THE IMPLEMENTATION OF FORMATIVE ASSESSMENT INTO CHEMISTRY EDUCATION AT SECONDARY SCHOOL

Mária Babinčáková
Pavol Jozef Šafárik University in Košice, Slovakia
Charles University in Prague, Czech Republic
E-mail: maria.babincakova@upjs.sk

Mária Ganajová, Ivana Sotáková, Veronika Jurková
Pavol Jozef Šafárik University in Košice, Slovakia
E-mail: maria.ganajova@upjs.sk, ivana.sotakova@upjs.sk,
veronika.kopcova@student.upjs.sk

Abstract

The results of the implementation of formative assessment into chemistry education at secondary school for the topic "Mixtures" are presented here. Students (12-14 years old, N=202) were divided into two groups – control (N=97) and experimental (N=105). Teachers of experimental group implemented formative assessment tools into ten lessons (a predictive card, Frayer model, self-assessment card, T-card, concept map, and exit card). Control group teachers taught without formative assessment. The Mann-Whitney U test confirmed statistically significant results ($p<.05$).
Keywords: *formative assessment, secondary school, mixtures.*

Introduction

Children´s learning is the main aim of the school. Assessment plays an inseparable role in that process (Black, 1993). Teachers use summative and formative assessment at their lessons. Michael Scriven has first used the term 'formative evaluation' in connection with the curriculum and teaching (Scriven, 1967). Bloom, Hastings, and Madaus (1971) have mentioned 'another type of evaluation which all who are involved - student, teacher, curriculum maker - would welcome because they find it so useful in helping them improve what they wish to do', which they have termed 'formative evaluation'. Such assessment becomes formative assessment when the evidence is actually used to adapt the teaching to meet student needs (Black & Wiliam, 1998). Evidence of the impact of formative assessment on students´ learning is clear and considerable (Black & Wiliam, 1998). Many studies have been devoted to assessment for learning (Zeng, Huang, Yu, & Chen, 2018). On the other side, there have been critical views on this assessment, arguing that there has been a limited body of scientifically based empirical evidence (Dunn & Mulvenon, 2009). As has been summarized in Dunn and Mulvenon, 2009, researchers have reproached the small size of experimental groups, self-selection, study without a control group.

In Slovak schools, summative assessment has the main role in students´ assessing. As has been mentioned in OECD Reviews of Evaluation and Assessment in Education in

the Slovak Republic (Shewbridge, Van Bruggen, Nusche, & Wright, 2014), the idea of formative assessment is not well understood by teachers, students, and parents. Results have shown that the integration of the assessment process in a formative way (especially with inquiry) has been a difficult task (Bernard, Dudek-Różycki, & Orwat, 2019). It is probably due to the fact that teachers have difficulties with finding a balance between giving to students the leading role at the lesson and guiding them by questioning.

The main aim of this research was to find out whether there are significant differences in the results of students taught with formative assessment and students taught without formative assessment (summative assessment only).

Research Methodology

The sample consisted of 202 students from five Slovak secondary schools. Of the 202 participants, there were 75 (37.1%) young men, 80 (39.6%) young women, and 47 (23.3%) participants with missing gender information. The sample consisted of the 7th grade secondary school students. They were randomly assigned either to a control group or an experimental group. There were 105 (52%) participants in the experimental group and 97 (48%) participants in the control group.

Experimental group teachers implemented formative assessment tools into ten lessons (predictive card 3x, Frayer model 1x, self-assessment card 3x, T-card 2x, concept map 1x, and exit card 1x) into the topic "Mixtures". Teachers in the control group did not use formative assessment tools, and they taught and assessed without formative assessment. The teachers were provided with the types of tools that were applied in the evaluation process. Both groups were taught by the same teachers.

Cognitive test was given to both groups after ten lessons. The test was built in Bloom´s revised taxonomy and consisted of ten items - remember 2x, understanding 3x, apply 2x, analyze 3x (Table 1). These questions were created and evaluated by the National Institute for Certified Educational Measurements of Ministry of Education, Science, Research and Sport of Slovak Republic.

Table 1. The assignment of the test items to the Bloom´s learning domains.

Number of item	Learning Domains
1	Understand
2	Analyze
3	Analyze
4	Understand
5	Understand
6	Apply
7	Apply
8	Remember
9	Remember
10	Analyze

First, a descriptive analysis of the items was performed to calculate the mean, median and the standard deviation. A Kolmogorov Smirnov test was used in order to determine whether the sample available corresponded to a normal distribution, and then independent comparison tests were performed with a significance level of $p<.05$. The sequence of Mann-Whitney U tests was conducted to compare the results of the control and experimental group. The statistical analysis was performed in SPSS ver. 18 (SPSS Inc, 2009).

Research Results

The Mann-Whitney U test was performed for the entire test, groups of the items according to Bloom´s learning domains, and each of the items in the test separately. The results are presented in Table 2.

Table 2. Results of the Mann-Whitney *U* test.

		Experimental Group	Control Group	*p*-value				
		Mean	Median	Std. Deviation	Mean	Median	Std. Deviation	
Items separated	Item 1	.56	1.00	.499	.31	.00	.465	<.0001
	Item 2	.53	1.00	.501	.19	.00	.391	<.0001
	Item 3	.70	1.00	.458	.29	.00	.455	<.0001
	Item 4	.80	1.00	.402	.66	1.00	.476	.025
	Item 5	.87	1.00	.342	.80	1.00	.399	.231
	Item 6	.48	.00	.502	.20	.00	.399	<.0001
	Item 7	.30	.00	.458	.08	.00	.277	<.0001
	Item 8	.42	.00	.496	.43	.00	.498	.842
	Item 9	.59	1.00	.494	.14	.00	.353	<.0001
	Item 10	.05	.00	.214	.10	.00	.306	.134
Items grouped	Remember	.50	.50	.350	.29	.00	.321	<.0001
	Understand	.74	.67	.282	.59	.67	.283	<.0001
	Apply	.39	.50	.388	.14	.00	.304	<.0001
	Analyze	.43	.33	.284	.19	.00	.240	<.0001
Entire test	.53	.50	.208	.32	.31	.204	<.0001	

Results indicated the significant differences (p<.05) between the control and experimental groups in the entire test. Describing the individual items, results show highly significant differences for the usability score in all items except item 5 (p=.231), item 8 (p=.842), which are focused on lower order thinking skills according to Bloom´s order (remember and understand) and item 10 (p=.134). Highly significant results were in items aimed to remember (Z=-4.360, p<.0001), items aimed to understand (Z=-3.895, p<.0001), items aimed to apply (Z=-5.103, p<.0001), items aimed to analyze (Z=-5.785, p<.0001). Depending on the results, mean of the score of the experimental group was higher than the score of the control group in all items except item 8 (.42<.43), and item 10 (.05<.10), but it is visible that the difference was very small.

Conclusions and Implications

Because of the results of this research, it can be said that using formative assessment in chemistry education statistically improves students´ knowledge. Mann-Whitney U test finds the difference to be highly significant in items aimed to remember, items aimed to understand, items aimed to apply, and items aimed to analyze. These results are preliminary, and it is important to confirm them after a longer period and after multiple interventions. In depth analysis could provide us with the information, which elements of formative assessment have influenced the scores. This gives us a field of how we could improve the next steps.

Acknowledgments

This contribution was created in the framework of the National project „IT Academy – Education for the 21st Century“, which is supported by the European Social Fund and the European Regional Development Fund in the framework of the Operational Programme Human Resources, ITMS code of the project: 312011F057 and by the grant VEGA No. 1/0265/17 *“Formative Assessment in the Teaching of Natural Sciences, Mathematics and Informatics”*.

References

Bernard, P., Dudek-Różycki, K., & Orwat, K. (2019). Integration of inquiry- based instruction with formative assessment: The case of experienced chemistry teachers. *Journal of Baltic Science Education, 18*(2), 184–196.

Black, P. (1993). Formative and summative assessment by teachers. *Studies in Science Education, 21*(1), 49–97. https://doi.org/10.1080/03057269308560014.

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 118–119. https://doi.org/10.1002/hrm.

Bloom, B. S., Hastings, T., & Madaus, G. F. (1971). *Handbook on formative and summative evaluation of student learning.* New York: McGraw-Hill.

Dunn, K. E., & Mulvenon, S. W. (2009). A critical review of research on formative assessment: The limited scientific evidence of the impact of formative assessment in education. *Practical Assessment Research & Evaluation, 14*(7), 11.

Scriven, M. (1967). *The methodology of evaluation*. Washington, DC: American Educational Research Association.

Shewbridge, C., Van Bruggen, J., Nusche, D., & Wright, P. (2014). *OECD reviews of evaluation and assessment in education: Slovak Republic 2014*. Paris: OECD Publishing. https://doi.org/10.1787/9789264117044-en.

SPSS Inc. (2009). *PASW Statistics for Windows, Version 18.0*. Chicago.

Zeng, W., Huang, F., Yu, L., & Chen, S. (2018). Towards a learning-oriented assessment to improve students' learning - a critical review of literature. *Educational Assessment, Evaluation and Accountability, 30*(3), 211–250. https://doi.org/10.1007/s11092-018-9281-9.