

## The differences in formative assessment evaluation between teachers and students – a non- parametric analysis

**Lulzim Thaçi**<sup>a</sup>, UBT College – Higher Education Institution, 10000, Prishtinë, Republic of Kosovo, <https://orcid.org/0000-0003-3902-9145>

**Xhevat Sopi**<sup>b \*</sup>, Public University “Kadri Zeka”, Faculty of Economics, 60000 Gjilan, Republic of Kosovo, <https://orcid.org/0000-0003-0081-2079>

### Suggested Citation:

Thaçi, L. & Sopi, X. (2022). The differences in formative assessment evaluation between teachers and students – a non- parametric analysis. *Cypriot Journal of Educational Science*. 17(5), 1631-1646 <https://doi.org/10.18844/cjes.v17i5.7256>

Received from January 03, 2022; revised from April 01, 2022; accepted from May 12, 2022.

©2022 Birlesik Dunya Yenilik Arastirma ve Yayıncılık Merkezi. All rights reserved.

### Abstract

This study aims to assess the impact of the application of formative assessment on student success in the Republic of Kosovo and to highlight the differences in assessment between teachers and students. The research included 217 teachers and 245 students from 34 lower secondary schools, from different cities and villages of the Republic of Kosovo. Separate questionnaires were prepared for both groups, while 5-point Likert questions were used to measure the evaluation of the effects of formative assessment. The non-parametric one-sample Wilcoxon signed rank test and Mann–Whitney U test methods were used to analyse the results and test the hypothesis. The results of the Wilcoxon test with 95% confidence level show that both teachers and students evaluate positively the effects of formative assessment, while the Mann–Whitney U test results show that for particular aspects there are significant differences in assessment between these two groups for the formative assessment method.

**Keywords:** Formative assessment, learning outcomes, teaching practice, curriculum, non-parametric methods, Kosovo.

\*ADDRESS FOR CORRESPONDENCE: Xhevat Sopi, Public University “Kadri Zeka”, str. “Zija Shemsiu” n.n. 60000, Gjilan, Republic of Kosovo.  
E-mail address: [xhevat.sopi@uni-gjilan.net](mailto:xhevat.sopi@uni-gjilan.net)

## 1. Introduction

Nowadays, we cannot imagine improving the quality of education without constant changes in the use of modern forms and methods of teaching and assessment of student achievement. Modern schools focus, above all, on improving students' skills and abilities for lifelong development and establishing their role as active participants in the learning process, rather than as passive listeners and performers of activities. Therefore, the role of the teacher shifts to a mentor, guide and facilitator of students' activities, to enable them to develop independent critical thinking and to participate directly in research and problem-solving. Therefore, teaching and assessment are an integral part of the learning process, and in order to be effective and efficient, the most appropriate forms and methods must also be used for assessment.

In recent decades, Kosovo has faced major and ongoing changes in the educational process, mainly in teaching and assessment, and needs to use the most diverse assessment methods in the complex assessment process, such as formative, summative, diagnostic and so on, which enables the achievement of the highest standards in assessment at all levels of education. Since 2016, in Kosovo, formative assessment has become mandatory at all levels in the educational process.

Formative assessment, as Brown (2019) points out, offers a range of possibilities and alternatives, through various techniques and instruments to obtain relevant data that will help the teacher make a fair assessment of achievement, in terms of its results learning and about the many abilities of students, such as communication, responsibility, willingness to learn throughout the life etc., which are also provided in each curriculum.

Experience shows that during formative assessment students are not subjected to stress, which is a key factor in the assessment. Cullinane (2011) emphasises that this is mainly due to the fact that formative assessment is an ongoing process and uses a variety of techniques and instruments, including observation, practical activity and so on.

According to administrative instruction (MEST, 2016), it is specified that, *assessment aims at collecting information for supporting students in mastering learning outcomes, determining the performance level, regular progress reporting, and orientation for further education.*

At this stage, it is important to see what the results of the application of this evaluation method in Kosovo are, under the limitation caused by the short time of its application, keeping in mind that any change is not simple and easy. Being a new approach, Kosovo has a serious lack of research in the field of formative assessment. Even these few works have been realised earlier when not long ago the application of formative evaluation had started, and their focus is how much this form of evaluation has found application (Ahmedi, 2019). So far, there is no in-depth research on understanding the effects of formative assessment in Kosovo from the perspective of teachers, especially from the perspective of students. There is also no research that has assessed whether there is a difference between the attitudes of teachers and students about formative assessment in general and its specific effects.

### 1.1. Theoretical framework

Knowledge of some basic characteristics of assessment is, of course, a condition for determining its impact on student achievement. According to Leite et al. (2006), formative assessment contributes to improving and strengthening knowledge, as well as increasing responsibility in the learning process. The importance of gathering and interpreting

information during the formative assessment process was confirmed by Hanna and Dettmer (2004), who emphasise that the data serve mainly to not only inform teachers and students but also parents and other subjects involved.

Scriven (1967) uses the terms formative and summative evaluation to distinguish between information about the results to be achieved and how to achieve them. Therefore, there are learning outcomes (formative assessment) and outcomes that are measurable at the end of a given period (summative assessment).

Janusheva et al. (2017) confirm the great role of formative assessment as evidence-based assessment. Many researchers, such as Sawchuk (2015), Diloreto et al. (2017) and Othman (2018), believe in the positive effects of formative assessment and point out that in many cases, when education policy changes, teachers need to be trained to be able to evaluate student achievement with reliability and punctuality.

There are several definitions to describe formative assessment, such as William (2011), who describes it as a process of collecting and interpreting data (evidence) on students' current level of learning. The teacher uses the information gathered to 'adapt' teaching and learning for students, i.e., to elevate further learning. These views are complemented by Clark (2012), who emphasises that formative assessment is associated with two purposes, namely assessment for learning and assessment of learning. Therefore, according to him, formative assessment can be used to not only assess the learning process of students but also teach them to achieve the required results.

Cronbach (1963) and Scriven (1967) point out that formative assessment refers to the collection of information to evaluate the effectiveness of the curriculum and how it can be adapted and improved. Van Diggelen et al. (2016) and Voinea (2018) describe formative assessment as an educational element that enriches teaching and learning and helps educational institutions and educational policy, both theoretically and practically. Prashanti and Ramnarayan (2019) argue that effective implementation of formative assessment can transform a classroom culture into an approach that echoes a learning triumph.

From all these research studies, it turns out that formative assessment is a successful concept because it provides students with greater responsibility in lifelong learning. This also increases their motivation to learn, because it makes the learners co-owners of knowledge and learning, which, in turn, undoubtedly leads to increased learning outcomes.

Blum (1968) also mentions the term formative assessment in his book *Mastery Learning* and defines it as an instrument that helps improve student teaching and learning. There are also opinions that the concept of formative evaluation needs to be revised and even reformulated. Thus, Black and William (2009) define the concept of formative assessment, based on the differences between the goals and methods of summative and formative assessment, but believe that this concept should be expanded because, as they say, 'Classroom assessment is formative if the notes of teachers and students for student achievement are interpreted and used to decide the next steps to improve learning'.

### *1.2. Related research – effects of formative assessment*

Formative assessment contributes to raising students' awareness that learning is a process in which learning skills are built. Students should receive concrete and constructive instructions to improve their learning skills. In this sense, the teacher gives great support through formative assessment. Managing student engagement becomes easier if they have clear and

accurate performance criteria and are familiar with them in advance. Students themselves can set criteria, depending on the goal to be achieved.

According to Black and William (2001) and Menendez et al. (2019), formative assessment is a good strategy for increasing collaboration between students during the learning process as a tool that enhances the quality of student achievement. In this context, formative evaluation is attributed to a number of special effects which is a concise way to fulfil the purpose.

#### *1.2.1. Students motivation*

Yin et al. (2008) and Cauley and McMillan (2010) confirm the impact of implementing formative assessment on motivating students to engage in achieving set goals. Research by Weurlander et al. (2012) also confirms that formative assessment motivates students to learn. Even in the study of Thaci et al. (2020), it is confirmed that formative assessment increases students' motivation to learn, which concludes that formative assessment can act as a motivating tool in the learning process and learning outcomes.

However, some studies emphasise that to get a complete picture of the impact of formative assessment on students' motivation to learn. Shaik et al. (2020) are not clear enough, which are the best examples, in terms of improving student performance.

#### *1.2.2. Decreasing student stress*

Although teachers are constantly trying to find the most appropriate assessment techniques and forms, of course, the assessment process is very stressful for any student assuming the final result. In this sense, Pope (2005) concludes that the level of stress in students during a formative assessment is lower. Research by Cardozo et al. (2020), on the impact of formative assessment on student stress, shows that students learning with traditional teaching methods show more stress and anxiety during assessment and exams than those learning with modern methods and continued use of active formative assessment methods.

#### *1.2.3. Variety of techniques used*

During formative assessment as a learning process, students face different requirements and different tasks, so it is reasonable to use different types of forms and methods or techniques for formative assessment that we encounter in everyday learning, for example, quizzes, debates, diagrams, self-assessment and so on. According to Cauley and McMillan (2010), different techniques help students feel supported during the lesson. Caraivan (2012) analyses the role of constructive assessment, formative techniques in higher education, and finds that the use of different formative assessment techniques can reach better results.

#### *1.2.4. Feedback*

Feedback is one of the key elements of formative evaluation. According to Bauer-Ramazani et al. (2016), the main advantage of feedback is that it gives teachers more opportunities to provide up-to-date information on student achievement. Graney (2018) shows that the characteristic of feedback is for students to think independently about the quality of their work, self-judgment and the degree of achievement. The impact of feedback on improving outcomes was confirmed by the research of Aridah and Iswari (2021) as well.

Some researchers, such as Brunit et al. (2000), argue that there is ample evidence to suggest that negative feedback may be more powerful than positive feedback, but this depends on commitment, goal-orientation and student self-efficiency.

### *1.2.5. Objectivity and reliability in evaluation*

In teaching practice, making the assessment is more important than teaching the students because assessment depends on the assessor and the assessor can negatively affect perceptions of learning. Stiggins et al. (2006) suggest that the chances of inaccurate assessment are high, so students should be allowed to form their assessment prediction before receiving a grade because their grade is often more realistic than that given by teachers.

Research by Tridane et al. (2015) shows that the reliability of the assessment during formative assessment time is much greater. According to Pettifor and Saklofske (2011), evaluation cannot be a recipe, so the best ethical examples from practice should be used to encourage teachers to open a dialogue that reflects a comprehensive picture of 'who, what, where, when and why', for student assessment.

### *1.2.6. Increasing student success*

Many studies show that formative assessment helps students improve learning and achieve very high results. According to Hammonds et al. (2017), it offers the opportunity to improve classroom progress and assess student learning, through their involvement in the process.

Huisman (2018) states that formative assessment methods used in the classroom positively affect student learning and improve student achievement. Andersson and Torulf (2017) and Zainuddin et al. (2020) point out that formative assessment can also help students identify their strengths and weaknesses and the target areas for which they should work.

## *1.3. Purpose of the study – objectives and hypothesis*

As a result of the above-mentioned literature review that was conducted, we can see that there is a lack of studies regarding formative evaluation in Kosovo. Based on this and also on the findings of related research, the purpose of the study is to assess the attitudes of teachers and students regarding the effects of formative assessment on specific aspects of it and to estimate whether there are differences in assessments between teachers and students.

### *1.3.1. Objectives*

The general objective of this research is to evaluate the results of the application of formative assessment in the educational process based on the attitudes of teachers and students. In a more detailed analysis, the research aims to look at the specific assessments of students and teachers for particular effects of formative assessment, as well as to highlight whether or not there are differences in assessment between these two groups.

### *1.3.2. Hypotheses*

In order to achieve the general and specific objectives, alternative hypotheses are formulated as follows:

- H<sub>1</sub>: Formative assessment contributes to increasing student motivation.
- H<sub>2</sub>: Formative assessment contributes to reducing stress in students.
- H<sub>3</sub>: Formative assessment contributes to more meaningful feedback.
- H<sub>4</sub>: Formative assessment is carried out through various assessment techniques.
- H<sub>5</sub>: Formative evaluation contributes to increasing the objectivity of evaluation.

H<sub>6</sub>: Formative assessment contributes to improving student achievement.

For all these effects of formative assessment that have been hypothesised above, a hypothesis has been raised on the existence of differences in assessment between the two groups taken in the study, i.e., students and teachers, generally formulated as follows:

H<sub>(1-6)</sub>: There are differences between students and teachers in assessing the *particular effects* (1–6) of applying formative assessment.

## 2. Methods and procedures

The research is based on the application of non-parametric quantitative methods which come as a result of the nature of the data. Quantitative analysis is based on structured interview schemes and models where questionnaires were prepared for surveying teachers and students. The questions address the hypotheses raised, which examine the views and opinions of teachers and students on some of the main concepts of formative assessment, namely motivation, stress, feedback, techniques and instruments used in formative assessment; the objectivity of assessment; and in increasing the results of students in general.

### 2.1. Methods

In order to produce the most accurate and valid results, quantitative methods with empirical data are used. Using the 5-point Likert scale assessment implies the use of non-parametric methods (Anderson et al., 2017; Field, 2018; Gujarat, 2004; Osmani, 2013), which enable to make inferences about a population without requiring an assumption about the specific form of the population's probability distribution. For this reason, these non-parametric methods are also called distribution-free methods (Anderson et al., 2017). The methods we use in this study overcome distributional problems by ranking the data, i.e., finding the lowest score and giving it a rank of 1, then finding the next highest score and giving it a rank of 2 and so on (Field, 2018).

The Likert scale is used in such a way that teachers and students evaluate statements about the positive effects of formative assessment. The five evaluation alternatives are 'strongly disagree', 'disagree', 'have not clearly thought', 'agree' and 'strongly agree'. These alternatives are quantified with numbers from 1 to 5 where number 1 corresponds to the strongly disagree alternative and 5 to the strongly agree alternative. In the case of hypothesis testing, if the groups positively evaluate the impact of the formative evaluation, then by using the IBM Statistical Package for the Social Sciences (SPSS) software package we will apply the one-sample Wilcoxon signed rank test, which is based on the median.

In our case, alternative hypotheses will be accepted as statistically significant if most teachers and students have a positive assessment of the specific effect of formative assessment. This means that alternative hypotheses for each case are accepted if most of the answers fall on alternatives 4 or 5, respectively:

H<sub>0</sub>: Median ≤ 3.5;

H<sub>1</sub>: Median > 3.5.

This hypothesis formulation implies a one-way test for hypothesis testing. Since the statistical software (in our case SPSS) provides data only for the two-sided test, then we divide the probability (*p*) value by 2 and thus obtain the value of the one-sided test. Care must be taken that the value of *t*-statistics is in the direction of one-sided proof so that in our case it should

be positive. The significance level is  $\alpha = 0.05$  (confidence level of 95%), which means that the alternative hypothesis is accepted if  $p < \alpha$ .

Testing hypotheses about the difference in evaluation between the two groups, which is a continuation of the hypotheses about the evaluation that groups make of the effects of formative evaluation, the most suitable is the Mann–Whitney U test, which is the non-parametric equivalent of the independent *t*-test. As with the above-mentioned method (Wilcoxon test), this method also tests the difference between the two groups based on the ranks, respectively, in the mean ranks. It shows how identical or non-identical the two groups are in terms of the views of a particular statement that supports a particular hypothesis. For hypothesis testing, we will use a level of significance of  $\alpha = 0.05$  (confidence level of 95%). For each particular test effect, there are two hypotheses where:

$H_0$ : There are no differences in assessment between groups.

$H_1$ : There are differences in assessment between groups.

If  $p < \alpha$ , then  $H_1$  is accepted, otherwise we hold  $H_0$ .

## 2.2. The data

Two questionnaires were prepared for data collection for both study groups: teachers and students. The questionnaires were constructed according to the Likert scale, based on the results of the interview after analysis and coding of the results (Creswell, 2014). Both questionnaires contain six statements/claims. Teachers and students have the opportunity to choose one of the offered alternatives, regarding the effects of formative assessment.

The sample consisted of 217 teachers and 245 students from 34 secondary schools in rural and urban municipalities of the Republic of Kosovo. After the data were cleaned, they were encoded and transferred to the SPSS programme, which is used to analyse the results and test the hypotheses.

## 2.3. The reliability and validity tools

To ensure the validity of data analysis and questionnaires, we will apply the reliability check through the application of Cronbach's alpha coefficient, which is the most common measure of internal consistency reliability. When using Likert-type scales, it is imperative to calculate and report Cronbach's alpha coefficient for internal consistency reliability for any scales or subscales one may be using (Gliem & Gliem, 2003). We often see in books or journal articles, or are told by people, that a value of 0.7–0.8 is an acceptable value for Cronbach's alpha and that values substantially lower indicate an unreliable scale (Field, 2018). According to Gliem and Gliem (2003), the analysis of the data must use these summated scales or subscales and not individual items because Cronbach's alpha does not provide reliable estimates for single items but for the overall reliability.

For the reliability check of our data, we will use the SPSS programme which has several options. One of the most useful options for questionnaire reliability is *Cronbach's Alpha if Item Deleted*, which tells us what the value of alpha would be if each item were deleted (Field, 2018). Regarding these options, if our questionnaire is reliable, then we would not expect any item to greatly affect the overall reliability. The second option *Corrected Item-Total Correlation* shows the correlations between each item and the total score from the questionnaire. If any of these values are less than about 0.3, then we have got problems,

because it means that a particular item does not correlate very well with the scale overall (Field, 2018).

### 3. Results and discussion

In this section, first, the results of descriptive statistics related to the characteristics of the sample are presented and then the results of the statistical analysis of empirical data collected through questionnaires are presented, which are those for the frequencies of teachers and students' attitudes and hypotheses' test results.

#### 3.1. Descriptive statistics

From the analysis of the data in Table 1, we see that among teachers the gender representation is approximately equal where 47.5% are women and 52.5% are men, while among students, women (56.7%) are more than men (43.3%). Regarding the age groups, in Table 1 we see in the teacher category that the dominant are the age groups from 26 to 55 years wherein each of the three groups are represented approximately the same number of teachers, about 25% per group, which in total give 75% of the teachers, while the other 25% is made up of the other two age groups, under 26 and over 55, respectively. The distribution among students is more diverse where almost half are in the IX grade (46.5%), while 29% are in the VIII grade and 24.5% are in the VII grade.

Table 1. Descriptive statistics of sample characteristics

Gender			Frequency	Percent	Cumulative Percent
<b>Teacher</b>	Women		103	47.5	47.5
	Man		114	52.5	100.0
	Total		217	100.0	
<b>Student</b>	Women		139	56.7	56.7
	Men		106	43.3	100.0
	Total		245	100.0	
<b>Age</b>	<b>Teacher</b>	Up to 25 years	20	9.2	9.2
		26–35	54	24.9	34.1
		36–45	56	25.8	59.9
		46–55	52	24.0	83.9
		56–65	35	16.1	100.0
		Total	217	100.0	
		<b>Grade</b>	<b>Student</b>	7	60
8	71			29.0	53.5
9	114			46.5	100.0
Total	245			100.0	

Source: Field survey – processing by SPSS.

Table 2 summarises the results for the statements of students and teachers separately for each of the six questions through which the six hypotheses raised are tested. For efficiency, only the relative frequencies from which the absolute frequencies can be found are pre-recorded. Regarding the central tendency in this table, we have included only the median because the testing of hypotheses regarding the position of groups for certain effects is carried out through the non-parametric Wilcoxon signed rank test method.



In summary, from Table 2, we see that in all cases the median falls into positive evaluation modalities for the effects of formative evaluation. In most cases, the median falls into mode 4 (agree) and there are three cases when the median falls into mode 5 (strongly agree). Over 50% of the students chose the ‘strongly agree’ option in terms of claims that formative assessment reduces stress and increases student engagement. On the other hand, teachers value the effect that formative assessment enables continuous information on the result/success of students. Precisely for the effect of informing about the result, students chose the most negative modalities (4.5% strongly disagree and 10.6% disagree), which indicates a significant difference between the assessments of students and teachers.

Table 2. Descriptive statistics of questions and answers by groups

Questions <sup>a</sup> :		Q <sub>1</sub>	Q <sub>2</sub>	Q <sub>3</sub>	Q <sub>4</sub>	Q <sub>5</sub>	Q <sub>6</sub>						
Groups	Modalities	Fr %	Me d	Fr %	Me d	Fr %	Med	Fr %	Med	Fr %	Me d	Fr %	Med
T E A CH E R S	Strongly disagree	0.5		0.0		.9		0.5		1.4		0.0	
	Disagree	2.8		6.9		4.6		3.2		0.0		1.8	
	Neutral <sup>b</sup>					11.		18.					
	Agree	10.6	4	15.7	4	5	4	9	4	5.5	5	12.4	4
	Strongly agree	47.9		36.4		9		5		6		42.9	
	Total		100		100		100		100		100		100
S T U D E N T S	Strongly disagree	3.7		.4		4.5		4.5		4.5		0.8	
	Disagree									10.			
	Neutral	7.3		1.6		7.8		7.3		6		2.9	
	Agree	18.4	4	9.0	5	1	4	7	4	4	4	4.1	5
	Strongly agree	39.6		15.1		7		8		6		23.3	
	Total		100		100		100		100		100		100

<sup>a</sup> Q<sub>1</sub>: When cooperating with friends I am (they are) more motivated; Q<sub>2</sub>: During the hourly activities the stress of the evaluator is not felt; Q<sub>3</sub>: During the lesson we can discuss and ask (they can ask) questions more easily; Q<sub>4</sub>: More than 5 five forms of assessment are used in formative assessment classes; Q<sub>5</sub>: I am constantly (students are) informed about the level of achievement and results; Q<sub>6</sub>: During the hours I am (they are) more engaged.

<sup>b</sup> The neutral option, in this case, means that teachers and students do not have a clear opinion about the question/statement of the questionnaire.

Source: Field survey – processing by SPSS.

### 3.2. The reliability and validity results

As elaborated in Section 2.3, Cronbach’s alpha coefficient was used to assess the validity and reliability of the data collection. The results from the processing in SPSS show that we have a good degree of overall internal consistency reliability, where the value of Cronbach’s alpha coefficient is greater than 0.7 ( $\alpha = 0.747$ ).

Even in terms of detailed analysis for individual items, from the results presented in Table 3, we see that we have a good degree of internal consistency reliability. In the column labelled

*Cronbach's Alpha if Item Deleted*, all individual values are less than Cronbach's alpha coefficient value (0.747) for overall reliability. This means that none of these items would increase the overall alpha value if removed from the analysis. Also, the results in the column labelled *Corrected Item-Total Correlation* are all above the value of 0.3, which means that all items correlate very well.

Table 3. Item-total statistics

Items	Scale mean if item deleted	Scale variance if item deleted	Corrected item-total correlation	Cronbach's alpha if item deleted
Q1	21.31	7.457	0.610	0.677
Q2	21.41	7.437	0.476	0.715
Q3	21.35	7.246	0.558	0.689
Q4	21.18	7.691	0.428	0.729
Q5	21.09	8.112	0.486	0.712
Q6	21.27	8.634	0.372	0.738

Source: Field survey – processing by SPSS.

### 3.3. Non-parametric tests results

#### 3.3.1. One-sample Wilcoxon signed rank test

To be able to estimate the whole population through the results from the samples presented in Table 2, the non-parametric test one-sample Wilcoxon signed rank test was applied, through which we test the hypotheses for the effects of formative evaluation in the population. From the test results applied separately for students and teachers and also separately for each hypothesis, the test shows that the results in all cases are statistically significant:  $p < \alpha$  for  $\alpha = 0.05$  (see Table 4).

Table 4. One-sample Wilcoxon signed rank test

	Q1		Q2		Q3		Q4		Q5		Q6	
	Teach.	Stud.	Teach.	Stud.	Teach.	Stud.	Teach.	Stud.	Teach.	Stud.	Teach.	Stud.
<b>Test Statistic</b>	20,932	21,378	19,233	28,622	20,098	19,812	19,824	21,690	22,399	20,335	21,344	28,218
<b>Standard error</b>	894	1,075	897	1,047	897	1,077	896	1,077	898	1,079	896	1,055
<b>Stand. test statistic</b>	10.18	5.83	8.26	12.94	9.22	4.41	8.92	6.15	11.78	4.88	10.62	12.46
<b>Sig. (2-tailed)</b>	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Source: Field survey – processing by SPSS.

These results show that in all cases for 95% confidence level, the median  $> 3.5$ , which means that both students and teachers, to the extent of over 50%, hold the view that formative assessment has positive effects on the learning process. So positive modalities for the effects of formative assessment (4 = agree and 5 = strongly agree) are chosen by more than 50% of both students and teachers.

### 3.3.2. Results on hypotheses testing for formative effects

Based on the results of the one-sample Wilcoxon signed rank test, we conclude that for significance level at  $\alpha = 0.05$ , we have sufficient evidence to reject the null hypotheses and to accept the alternative hypotheses for all six hypotheses raised, respectively, we accept the following hypotheses:

- H<sub>1</sub>: Formative assessment contributes to increasing student motivation.
- H<sub>2</sub>: Formative assessment contributes to reducing stress in students.
- H<sub>3</sub>: Formative assessment contributes to more meaningful feedback.
- H<sub>4</sub>: Formative assessment is carried out through various assessment techniques.
- H<sub>5</sub>: Formative evaluation contributes to increasing the objectivity of evaluation.
- H<sub>6</sub>: Formative assessment contributes to improving student achievement.

### 3.3.3. Mann–Whitney U Tests

Regarding the testing of hypotheses and the difference in assessment between students and teachers for the special effects of formative assessment, respectively, for the hypotheses tested above, the following are the results from the application of the Mann–Whitney U test. This method tests the difference between the ranks of the two groups. In Table 5, we can see the rank averages and the differences between the two groups for each of the statements/questions posed in the questionnaire. As we can make comparisons, in all cases we have differences in mean ranks between the two groups.

Table 5. Mean and sum of ranks

		N	Mean rank	Sum of ranks
<b>Increasing motivation-Q1</b>	Teachers	217	251.69	54,617.00
	Students	245	213.62	52,336.00
<b>Stress reduction-Q2</b>	Teachers	217	191.65	41,587.00
	Students	245	266.80	65366.00
<b>Significant feedback-Q3</b>	Teachers	217	255.85	55,518.50
	Students	245	209.94	51,434.50
<b>At least five forms of evaluation-Q4</b>	Teachers	217	245.32	53,235.50
	Students	245	219.26	53,717.50
<b>Ongoing information on the result-Q5</b>	Teachers	217	264.09	57,308.50
	Students	245	202.63	49,644.50
<b>Engagement increases during the hour-Q6</b>	Teachers	217	199.99	43,397.00
	Students	245	259.41	63,556.00

Source: Field survey – processing by SPSS.

How much are the mean rank differences between the groups from Table 5 are statistically significant we will see in the results of the Mann–Whitney U tests, presented in Table 6. According to the test statistics, it turns out that in all cases the differences in attitudes between the two groups are statistically significant:  $p < \alpha$  for  $\alpha = 0.05$ . These results show that for a 95% level of reliability, we can conclude that the attitudes of the groups regarding the effects of formative assessment are not identical.

Table 6. Mann–Whitney U test about differences between groups

Test statistics <sup>a</sup>						
	Q <sub>1</sub>	Q <sub>2</sub>	Q <sub>3</sub>	Q <sub>4</sub>	Q <sub>5</sub>	Q <sub>6</sub>
<b>Mann-Whitney U</b>	22,201.000	17,934.000	21,299.500	23,582.500	19,509.500	19,744.000
<b>Wilcoxon W</b>	52,336.000	41,587.000	51,434.500	53,717.500	49,644.500	43,397.000
<b>Z</b>	-3.273	-6.825	-3.894	-2.222	-5.301	-5.396
<b>Sig. (2-tailed)</b>	<b>0.001</b>	<b>0.000</b>	<b>0.000</b>	<b>0.026</b>	<b>0.000</b>	<b>0.000</b>

<sup>a</sup> Grouping variable: teachers/students.

Source: Field survey – processing by SPSS.

### 3.3.4. Results on hypotheses testing for differences between groups about attitudes to the effects of formative evaluation

The results of the Mann–Whitney U test statistics show that for the significance level of  $\alpha = 0.05$  of the tailed tests, we have sufficient evidence for all cases to reject the null hypotheses and accept the alternative hypotheses, respectively; we accept the hypotheses that:

*There are statistically significant differences between teachers and students' attitudes regarding the effects of formative assessment. These differences are statistically significant separately for each effect tested according to the hypotheses in Section 3.3.2.*

## 4. Conclusion

From the analysis of the results, we conclude that formative assessment is a process that is welcomed by teachers and students and that both groups are very satisfied with the effects of its implementation on the learning process. They generally positively evaluate the effects of formative assessment, in terms of both motivation, which is followed by increased student engagement, and reduced stress and increased objectivity in assessment, which together with the application of a greater number of techniques assessment bring better results for student success and the learning process in general.

On the other hand, research shows that although both students and teachers positively evaluate these aspects of formative assessment, there are statistically significant differences in their attitudes to all hypotheses. This indicates the need for in-depth analysis of the use of formative assessment, which will certainly contribute to improving students' achievement, their motivation and other aspects of formative assessment tested with the working hypotheses.

Differences in assessment between students and teachers imply the need for deeper and more comprehensive research into the causes of these differences. This would give a clear picture of why students do not think the same with teachers about these issues as well as teachers' willingness to adapt even more to students' needs and requirements. Further research should also be conducted at the other two levels of pre-university education which are upper secondary school and primary school to see if the results are similar to this research.

Outcomes are important for education policy and institutions in lower secondary education. They imply ongoing commitments to support implementation and create conditions for its development, such as various teacher trainings and other forms of support.

The research has its limitations, especially in terms of comparison with results before the start of the application of formative assessment in Kosovo (as there is no such research for formative assessment in Kosovo) or the possibility of conducting a comparative experiment between groups where formative assessment was applied and those where formative assessment does not apply.

## References

- Ahmedi, V. (2019). Teachers' attitudes and practices towards formative assessment in primary schools. *Journal of Social Studies Education Research*, 10(3), 161–175. <https://jsser.org/index.php/jsser/article/view/1233/392>
- Anderson, D., Sweeney, D., Williams, T., Camm, J., & Cochran, J. (2017). *Statistics for business and economics* (13th ed.). Cengage Learning. [http://lib.yzu.am/disciplines\\_bk/efbc179b18dd97d124c6ccd17bd58d28.pdf](http://lib.yzu.am/disciplines_bk/efbc179b18dd97d124c6ccd17bd58d28.pdf)
- Andersson, C., & Torulf, P. (2017). The impact of formative assessment on student achievement: A study of the effects of changes to classroom practice after a comprehensive professional development programme. *Learning and instruction*, 49, 92–102. <https://doi.org/10.1016/j.learninstruc.2016.12.006>
- Aridah, A., & Iswari, W. P. (2021). The effect of indirect feedback on students' writing performance across different learning strategies. *Cypriot Journal of Educational Sciences*, 16(3), 1021–1035. <https://doi.org/10.18844/cjes.v16i3.5820>
- Bauer-Ramazani, C., Graney, J. M., Marshall, H. W., & Sabieh, C. (2016). Flipped learning in TESOL: Definitions, approaches, and implementation. *Tesol Journal*, 7(2), 429–437. [https://www.researchgate.net/profile/Helaine-Marshall-2/publication/299554197\\_Flipped\\_Learning\\_in\\_TESOL\\_Definitions\\_Approaches\\_and\\_Implementation/links/5a62b53eaca272a15819a783/Flipped-Learning-in-TESOL-Definitions-Approaches-and-Implementation.pdf](https://www.researchgate.net/profile/Helaine-Marshall-2/publication/299554197_Flipped_Learning_in_TESOL_Definitions_Approaches_and_Implementation/links/5a62b53eaca272a15819a783/Flipped-Learning-in-TESOL-Definitions-Approaches-and-Implementation.pdf)
- Black, P., & Wiliam, D. (2001). *Inside the black box, raising standards through classroom assessment*. <http://weaeducation.typepad.co.uk/files/blackbox-1.pdf>
- Black, P. J., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5–31. [https://www.researchgate.net/publication/225590759\\_Developing\\_the\\_theory\\_of\\_formative\\_assessment](https://www.researchgate.net/publication/225590759_Developing_the_theory_of_formative_assessment)
- Bloom, B. S. (1968). Learning for mastery. *UCLA Evaluation Comment*, 1(2), 1–12. <https://files.eric.ed.gov/fulltext/ED053419.pdf>
- Brown, G. T. L. (2019). Is assessment for learning really assessment? *Frontiers in Education*, 4(64). <https://www.frontiersin.org/articles/10.3389/educ.2019.00064/full>
- Brunit, S., Huguet, P., & Monteil, J. M. (2000). Performance feedback and self-focused attention in the classroom: When past and present interact. *Social Psychology of Education*, 3, 277–293. [https://www.researchgate.net/profile/Sophie-Brunot/publication/227145090\\_Performance\\_feedback\\_and\\_self-focused\\_attention\\_in\\_the\\_classroom\\_When\\_past\\_and\\_present\\_interact/links/58a2203245851598babae944/Performance-feedback-and-self-focused-attention-in-the-classroom-When-past-and-present-interact.pdf](https://www.researchgate.net/profile/Sophie-Brunot/publication/227145090_Performance_feedback_and_self-focused_attention_in_the_classroom_When_past_and_present_interact/links/58a2203245851598babae944/Performance-feedback-and-self-focused-attention-in-the-classroom-When-past-and-present-interact.pdf)
- Caraivan, L. (2012). *(In)formative assessment techniques in blended and e-learning environments. Leveraging technology for learning* (vol. 1, pp. 74–77). Proceedings of the 8th international scientific conference “eLearning and software for education”, Bucharest. <https://proceedings.elseconference.eu/index.php?r=site/index&year=2012&index=papers&vol=1&paper=9616b117c11713a5f05425738b8367d7>
- Cardozo, L. T., Azevedo, M. A. R. D., Carvalho, M. S. M., Costa, R., De Lima, P. O., & Marcondes, F. K. (2020). Effect of an active learning methodology combined with formative assessments on performance, test anxiety, and stress of university students. *Advances in Physiology Education*, 44(4), 744–751. <https://journals.physiology.org/doi/full/10.1152/advan.00075.2020>

- Thaçi, L. & Sopi, X. (2022). The differences in formative assessment evaluation between teachers and students – a non-parametric analysis. *Cypriot Journal of Educational Science*. 17(5), 1631-1646 <https://doi.org/10.18844/cjes.v17i5.7256>
- Cauley, M. K., & McMillan, H. J. (2010). Formative assessment techniques to support student motivation and achievement. *A Journal of Educational Strategies, Issues and Ideas*, 83(1), 1–6. <https://doi.org/10.1080/00098650903267784>
- Clark, I. (2012). Formative assessment: Assessment is for self-regulated learning. *Educational Psychology Review*, 24(2), 205–249. <https://doi.org/10.1007/s10648-011-9191-6>.
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative and mixed methods approaches*. University of Nebraska: Sage publications.  
[https://books.google.com/books?hl=en&lr=&id=51UXBAAAQBAJ&oi=fnd&pg=PR1&dq=Creswell,+J.+W.++\(2014\).+Research+design:+qualitative,+quantitative+and+mixed+methods+approaches.+University+of+Nebraska:+Sage+publications.&ots=6aHnL9YoMx&sig=gGQsrWdU-K1hM02GHTXeVTkSJ0w#v=onepage&q&f=false](https://books.google.com/books?hl=en&lr=&id=51UXBAAAQBAJ&oi=fnd&pg=PR1&dq=Creswell,+J.+W.++(2014).+Research+design:+qualitative,+quantitative+and+mixed+methods+approaches.+University+of+Nebraska:+Sage+publications.&ots=6aHnL9YoMx&sig=gGQsrWdU-K1hM02GHTXeVTkSJ0w#v=onepage&q&f=false)
- Cronbach, L. J. (1963). Course improvement through evaluation. *Teacher's College Record*, 64(8), 672–683. <https://avys.omu.edu.tr/storage/app/public/ismailgelen/116687/14.PDF>
- Cullinane, A. (2011). Formative assessment classroom techniques. *Research and Resource Guides*, 2(13), 1–4. <https://www.researchgate.net/publication/283328368> [Formative Assessment Classroom Techniques](https://www.researchgate.net/publication/283328368).
- Diloreto, M., Pellow, C., & Stout, D. (2017). Exploration of conceptions of assessment within high-stakes U.S. Culture, 16(7), 1–9. [https://www.researchgate.net/publication/319332750\\_Exploration\\_of\\_Conceptions\\_of\\_Assessment\\_within\\_High-Stakes\\_US\\_Culture](https://www.researchgate.net/publication/319332750_Exploration_of_Conceptions_of_Assessment_within_High-Stakes_US_Culture)
- Field, A. (2018). *Discovering statistics using IBM SPSS Statistics* (5th ed.). SAGE Edge. [https://www.academia.edu/10707181/Discovering\\_Statistics\\_using\\_Spss](https://www.academia.edu/10707181/Discovering_Statistics_using_Spss)
- Gliem, J. A., & Gliem, R. R. (2003). *Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for Likert-type scales*. Midwest Research-to-Practice Conference in Adult, Continuing, and Community Education. <https://scholarworks.iupui.edu/bitstream/handle/1805/344/Gliem%20%26%20Gliem.pdf?sequence=1&isAllowed=y>
- Graney, J. M. (2018). Flipped learning and formative assessment in an English language class. In: Mwehring, J., Leis, A. (eds) *Innovations in flipping the language classroom* (pp. 59–68). Springer, Singapore. [https://doi.org/10.1007/978-981-10-6968-0\\_5](https://doi.org/10.1007/978-981-10-6968-0_5)  
[https://www.google.com/books/edition/Innovations\\_in\\_Flipping\\_the\\_Language\\_Cla/R5xBDwAAQBAJ?hl=en&bpv=1&pg=PA59&printsec=frontcover](https://www.google.com/books/edition/Innovations_in_Flipping_the_Language_Cla/R5xBDwAAQBAJ?hl=en&bpv=1&pg=PA59&printsec=frontcover)
- Gujarati, D. N. (2004). *Basic econometrics* (4th ed.). The McGraw-Hill Companies. <http://zalamyah.staff.unja.ac.id/wp-content/uploads/sites/286/2019/11/7-Basic-Econometrics-4th-Ed.-Gujarati.pdf>
- Hammonds, F., Mariano, G. J., Ammons, G., & Chambers, S. (2017). Student evaluations of teaching: improving teaching quality in higher education. *Perspectives, Policy and Practice in Higher Education*, 21(1), 26–33. <https://doi.org/10.1080/13603108.2016.1227388>
- Hanna, G. S., & Dettmer, P. (2004). *Assessment for effective teaching: Using context-adaptive planning*. Pearson A and B. <https://www.pearson.com/us/higher-education/program/Hanna-Assessment-for-Effective-Teaching-Using-Context-Adaptive-Planning/PGM2811397.html>
- Huisman, M. (2018). *Formative assessment and the impact on student learning*. National center for research on evaluation, standards, and student testing (CRESST). [https://nwcommons.nwciowa.edu/cgi/viewcontent.cgi?article=1080&context=education\\_masters](https://nwcommons.nwciowa.edu/cgi/viewcontent.cgi?article=1080&context=education_masters)
- Janusheva, V., Pejchinovska, M., Talevski, J. (2017). The external assessment in the republic of macedonia – Does it justify its existence? *Teacher - International Journal of Education*, 13, 74–82. [https://www.researchgate.net/profile/Milena-Pejchinovska/publication/315798842\\_THE\\_EXTERNAL\\_ASSESSMENT\\_IN\\_THE\\_REPUBLIC\\_OF\\_MACEDONIA\\_DOES\\_IT\\_JUSTIFY\\_ITS\\_EXISTENCE/links/58e61ffaa6fdcc6800b449e3/THE-EXTERNAL-ASSESSMENT-IN-THE-REPUBLIC-OF-MACEDONIA-DOES-IT-JUSTIFY-ITS-EXISTENCE.pdf](https://www.researchgate.net/profile/Milena-Pejchinovska/publication/315798842_THE_EXTERNAL_ASSESSMENT_IN_THE_REPUBLIC_OF_MACEDONIA_DOES_IT_JUSTIFY_ITS_EXISTENCE/links/58e61ffaa6fdcc6800b449e3/THE-EXTERNAL-ASSESSMENT-IN-THE-REPUBLIC-OF-MACEDONIA-DOES-IT-JUSTIFY-ITS-EXISTENCE.pdf)

- Thaçi, L. & Sopi, X. (2022). The differences in formative assessment evaluation between teachers and students – a non-parametric analysis. *Cypriot Journal of Educational Science*. 17(5), 1631-1646 <https://doi.org/10.18844/cjes.v17i5.7256>
- Leite, C., Rodrigues, L., & Fernandes, P. (2006). A auto-avaliação das escolas e a melhorias da qualidade da educação: Um olhar reflexivo a partir de uma situação. *Revista de estudos curriculares*, 4, 21–45. <https://www.redalyc.org/pdf/275/27554785024.pdf>
- Menendez, C. Y. I., Napa, M., Moreira, M., & Zambrano, G. (2019). The importance of formative assessment in the learning teaching process. *International Journal of Social Sciences and Humanities*, 3, 238–249. <http://www.sciencescholar.us/journal/index.php/ijssh/article/view/322/326>.
- MEST (2016). *Administrative instruction*. [https://masht.rks.gov.net/uploads/2016/08/ua-nr-08-2016-per-vleresimin-e-nxeneseve\\_1.pdf](https://masht.rks.gov.net/uploads/2016/08/ua-nr-08-2016-per-vleresimin-e-nxeneseve_1.pdf)
- Osmani, M. (2015). *Metodat e Statistikës (eng:Statistical Methods)*. Pegi, Tiranë (printed book). <https://www.libriarubt.com/produkti/metodat-e-statistikes-myslym-osmani/>
- Othman, J. (2018). Reform in assessment: teachers' beliefs and practices. *Journal of Teaching English for Specific and Academic Purposes*, 6(3), 501–512. <http://espeap.junis.ni.ac.rs/index.php/espeap/article/view/711>.
- Pettifor, J. L., & Saklofske, D. H. (2011). Fair and ethical student assessment practices. In *Leading student assessment* (pp. 87–106). Springer. [https://link.springer.com/chapter/10.1007/978-94-007-1727-5\\_5](https://link.springer.com/chapter/10.1007/978-94-007-1727-5_5)
- Pope, N. K. L. (2005). The impact of stress in self-and peer assessment. *Assessment & Evaluation in Higher Education*, 30(1), 51–163. <https://doi.org/10.1080/0260293042003243896>
- Prashanti, E., & Ramnarayan, K. (2019). Ten maxims of formative assessment. *Advances in Physiology Education*, 43, 99–102. <https://www.physiology.org/doi/full/10.1152/advan.00173.2018>.
- Sawchuk, S. (2015). Teacher evaluation: an issue overview. *Education Week*, 35(3). <https://www.edweek.org/ew/section/multimedia/teacher-performance-evaluation-issue-overview.html>.
- Scriven, M. (1967). *The methodology of evaluation*. <https://fdocuments.in/document/scriven-the-methodology-of-evaluation-1967.html?page=1>
- Shaik, N., Shah, D., & Mirza, N. (2020). Impact of formative assessment on students' learning at private schools at district Sanghar. Sindh. *Pakistan Social Sciences Review*, 4(2), 473–482. <https://pssr.org.pk/issues/v4/2/impact-of-formative-assessment-on-students-learning-at-private-schools-in-district-sanghar-sindh.pdf>
- Stiggins, R. J., Arter, J. A., Chappuis, J., & Chappuis, S. (2007). *Classroom assessment for student learning: Doing it right – Using it well* (pp. 3–40). Educational Testing Service.. [https://www.bostonpublicschools.org/cms/lib/MA01906464/Centricity/Domain/99/stiggins\\_assessment.pdf](https://www.bostonpublicschools.org/cms/lib/MA01906464/Centricity/Domain/99/stiggins_assessment.pdf)
- Thaci, L., Janusheva, V., & Talevski, J. D. (2020); The effect of the formative assessment on students' motivation in the Republic of Kosovo. *International Journal of Education TEACHER, [S.I.]*, 10(19), 5–13. ISSN 1857-8888. <http://eprints.uklo.edu.mk/5440/1/document.pdf>
- Tridane, M., Belaaouad, S., Benmokhtar, S., Buhra, G., & Mohamed, R. (2015). The impact of formative assessment on the learning process and the unreliability of the mark for the summative evaluation. *Procedia – Social and Behavioral Sciences*, 197, 680–685. <https://www.iosrjournals.org/iosr-jrme/papers/Vol-5%20Issue-2/Version-2/G05224851.pdf>
- Van Diggelen, M. R., Morgan, C. M., Funk, M., & Bruns, M. (2016). *Formative assessment: enriching teaching and learning with formative assessment*. Eindhoven University of Technology. <https://docplayer.net/104547753-Document-version-publisher-s-pdf-also-known-as-version-of-record-includes-final-page-issue-and-volume-numbers.html>.
- Voinea, L. (2018). Formative assessment as assessment for learning development. *Journal of Pedagogy*, 1, 7–23. <https://www.researchgate.net/publication/326190241>  
[FORMATIVE ASSESSMENT AS ASSESSMENT FOR LEARNING DEVELOPMENT](https://www.researchgate.net/publication/326190241).

- Thaçi, L. & Sopi, X. (2022). The differences in formative assessment evaluation between teachers and students – a non- parametric analysis. *Cypriot Journal of Educational Science*. 17(5), 1631-1646 <https://doi.org/10.18844/cjes.v17i5.7256>
- Weurlander, M., Soderberg, M., Scheja, M., Hult, H., & Wernerson, A. (2012). Exploring formative assessment as a tool for learning: students' experiences of different methods of formative assessment. *Assessment & Evaluation in Higher Education*, 37(6), 747–760. <https://doi.org/10.1080/02602938.2011.572153>
- William, D. (2011). *Embedded formative assessment*. Solution Tree Press. <http://www.tdschools.org/wp-content/uploads/2012/07/The-Main-Idea-Embedded-Formative-Assessment-March-2013.pdf>
- Yin, Y., Shavelson, R., Ayala, C., Ruiz-Primo, M., Brandon, P., Furtak, E., Tomita, M., & Young, D. (2008). On the impact of formative assessment on student motivation, achievement, and conceptual change. *Applied Measurement in Education*, 21(4), 335–359. [https://www.researchgate.net/publication/233445437\\_On\\_the\\_Impact\\_of\\_Formative\\_Assessment\\_on\\_Student\\_Motivation\\_Achievement\\_and\\_Conceptual\\_Change](https://www.researchgate.net/publication/233445437_On_the_Impact_of_Formative_Assessment_on_Student_Motivation_Achievement_and_Conceptual_Change).
- Zainuddin, Z., Muhammad, S., Haruna, H., & Chu, S. (2020). The role of gamified e-quizzes on student learning and engagement: An interactive gamification solution for a formative assessment system. *Computers & Education*, 145. [https://www.academia.edu/40802741/The\\_role\\_of\\_gamified\\_e\\_quizzes\\_on\\_student\\_learning\\_and\\_engagement\\_An\\_interactive\\_gamification\\_solution\\_for\\_a\\_formative\\_assessment\\_system](https://www.academia.edu/40802741/The_role_of_gamified_e_quizzes_on_student_learning_and_engagement_An_interactive_gamification_solution_for_a_formative_assessment_system)