | 4.22 | 0.759 | 0.064      | 4.09                             | 4.35        |

Table 20

*ANOVA of the Fourth Hypothesis*

| Students are also assessed on the basis of experimental work, projects, handmade works |                |     |             |       |       |
|----------------------------------------------------------------------------------------|----------------|-----|-------------|-------|-------|
|                                                                                        | Sum of squares | df  | Mean square | F     | Sig.  |
| Between groups                                                                         | 5.790          | 4   | 1.448       | 2.629 | 0.037 |
| Within groups                                                                          | 74.345         | 135 | 0.551       |       |       |
| Total                                                                                  | 80.136         | 139 |             |       |       |

## Hypothesis 5:

Alternative hypothesis ( $H_a$ ): Teachers who apply each other's grading system also practice self-assessment.

Hypothesis zero ( $H_0$ ): There is no correlation between each other's assessment and students' self-assessment.

Testing was done through Pearson correlation. Results showed that Pearson correlation coefficient,  $r = 0.647^{**}$ ,  $p = 0.000 < 0.05$ . From these results a correlation of the average level was found between the implementation of each other's assessment by students and self-assessment, so we can conclude that the alternative hypothesis is endorsed.

Table 21

*Testing of the Fifth Hypothesis*

| Correlations                |                     |                            |                             |
|-----------------------------|---------------------|----------------------------|-----------------------------|
|                             |                     | Students assess each other | Students do self-assessment |
| Students assess each other  | Pearson correlation | 1                          | 0.647**                     |
|                             | Sig. (2-tailed)     |                            | 0.000                       |
|                             | N                   | 140                        | 140                         |
| Students do self-assessment | Pearson correlation | 0.647**                    | 1                           |
|                             | Sig. (2-tailed)     | 0.000                      |                             |
|                             | N                   | 140                        | 140                         |

\*\* Correlation is significant at the 0.01 level (2-tailed).

### Qualitative Results

1. Possibly tell us, have you encountered difficulties in evaluating students?

Teacher 1: So far I have not encountered any difficulties.

Teacher 2: During the assessment of the students, I did not encounter any difficulties.

Teacher 3: I did not encounter any difficulties in assessing the students.

Teacher 4: I have not encountered any difficulties in assessing students.

Teacher 5: I did not encounter any difficulties in assessing the students.

Teacher 6: I did not encounter any difficulties in assessing the students.

Teacher 7: I have not had great difficulties, they have all been surmountable.

2. What forms of evaluation do you apply with students?

Teacher 1: To assess the student with a final grade, I rely on classroom activities, tests, homework.

Teacher 2: To assess students, I rely on classroom activities, tests, student files, homework, etc.

Teacher 3: I apply different forms of assessment.

Teacher 4: For the assessment of students I use objective tests, based on students' activities in the classroom, based on the students' files, homework, etc.

Teacher 5: I apply a large number of forms: tests, homework, oral assessment, writing, etc.

Teacher 6: I assess students orally, with tests, continuous monitoring, formative assessment, etc.

Teacher 7: I assess students continuously for tests, oral assessment, assignments, etc.

3. How much do you take student's activities into account when grading them?

Teacher 1: Every time I take into account the activities of the students to carry out the assessment of students with a final grade.

Teacher 2: I take student's activities as a basis.

Teacher 3: I take student's activities as a basis.

Teacher 4: I take very seriously the different activities of the students for their assessment with a final grade.

Teacher 5: Student engagement is a part of their assessment.

Teacher 6: I take very seriously the commitment of students to student assessment.

Teacher 7: Student activities are very important for student assessment.

4. Have you ever had a case where students felt they were more active and you did not do assessment it properly?

Teacher 1: I have not had a case where the student considered that I rated him/her with the lowest grade.

Teacher 2: I have not had a case where students have considered that they have been more active and I have not assessed it properly.

Teacher 3: I have not had any such case.

Teacher 4: Personally I have not had any such case.

Teacher 5: I have not had any such case, as I am very careful in this regard.

Teacher 6: When I assess a student, I try to convince him that he deserves that grade. I do this using all the arguments, including tests, tasks, files, and so on.

Teacher 7: I have not had any such case, as I try to be correct with all the students.

5. What have you done to improve assessment practice towards students?

Teacher 1: Because during the assessment was fine, I did not make any major changes. Maybe I started making assessments more often to come up with the final assessment.

Teacher 2: I have tried to improve the assessment constantly. Currently, after each chapter, I do a written test or oral assessment.

Teacher 3: I have attended various trainings and seminars for students' assessment and I have tried to put into practice the knowledge gained from them.

Teacher 4: I have tried to make some improvements. During my experience, I have introduced some innovations in student assessment. Currently, after each Chapter, I assess the students through the test, follow their activities, and also use the students' files.

Teacher 5: I try to follow the reforms in education, as well as in the students' assessment.

Teacher 6: Because everything was fine during the evaluation, I didn't make any major changes. Maybe I have started doing assessments more often to get to the final assessment.

Teacher 7: I have tried to build on the innovations in education and improve my assessment practice.

### Conclusions

The overall results showed that 89% of teachers agree and 2% disagree that the students rate them satisfactorily with the final grade. From these results it appears that the average perception of the teachers on the evaluation of the students with the final grade is quite high,  $M = 436$ .

Correlational results showed several cases of correlation of one action of Teachers with another action, for example, the teachers who take as a basis the commitment of Students, they also take as a basis the long-term commitment for assessment with a final grade,  $r = 0.77$ ,  $p = 0.000 < 0.05$ .

The results showed that there are differences between teachers with different education in the evaluation of students with final grades.

According to the empirical results, it was found that teachers who have completed evaluation training have a more positive approach to the student evaluation than teachers who have not completed training.

The results showed that there are differences between teachers with different experience in evaluating students with final grades. Most of them who take experimental work, projects, and handmade works as a basis are teachers who have 6-15 years of experience. From these results we can conclude that the alternative hypothesis is endorsed.

The results showed that there is an intermediate level correlation between the implementation of each other's evaluation by students and self-assessment, so we can conclude that the alternative hypothesis is endorsed.

In qualitative results, no teacher encountered any difficulty in evaluating students. For student evaluation, most teachers use a variety of forms, including tests, students' files, assignments, oral evaluation, and other activities. All teachers take as a basis the activities of students for the final grade assessment.

No teacher has had a case where students have considered themselves more active and they have not evaluated them properly.

Some teachers have not taken any action to improve assessment practice, while there are some who have used various trainings to improve assessment practice, some of them assess students after each chapter.

In general, teachers have a fairly high perception about the practice of students' assessment with a final grade.

### References

- Adams, M. E., & Ray, P. (2016). *Fundamentals of student learning assessment and grading procedures for middle and secondary school teachers*.  
 Black, L., & Wiliam, D. (1998). *Inside the Black Box-Raising Standards through Classroom Assessment*. London: King's Collage.

- Cohen, L., Manion, L., & Morrison, K. (2000). *Research methods in education* (5th edition). London and New York: Routledge/Falmer.
- Devetaku, H., & Mehmeti S. (2010). *Vlerësimi i nxënësve-praktikat dhe treguesit. Kërkime Pedagogjike*. Tiranë ISP.
- Dushku, Q. (2003). *Rreth vlerësimit të punës së mësuesit dhe të shkollës. Vlerësimi i Arsimit* (Nr. 3). Tiranë ISP.
- Gojani-Devetaku, H. (2014). *Ndikimi i vlerësimit të brendshëm në përmirësimin e cilësisë së mësimdhënies dhe nxënies. Kërkime Pedagogjike*. Prishtinë IPK.
- Joint Committee on Standards for Educational Evaluation. (1981). *Standards for Evaluations of Educational Programs, Projects, and Materials*. New York, NY: McGraw-Hill.
- Kraja, M. (2002). *Pedagogjia. Libri Universitar*. Tiranë ISP.
- Lewis, J. (2005). *Studime Akademike*. Tiranë CDE.
- Limanaqi, E. (2003). *Vlerësimi i nxënësit, një çështje në zhvillim* (Nr. 3). Tiranë ISP.
- MASHT. (2001). *Korniza e Kurrikulit të Ri të Kosovës*. Prishtinë UP.
- MASHT. (2015/2016). *Vlerësimi sipas kurrikulës së bazuar në kompetenca*.
- MASHT. (2016). *Korniza e Kurrikulit të Ri të Kosovës*. Prishtinë UP.
- MASHT. (2017). *Korniza e vlerësimit të nxënësit në arsimin parauniversitar*. Luanda: IZHA.
- MASHT-UNMIK. (2004). *Udhëzim administrativ për vlerësimin e shkollës së arritshmërisë së nxënësve*. Prishtinë UP.
- Mita, N. (2003). *Vlerësimi në bazë të shkollës, Qendra për Arsimin e Kosovës*. Prishtinë UP.
- Mula, M. (2011). *Matematika dhe mësimdhënia e matematikës, Udhëzues për klasat 6-9*. Prishtinë GIZ.
- Musai, B. (1999). *Psikologji edukimi*. Tiranë Pegi.
- Musai, B. (2003). *Metodologji e Mësimdhënies. "Pegi"*. Tiranë CDE.
- Nevo, D. (1995). *School based evaluation: A dialogue for school improvement*. Oxford: Pergamon.
- Pasku, V. (2003). *Ruajtja e cilësisë dhe të vlerësimit të shkollës. Vlerësimi i Arsimit* (Nr. 3). Tiranë ISP.
- Petrela, S. (2001). *Monitorimi dhe vlerësimi i arsimit shqiptar*, Tiranë MASH-AEDP/SOROS.
- Robo, M. (2003). *Diagnostikimi i përpërimit të nxënësve* (Nr. 3). Tiranë ISP.
- Scriven, M. (1967). *The methodology of evaluation*. In R. E. Stake (Ed.) *AERA, monograph series on curriculum evaluation* (No. 1). Chicago: Rand McNally.
- Tyler, R. W. (1950). *Basic principles of curriculum and instruction*. Chicago: University of Chicago Press.
- Woolfolk, A. (2011). *Psikologji edukimi. Botimi i njëmbëdhjetë*. Tiranë CDE.
- Zeneli, I. (2003). *Bazat e standardizimit në arsim, "Libri shkollor"*. Prishtinë UP.
- Zylfiu, N. (2005). *Didaktika (Teoria e mësimit dhe e mësimdhënies)*. Prishtinë UP.
- Zylfiu, N. (2011). *Didaktika*. Prishtinë UP.