

# The Impact of Formative Assessment on Results of Secondary School Pupils in Mathematics: One Case of Schools in Slovakia

Lucia Rumanová, Dušan Vallo, Júlia Záhorská

*Constantine the Philosopher University in Nitra, Slovak republic, Tr. A. Hlinku 1, Nitra, Slovakia*

**Abstract** – Teaching of mathematics at Slovak primary schools used to be based on the transmissive transfer of knowledge and skills from teachers to pupils. The currently used constructive and investigative approach in teaching mathematics in Slovakia requires a change in the formative assessment in mathematics based on the pedagogic experiment, which is carried out on a target group of respondents. The issue of measuring and evaluating pupils' performance is still topical for several reasons. Pupils' results in mathematics often do not meet the expectations among pupils themselves and their parents. Measuring pupil's level of knowledge and quality of education is also one of the topics that are very often discussed by the general public. We consider the use of formative assessment to correct teachers' teaching practices and pupils' learning practices as a suitable tool for the teaching process. This fact can be applied also in teaching mathematics. It helps to identify misconceptions in pupils' knowledge, to diagnose various problems and their causes in the learning process of pupils.

**Keywords** – Formative assessment, Learning process, Mathematics, Statistical evaluation.

## 1. Introduction

In Since 1st September 2008 the Slovak education sector has carried out a wide content reform.

DOI: 10.18421/TEM93-47

<https://doi.org/10.18421/TEM93-47>

**Corresponding author:** Lucia Rumanová,  
*Department of Mathematics, Constantine the Philosopher University in Nitra, Slovakia.*

**Email:** [lrumanova@ukf.sk](mailto:lrumanova@ukf.sk)

*Received:* 29 June 2020.

*Revised:* 04 August 2020.

*Accepted:* 11 August 2020.

*Published:* 28 August 2020.

 © 2020 Lucia Rumanová, Dušan Vallo & Júlia Záhorská; published by UIKTEN. This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 License.

The article is published with Open Access at [www.temjournal.com](http://www.temjournal.com)

In this context and with the view to the education level of the young generation a very important question has been raised. It is the question of evaluation of students, their understanding of curriculum, progress, identification of learning and correcting needs cover positive and negative outputs [1].

The traditional transmissive teaching from the past, based on the transfer of knowledge and skills from teachers to pupils [2], has become obsolete. The pupil absorbs a huge amount of information whose deep and thorough understanding is impossible. At the same time, these data lose its topicality and require updating [3].

The constructive and investigative approach to teaching is used much more frequently, which requires an adequate evaluation method.

Innovative teaching methods bring in modification in the assessment of pupils. The assessment of knowledge and quality of education are issues of interest to a wider audience. This is all the truer in mathematics as the results achieved are often short of pupils and parents' expectations. The problems lie in the importance attributed to mathematical knowledge and skills instead of mathematical literacy which is emphasized in the world ([4], [5]).

The recent changes in the assessment and evaluation of results with regard to Slovak pupils indicate development in the right direction. It seems that better results in assessment and knowledge levels are subject to the evaluation during the learning process. The evaluation, possibly with classification, and traditional knowledge verification (tests, identification of results) when the focus is put on increased mathematical literacy. One of the possibilities is the formative assessment, whose substance and impact on pupils' results in mathematics at the level II. (Key stage 3) of primary schools will be described in this article.

## 2. Theoretical Background

By screening the teaching process in mathematics we can measure its results, i.e. the degree of achievement. It involves testing and evaluation and

possibly classification of pupils. These two processes are closely related but not identical. By testing we evaluate the performance, but it is not the condition for evaluation, as evaluation can be done without tests. The screening of the teaching process consists of two basic activities:

- a) Evaluation of results during the teaching process, by which we mean screening and testing of pupils. The subject is performance evaluation coupled with quantitative values according to predetermined rules.
- b) Evaluation of the results of the teaching process is regarded as the evaluation of pupils. Its essence lies in the comparison of particular pupils' results with specified requirements, models, standards or their previous results.

In [6] we can read that the role of the teacher consists of two similar but not identical activities. They are based on monitoring and evaluating pupil's progress. The grade is the result of the teacher's check based on the comparison of the actual result and the existing benchmark. The teacher identifies the difference by using criteria in the Classification Regulations and Standards. The teacher evaluates the performance of the pupil according to how close the performance is to the general requirements and to what extent the performance meets the potential of a particular pupil.

Teacher's assessment, which also reflects his or her expectations, could be more arbitrary than the classification relying on objective target values (equal for all pupils). The problem lies in the level of objectivity and the veracity of the value judgment. According to [7], teacher's assessment is a systematic process evaluating the pupils, groups of pupils of educational program's quality and performance. This process is prepared, organized and carried out regularly and subject to revisions and corrections. The basic rule of school evaluation is that pupils are notified of the subject of assessment (standard, over-performance and underperformance criteria). The basic criteria for assessment is the target standard in the target category [8].

As mentioned before, evaluation and classification are not identical activities. By classification we understand grade assignment, classification of pupil's performance according to the performance groups by letters, points, percentage etc.

Evaluation of teaching outcomes involves diagnosing, measuring and comparing knowledge and skills of pupils acquired in tuition of subjects or topics at school. It does not need to be coupled with classification. For teaching mathematics in Slovakia an increasing importance is given to point evaluation. Individual performances are assigned a particular number of points and at the end of the classification

period a grade is given according to the number of points. Three main evaluation methods are applied:

- a) Comprehensive – which is the final evaluation in the mid school term and at the end of a school term, the end of a course etc.
- b) Diagnostic – the teacher selects a specific group of problems in order to identify the cause of their misunderstanding or inability to solve a task correctly [9].
- c) Formative – informal and formal assessment which should be made by teachers in every lesson in order to assess how many their pupils have learned, and how much work can be scheduled for every individual or a group in order to achieve their potential.

The primary objective of formative evaluation is to obtain a feedback. Therefore, remedying deficiencies and increasing efficiency are crucial. Information about the work, learning abilities of pupils is acquired and mistakes, misconceptions, problems and their causes are identified.

Formative assessment is therefore called testing against the requirements directly affecting social skills of pupils and their overall academic success. Learning as such is transformed into a creative process as a social phenomenon [10].

We have to recognize that the development of crucial skills in pupils is closely linked with the forms of evaluation which we call formative [11].

Formative assessment is usually disconnected from the classification of pupils and is efficient only when information is actually used for addressing the gaps in pupil's/student's work [12].

### 3. Research Methodology

The main objective of the research was to identify and evaluate the impact of formative evaluation on the level of results achieved in mathematics, with the emphasis put on comparing the results in mathematics between groups of pupils of the level II. Primary schools whose teaching process was enhanced by formative assessment and a group which did not have it.

The research sample consisted of students in the ninth grade of primary school and the experiment was carried out in two primary schools in Nitra (Slovakia), which we call A and B. Respondents were put into two equally performing classes of 9 - graders with the same teacher in both groups.

Both schools A and B had teachers of mathematics with the professional experience of 26 – 32 years who agreed not to mark their pupils in this experiment and did not influence the results of this experiment in any substantial way.

Each school selected two classes, one of which was experimental and the other was a control group (class). The primary school A was in the town center, but also attended by pupils from Nitra’s surrounding (about 10 km from the town center). Both classes in the school A had an equal number of pupils i.e. 20. The primary school B was located in a suburban settlement, the pupils originated for the catchment area of the school. The class with 18 pupils was defined as experimental and the class with 22 pupils was the control group.

**3.1. Research Materials and Methods**

From our practical experience we know that higher quality and more lasting results in pupils are achieved when formative assessment is applied. We established research hypothesis:

– Pupils in senior classes thought in standard conditions, whose interim results in mathematics were assessed formatively, on average perform better than those pupils who are evaluated in the standard way.

The following hypotheses were established for the statistical evaluation:

- We presume there is no statistically relevant difference in the pre-test results between the experimental and the control group. ( $H_{01}$ );
- We presume there is no statistically relevant difference in the post-test results between the experimental and the control group, ( $H_{02}$ ).

In the empirical part of the research we provided training for teachers of mathematics in selected classes in order to make them understand the problem of formative evaluation. The training was verbal, we explained the problem, briefed them on the planned course of the experiment and their role in the whole process. The length of the experiment was limited by the time period allocated for two thematic units – Functions. The linear function, the volume and the surface of bodies accounted 44 lessons in total.

**4. Analysis of Results**

Input knowledge and skills of respondents were evaluated in a pre-test. We accepted the results of Testing 9, the official set of national tests in mathematics applied in Slovak schools. In the light of their expertise and implementation method they can be used as a benchmark for selecting equal groups.

We underline that in the framework of Testing 9, the reliability of these tests is at least 0,80, and the NR test is applied (discriminatory, comparative and

relative performance tests). This is not a standard school test [13].

The results of respondents in Testing 9 are contained in Table 1. These data were provided by the teachers upon the agreement of the parents on condition that the personal data of pupils remain anonymous.

*Table 1. Results of pupils in the pre-test*

Order of pupils	Primary school 1		Primary school 2	
	Control group	Experimental group	Control group	Experimental group
1.	15,00%	55,00%	40,00%	85,00%
2.	65,00%	35,00%	35,00%	60,00%
3.	50,00%	55,00%	25,00%	30,00%
4.	55,00%	95,00%	55,00%	60,00%
5.	60,00%	55,00%	70,00%	15,00%
6.	60,00%	80,00%	25,00%	10,00%
7.	80,00%	80,00%	70,00%	80,00%
8.	55,00%	90,00%	10,00%	15,00%
9.	85,00%	65,00%	45,00%	30,00%
10.	80,00%	85,00%	25,00%	90,00%
11.	90,00%	90,00%	35,00%	15,00%
12.	65,00%	40,00%	50,00%	95,00%
13.	65,00%	80,00%	50,00%	70,00%
14.	65,00%	90,00%	25,00%	45,00%
15.	60,00%	35,00%	20,00%	80,00%
16.	50,00%	80,00%	5,00%	15,00%
17.	80,00%	75,00%	50,00%	55,00%
18.	55,00%	65,00%	70,00%	35,00%
19.	50,00%	60,00%	50,00%	
20.	35,00%	70,00%	65,00%	
21.			5,00%	
22.			35,00%	
	<b>61,00%</b>	<b>69,00%</b>	<b>39,09%</b>	<b>49,17%</b>

Twelve teaching materials were prepared for the experiment and their content was in line with education standards and requirements for school leavers [14]. Each material contained seven tasks, which were used as formative evaluation tools.

The experiment was implemented in two parallel classes in two schools and the teaching process was carried out as follows:

- a) The new curriculum was taught by teachers of mathematics in the usual manner and conditions. During that time experimenters prepared special teaching materials.
- b) Control groups were thought in the usual way, using standard mathematics textbooks, and collections of mathematic tasks and evaluated in the usual manner.
- c) The teaching process in the experimental groups contained extra teaching materials with formative evaluation, used in 3-5 lesson blocks according to the character and the level of difficulty of the curriculum.

Each two thematic units were evaluated. In every assessment the teacher also applied the summary evaluation method he/she was familiar with. Testing materials were given to the pupils from the experimental group in presence of experimenters, who evaluated their results. These results were marked directly into the teaching materials. Reasoning and arguments, errors, written evaluation were discussed with individual pupils at the following teaching lesson. Sample of formative assessment of the task No. 4 from the test No. 1 is shown in Figure 1.

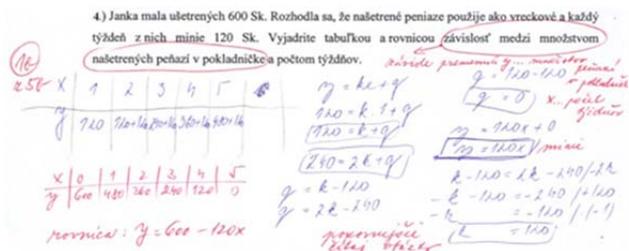


Figure 1. Sample of formative assessment of the one task from the pupils' test

Experimenters paid a particular attention to tasks with a so called “Find the mistake in the solution of the task!” type of request. These tasks are usually not part of textbooks or widely available in collections of tasks. Tasks of this kind are suitable for diagnosing informal knowledge in pupils [15]. The point is to identify a mistake in the presented result. The aim of the task is to develop argument skills, critical thinking, rise self-confidence and consolidate the knowledge by “learning from other people’s mistakes” (see in Figure 2).

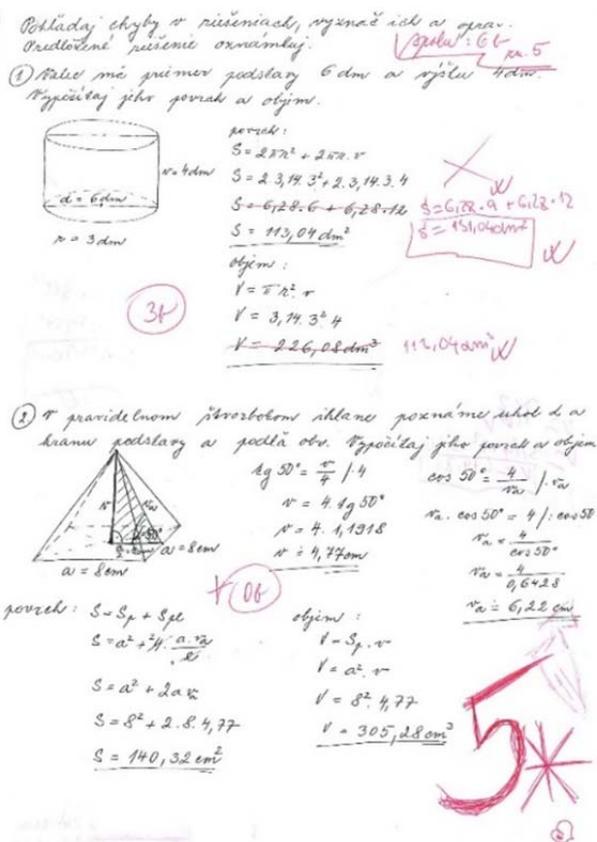


Figure 2. Sample of formative assessment of the task “Find the mistake ...”, from the pupils' test

At the end of the experiment a post-test was applied. It was a written test revision of the last three months' curriculum, which contained: calculations of the volume and surface of bodies, graphical solutions of two systems of linear equations and construction of a graph of a linear function. The results of the pupils in the post-test are presented first, so the results of the post-test are in Table 2.

Table 2. Results of pupils in the post-test

Order of pupils	Primary school 1		Primary school 2	
	Control group	Experimental group	Control group	Experimental group
1.	didn't write	didn't write	9,00 %	59,00 %
2.	54,00 %	11,00 %	77,00 %	64,00 %
3.	29,00 %	32,00 %	9,00 %	23,00 %
4.	39,00 %	96,00 %	95,00 %	95,00 %
5.	71,00 %	68,00 %	77,00 %	14,00 %
6.	61,00 %	43,00 %	23,00 %	36,00 %
7.	didn't write	7,00 %	73,00 %	82,00 %
8.	64,00 %	89,00 %	41,00 %	23,00 %
9.	79,00 %	75,00 %	100,00 %	18,00 %
10.	54,00 %	86,00 %	didn't write	100,00 %
11.	46,00 %	79,00 %	64,00 %	41,00 %
12.	39,00 %	50,00 %	41,00 %	100,00 %
13.	25,00 %	79,00 %	64,00 %	77,00 %
14.	71,00 %	100,00 %	36,00 %	77,00 %
15.	14,00 %	didn't write	didn't write	didn't write
16.	39,00 %	68,00 %	23,00	9,00 %
17.	32,00 %	54,00 %	23,00 %	91,00 %
18.	46,00 %	61,00 %	27,00 %	36,00 %
19.	54,00 %	100,00 %	36,00 %	
20.	14,00 %	79,00 %	82,00 %	
21.			5,00 %	
22.			55,00 %	
	<b>46,17 %</b>	<b>65,39 %</b>	<b>48,00 %</b>	<b>55,59 %</b>

5. Discussion

The following discussion is related to statistical evaluation of the pedagogical experiment. The results of pre-test (see Table 3) and post-test (see Table 4) in experimental and control groups were compiled in the following descriptive statistical tables. As an average we selected 95% confidence limit, i.e. the estimated average is <-95; 95>.

Table 3. Descriptive statistics (pre-test)

Effect	Descriptive Statistics							
	Level of Factor	N	Pre-test Mean	Std. Dev.	Std. Err.	Pre-test -95%	Pre-test 95%	
<b>total</b>		73	55,27	23,97	2,80	49,67	60,86	
<b>group</b>	control	38	51,05	20,70	3,35	44,24	57,85	
<b>group</b>	exp	35	59,85	26,63	4,50	50,70	69,00	
<b>school</b>	PS 1	36	67,08	16,22	2,70	61,59	72,57	
<b>school</b>	PS 2	37	43,78	24,87	4,08	35,49	52,07	
<b>group *school</b>	control PS 1	18	62,5	13,95	3,29	55,55	69,44	
<b>group *school</b>	control PS 2	20	40,75	20,60	4,60	31,10	50,39	
<b>group *school</b>	exp PS 1	18	71,66	17,40	4,10	63,01	80,32	
<b>group *school</b>	exp PS 2	17	47,35	29,37	7,12	32,24	62,45	

Table 4. Descriptive statistics (post-test)

Effect	Descriptive Statistics						
	Level of Factor	N	Post-test Mean	Post-test Std. Dev.	Post-test Std. Err.	Post-test -95%	Post-test 95%
<b>total</b>		73	53,60	28,16	3,29	47,03	60,17
<b>group</b>	control	38	47,13	24,80	4,02	38,97	55,28
<b>group</b>	exp	35	60,62	30,22	5,10	50,24	71,00
<b>school</b>	PS 1	36	55,77	25,56	4,26	47,12	64,42
<b>school</b>	PS 2	37	51,48	30,69	5,04	41,25	61,72
<b>group *school</b>	control PS 1	18	46,16	19,09	4,50	36,66	55,66
<b>group *school</b>	control PS 2	20	48	29,49	6,59	34,19	61,80
<b>group *school</b>	exp PS 1	18	65,38	28,01	6,60	51,45	79,32
<b>group *school</b>	exp PS 2	17	55,58	32,46	7,87	38,89	72,28

The result comparison in experimental and control groups shows that these two groups cannot be considered as equal (can see in Figure 3).

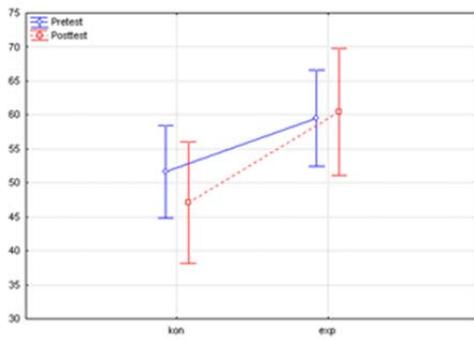


Figure 3. Results of the experimental and control group in the pre-test and post-test

We verified the following assumptions: variable is normally distributed in both groups; and equality of variances.

In the Figure 4 and Figure 5 there are graphs describing diagnosis predicted in the tests of pupils (pre-test and post-test). We can see obvious deviations from normality.

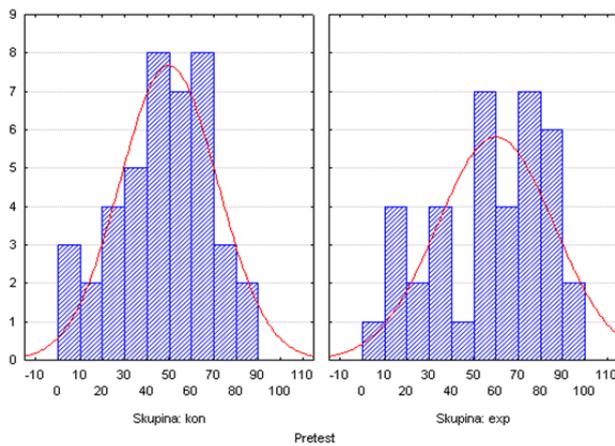


Figure 4. Diagnosis of predicted validity in the pre-test

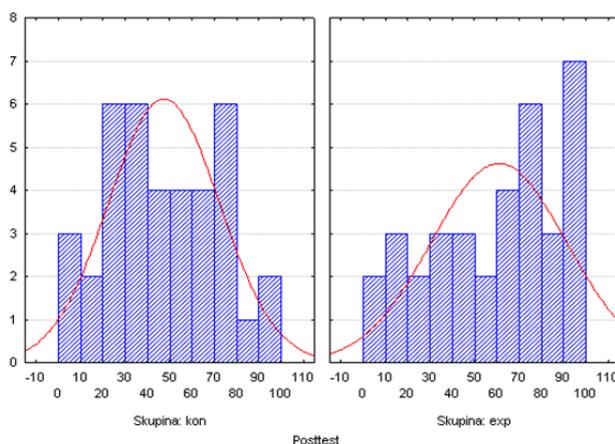


Figure 5. Diagnosis of predicted validity in the post-test

Now we present Table 5 and Table 6 in which the data are processed by the Levene's nonparametric test for homogeneity. This implies that the second assumption is not valid, i.e. the assumption of the same deviations for the parametric use test ( $p < 0,05$ ).

Table 5. Test for homogeneity of the experimental and control group

Control and experimental group	Levene's Test for Homogeneity of Variances			
	MS Effect	MS Error	F	p
Pre-test	742,0552	187,3728	3,960315	0,050437
Post-test	489,9648	197,1362	2,485413	0,119352

Table 6. Primary schools test for homogeneity

Primary school 1 and primary school 2	Levene's Test for Homogeneity of Variances			
	MS Effect	MS Error	F	p
Pre-test	937,2201	127,0455	7,37704	0,008291
Post-test	729,4559	187,9337	3,881454	0,052723

Therefore, we used a non-parametric Median test for homogeneity in statistical data analysis [16]. The dependent variable is the pre-test results and the independent variable is the group (see Table 7).

Table 7. Median test for homogeneity of the experimental and control groups in the pre-test

Dependent: Pretest Independent: group	Median Test, Overall Median = 55,000 Chi-Square = 3,136161 df = 1 p = 0,0766		
	control	exp	total
<= Median: observed	26	16	42
expected	22,05	19,95	
obs.- exp.	3,95	-3,95	
> Median: observed	16	22	38
expected	19,95	18,05	
obs.- exp.	-3,95	3,95	
Total: observed	42	38	80

Since  $p > 0,05$  applies, the hypothesis  $H_{01}$  about the equivalence of the sets cannot be unequivocally rejected. So we conclude that both groups are equal in the pre-test. This is also illustrated in the Figure 6.

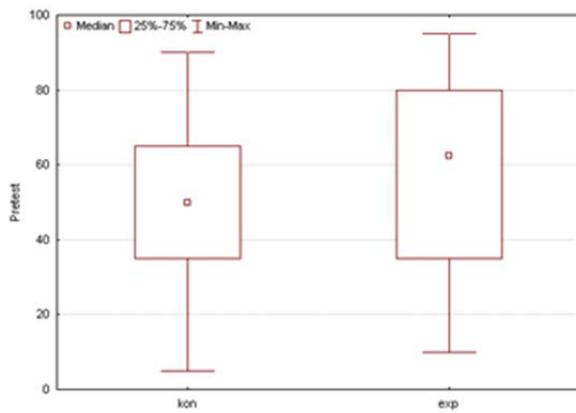


Figure 6. Median test for homogeneity of the experimental and control groups in the pre-test

We also used the median test for further processing. The dependent variable is the results of the post-test and the independent variable is the group (see Table 8 and Figure 7).

Table 8. Median test for homogeneity of the experimental and control groups in the post-test

Dependent: Posttest Independent: group	Median Test, Overall Median = 54,0000 Chi-Square = 3,914903 df = 1 p = 0,0479		
	control	exp	total
$\leq$ Median: observed	24	14	38
expected	19,78082	18,21918	
obs.-exp.	4,21918	-4,21918	
$>$ Median: observed	14	21	35
expected	18,21918	16,78082	
obs.-exp.	-4,21918	4,21918	
<b>Total: observed</b>	<b>38</b>	<b>35</b>	<b>73</b>

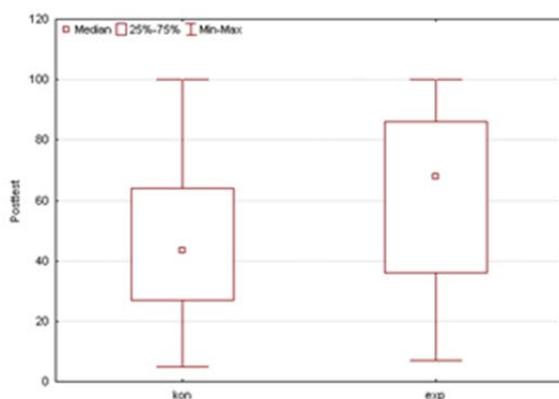


Figure 7. Median test for homogeneity of the experimental and control groups in the post-test

Since  $p < 0,05$ , we reject the hypothesis  $H_0$  about the equivalence of sets. Thus, the groups are not equivalent in the post-test results. This implies that there is a statistically significant difference in favour of the experimental group with 95% confidence.

## 6. Conclusions

The experiment confirmed that formative assessment positively impacts academic results of pupils in mathematics. The selection of the research sample – senior pupils in the lower secondary education cycle (Level II. of the primary school) was intentional. One of the reasons was Testing 9 – 2008, whose credible results were used as pre-test results in order to identify comparable experimental and control groups. The second, and even more relevant reason, was the educational impact of the formative assessment on this age group, whose members are more difficult to motivate to better performance in general.

We have to mention several negative external factors present in this experiment which complicated its implementation (external disturbances which led to disruptions, non-participation of individual pupils in classes etc.).

Based on the results obtained in this research we formulated significant recommendations for the educational theory and practice. Hierarchically, we could divide them into recommendations for the curriculum and for the personal development of pupils.

It would be suitable for the teachers of mathematics to include more activities developing mathematical literacy, reading with understanding, and continuous evaluation of pupil's work emphasizing the positive achievements. It is equally important in our view to stress positive aspects of the performance for underachievers and encourage them to improve their performance in future.

The experiment's main conclusion is that using diverse teaching materials in mathematics aimed at formative assessment that has a positive impact on the knowledge level of pupils. Integrating these materials is especially important in the lower secondary education. We suggest using them after every thematic unit is explained in line with the current legislation and training standards.

## Acknowledgements

*This work was partially supported by the Ministry of Education, Science, Research and Sport of the Slovak republic, name of the project "Inspirational Didactic Processes in Teaching of Projection Methods in Secondary Mathematics Education with a Focus on Requirements of Society and Practice", project number 019UKF-4/2020.*

## References

- [1]. Fulier, J., Vrábellová, M., Vrábello, P., Kmeťová, M., Rybanský, L., Lehocá, Z. N., ... & Csáky, A. (2014). Zvyšovanie matematických kompetencií žiakov nižšieho sekundárneho vzdelávania (ISCED 2). Univerzita Konštantína Filozofa v Nitre, Nitra, Slovakia. Retrieved from: [http://www.kega.fss.ukf.sk/Odborna\\_publ.pdf](http://www.kega.fss.ukf.sk/Odborna_publ.pdf) [accessed: 10 February 2020].
- [2]. Švec, Š. (2002). Základné pojmy v pedagogike av andragogike:(S anglicko-slovenským slovníkom termínov s definíciami as registrom ich slovensko-anglických ekvivalentov). Iris.
- [3]. Kalaš, R. I. (2010). Premeny školy v digitálnom veku. *DidInfo 2010*.
- [4]. Turek, Ivan. Škola a kvalita. Bratislava, 2015. ISBN 978-80-8168-221-6
- [5]. Hrmo, R., & Krpáľková-Krelová, K. (2010). Zvyšovanie kvality vyučovacieho procesu. *Bratislava: STU*.
- [6]. Rosa, V. (2007). Metodika tvorby didaktických testov. *ŠPÚ: Bratislava*.
- [7]. Pasch, M., Gardner, T. G., Sparks-Langerová, G., Starková, A. J., & Moodyová, C. D. (1998). *Od vzdelávacieho programu k vyučovaci hodine*. Portál.
- [8]. Kolář, Z., & Šikulová, R. (2005). *Hodnocení žáků: formy hodnocení: učitel a žák: sebehodnocení: praktické ukázky*. Grada Publishing.
- [9]. Jirotková, D., & Littler, G. (2003). Student's Concept of Infinity in the Context of a Simple Geometrical Construct. *International Group for the Psychology of Mathematics Education*, 3, 125-132.
- [10]. Prucha, J., Walterová, E., & MAREŠ, J. (2003). Pedagogický slovník. *Praha: Portál*, 322.
- [11]. Starý, K. (2006). *Sumativní a formativní hodnocení*. Retrieved from: <http://www.rvp.cz/clanek/457/992> [Accessed: 16 March 2020]
- [12]. OECD's Centre for Educational Research and Innovation (2005). *Formative Assessment: Improving Learning in Secondary Classrooms*. Retrieved from: <http://www.oecd.org/education/ceeri/35661078.pdf> [Accessed: 10 February 2020].
- [13]. NÚCEM (2019). *Testovanie 9*. Retrieved from: <https://www.nucem.sk/sk/merania/narodne-merania/testovanie-9/o-testovani-9> [Accessed: 16 May 2020].
- [14]. ŠPU (2015). *Inovovaný ŠVP pre druhý stupeň ZŠ*. Retrieved from: <https://www.statpedu.sk/sk/svp/inovovany-statny-vzdelavaci-program/> [accessed: 22 March 2020].
- [15]. Hejný, M., & Kuřina, F. (2009). *Dítě, škola a matematika—druhé, aktualizované vydání*. *Praha: Portál*.
- [16]. Chráška, M. (2007). *Metódy pedagogického výzkumu: Základy kvantitativního výzkumu*. Praha: Grada Publishing.