



Co-funded by
the European Union



Report on the results of testing students' knowledge within the Out and about project

The aim of the research was to examine the impact of outdoor teaching on the quality of acquired knowledge of fourth and seventh grade students. The quality of knowledge in Science and Mathematics was measured on a sample of fourth-grade students, while the quality of knowledge was measured on a sample of seventh-grade students in the subjects of Physics, Chemistry, Biology, Geography and Mathematics. Based on the set goal, the following research tasks were defined:

1. To determine whether there is a statistically significant difference between the quality of knowledge of fourth grade students on the initial and final test of knowledge from Science at the level of recognition, reproduction and practical application;
2. Determine whether there is a statistically significant difference between the quality of knowledge of fourth-grade students on the initial and final test of knowledge in Mathematics at the level of recognition, reproduction and practical application;
3. To determine whether there is a statistically significant difference between the quality of knowledge of seventh grade students on the initial and final test of knowledge in Physics at the level of recognition, reproduction and practical application;
4. Determine whether there is a statistically significant difference between the quality of knowledge of seventh grade students on the initial and final test of knowledge in Chemistry at the level of recognition, reproduction and practical application;
5. Determine whether there is a statistically significant difference between the quality of knowledge of seventh grade students on the initial and final test of knowledge in Biology at the level of recognition, reproduction and practical application;
6. To determine whether there is a statistically significant difference between the quality of knowledge of seventh grade students on the initial and final test of knowledge in Geography at the level of recognition, reproduction and practical application;
7. To determine whether there is a statistically significant difference between the quality of knowledge of seventh grade students on the initial and final test of knowledge in Mathematics at the level of recognition, reproduction and practical application.

Sample research

The sample of respondents (N=69) consisted of 20 students of the fourth and 22 students of the seventh grade of the elementary school "Olga Milošević" in Smederevska Palanka (Serbia), then 16 students of the fourth grade from the elementary school _____ in Mersin (Turkey) and 11 students from Greece who attend the ROBOLEND center.

The content sample consisted of 40 teaching units that were implemented outside the classroom during the Out and about project activities. (Lesson models and contents of teaching units are given within the project result R1 - Out and About Treasure Chest - a collection of lesson plans _____ and _____ are _____ available _____ at _____ this _____ link <https://drive.google.com/file/d/1cRxKjardnJrHBji1sCu7UNPFNTSIVPIW/view?usp=sharing>)

Research flow

At the beginning of the research (October 2022), all students who made up the research sample took initial knowledge tests. Students of the fourth grade took an initial test in nature and society and mathematics, and students of the seventh grade took an initial test of knowledge in mathematics, biology, physics, chemistry and geography. After the initial testing, the students attended the classes that were held as part of the Out and about project activities during the 2022/2023 school year. in Serbia, Greece and Turkey. After the realization of all the hours outside the classroom that are foreseen by the project activities (a total of 40 hours outside the classroom), the final testing of the students was also carried out.

Research instruments

Initial and final knowledge tests were used to examine the quality of students' knowledge, which were created and validated as part of a pilot study conducted on a sample of students from Jagodina. The results of the pilot study and the determination of the metric characteristics of the instruments are described in detail in Appendix 11. Metric characteristics of the research instruments. As part of the mentioned pilot research, the reliability of their measuring scales was checked, and the values of the Cronbach's alpha coefficient (Cronbach's α) indicate that they are acceptable ($\alpha=0.8$ to $\alpha=0.9$), which indicates that the mentioned instruments are reliable for using and measuring the quality of knowledge from the previously mentioned subjects (All knowledge tests, in Serbian, English and Turkish, are given in attachments 1, 4, 5, 6, 7, 8, 9).

The following classification was used to determine the quality of knowledge, a dependent variable that was measured using knowledge tests (Mirkov, 1998). Level of recognition – The student is able to name and recognize the characteristics and elements of the content presented to him. Knowledge of facts, concepts, principles and generalizations that the student is only able to reproduce, without going deeper into their meanings; Level of understanding - The student understands the importance of facts, terms and concepts as part of a larger whole; observes dependencies between concepts; separates the essential from the unimportant; explains definitions,

laws and rules; connects facts and independently presents learned material. The student learns what is important, connects the learned material and concludes, understands cause-and-effect relationships, interprets various phenomena and events. The student gives learned examples of the principle or restates the principle, but in his own words. He is able to use learned principles and rules when he is presented with already known examples, but is unable to apply them to solving completely new and unknown problems; Level of application – Generalized knowledge, contains insight into essential relationships, understanding of general rules and principles. In order to be applicable to new situations, knowledge must be sufficiently general and free from the specific context in which it was first acquired. The student is able to apply the learned generalization to a specific situation, that is, when solving new and unknown problems. The student independently and independently uses and transfers the acquired knowledge in new situations and on unfamiliar material. Individually and critically approaches assigned problems, abstracts and generalizes them, comes to principles and legality, shows affection for research.

Based on the given classification of knowledge levels, the items in the research instruments were formulated based on the tasks from the international knowledge tests used in the PISA and TIMSS research.

Research results

After assessing the normality of the distribution of the results obtained on the final knowledge tests, using the Shapiro-Wilk test (Table 1), we come to the conclusion that for all the examined variables the level of statistical significance is lower than the set critical value ($\alpha < 0.05$). because the significance of the Shapiro-Wilk test is $r = 0.007$ for the final test from nature and society as a whole, and for sub-levels recognition $r = 0.000$, reproduction $r = 0.002$ and application $r = 0.014$, which indicates that the assumption of normality of distribution is not confirmed.

Table 1. *Examining the normality of the distribution of results on the initial and final test of knowledge from nature and society*

	Shapiro-Wilk		
	Statistic	df	Sig.
Total points on the recognition level	0.658	47	.000
Total points at the reproduction level	0.932	47	.002
Total points at the application level	0.950	47	.014
Total points on the final test	0.943	47	.007

When it comes to the mathematics test for the 4th grade (Table 2), we come to the conclusion that for all the examined variables the level of statistical significance is lower than the set critical value ($r < 0.05$) because the significance of the Shapiro-Wilk test is $r = 0.004$ for final test in mathematics as a whole, and for sub-levels recognition $r = 0.000$, reproduction $r = 0.001$ and application $r = 0.012$, which indicates that the assumption of normality of distribution is not confirmed.

Table 2. Examining the normality of the distribution of results on the initial and final test of knowledge in mathematics

	Shapiro-Wilk		
	Statistic	df	Sig.
Total points on the recognition level	0.581	47	.000
Total points at the reproduction level	0.734	47	.001
Total points at the application level	0.856	47	.012
Total points on the final test	0.723	47	.004

When it comes to the 7th grade (Table 3), we come to the conclusion that for all the examined variables the level of statistical significance is lower than the set critical value ($r < 0.05$) because the significance of the Shapiro-Wilk test is $r = 0.003$ for the final test in physics overall, and for sub-levels recognition $r = 0.000$, reproduction $r = 0.014$ and application $r = 0.016$, which indicates that the assumption of normality of distribution is not confirmed.

Table 3. Examining the normality of the distribution of results on the initial and final test of knowledge in physics

	Shapiro-Wilk		
	Statistic	df	Sig.
Total points on the recognition level	0.253	22	.000
Total points at the reproduction level	0.547	22	.014
Total points at the application level	0.761	22	.016
Total points on the final test	0.586	22	.003

We obtained similar data when it comes to chemistry and biology tests for the 7th grade (Tables 4 and 5), and we come to the conclusion that for all the examined variables the level of statistical significance is lower than the set critical value ($r < 0.05$), which indicates that the assumption of normality of distribution is not confirmed.

Table 4. Examining the normality of the distribution of results on the initial and final test of knowledge in chemistry

	Shapiro-Wilk		
	Statistic	df	Sig.
Total points on the recognition level	0.496	22	.000
Total points at the reproduction level	0.837	22	.000
Total points at the application level	0.622	22	.006
Total points on the final test	0.763	22	.005

Table 5. Examining the normality of the distribution of results on the initial and final test of knowledge in biology

	Shapiro-Wilk		
	Statistic	df	Sig.
Total points on the recognition level	0.844	22	.016
Total points at the reproduction level	0.547	22	.023
Total points at the application level	0.153	22	.019
Total points on the final test	0.788	22	.011

When it comes to the geography and mathematics tests for the 7th grade, we come to the conclusion that for all the examined variables the level of statistical significance is lower than the set critical value ($r < 0.05$) because the significance of the Shapiro-Wilk test is $r = 0.029$ for the final test from geography, and $r = 0.036$ for the mathematics test as a whole, and for the sub-levels recognition $r = 0.000$, reproduction $r = 0.042$ and application $r = 0.038$ for the geography test, and for the mathematics test the level of recognition $r = 0.027$, reproduction $r = 0.031$ and the application of $r = 0.014$, which indicates that the assumption of normality of distribution is not confirmed in this case either.

Due to the preliminary results of the normality test, the non-parametric Mann-Whitney U test was used in the analysis of achievement differences on the initial and final test in nature and society and all sublevels. The test results (Table 6) show that there is a statistically significant difference in the total number of points on the initial and final knowledge test. On the final knowledge test, a higher number of points is recorded compared to the results on the final test, and according to Cohen's criterion for the total number of points on the final test ($r = -5,093/\sqrt{61}$), the influence of the experimental factor is large and amounts to $r = 0.65$.

Table 6. The difference between the results on the initial and final test in nature and society and all sub-levels of knowledge

	ТЕСТ	Mdn	N	Mann–Whitney U	Wilcoxon W	Z	p
Укупно	Финални	49	47	112.500	608.500	-5.093	.000
	Иницијални	37	47				
Ниво препознавање	Финални	14	47	172.000	668.000	-4.700	.000
	Иницијални	13	47				
Ниво репродукције	Финални	15	47	233.000	729.000	-3.383	.001
	Иницијални	12	47				
Ниво примене	Финални	20	47	118.500	614.500	-5.018	.000
	Иницијални	14	47				

The results of the initial and final test in mathematics (Table 7) show that there is also a statistically significant difference in the total number of points on the initial and final test of knowledge. On the final test of knowledge, a higher number of points is recorded compared to the results on the final test, and according to Cohen's criterion for the total number of points on the final test, the influence of the experimental factor is large and amounts to $r=0.61$.

Table 7. The difference between the results on the initial and final test in mathematics and all sublevels of knowledge

	ТЕСТ	<u>Mdn</u>	N	Mann–Whitney U	Wilcoxon W	Z	p
Укупно	Финални	50	47	115.400	605.230	-5.183	.000
	Иницијални	39	47				
Ниво репродукције	Финални	15	47	183.000	671.000	-4.501	.000
	Иницијални	13	47				
Ниво разумевања	Финални	14	47	230.000	734.000	-3.405	.001
	Иницијални	11	47				
Ниво примене	Финални	19	47	116.820	616.810	-5.120	.000
	Иницијални	14	47				

By analyzing the results obtained by testing seventh grade students, we come to the conclusion that in the field of physics (Table 8.) there is a statistically significant difference in the results achieved by students on the initial and final knowledge test, in favor of the final test, which

practically means that the students on the final achieved better results in the physics test compared to the initial test.

Table 8. *The difference between the results on the initial and final test in physics and all sub-levels of knowledge*

Ман-Витнијев тест независних узорака Independent-Samples Mann/Whitney U Test	Значајност Sig.
Ниво препознавања	0.03
Ман-Витнијев тест независних узорака Independent-Samples Mann/Whitney U Test	Значајност Sig.
Ниво репродукције	0.042
Ман-Витнијев тест независних узорака Independent-Samples Mann/Whitney U Test	Значајност Sig.
Ниво практичне примене	0.0251

We got similar results when it came to the chemistry test (Table 9). The significance of the tests as a whole, as well as for each level of knowledge individually, indicates the fact that there is a statistically significant difference between the results of the initial and final knowledge tests in favor of the final knowledge test. Accordingly, it can be concluded that students on the final test in chemistry achieved better results on all levels of complexity, both at the level of recognition and at the level of reproduction and practical application.

Табела 9. *Разлика између резултата на иницијалном и финалном тесту из хемије и свим поднивоима знања*

Ман-Витнијев тест независних узорака Independent-Samples Mann/Whitney U Test	Значајност Sig.
Ниво препознавања	0.001
Ман-Витнијев тест независних узорака Independent-Samples Mann/Whitney U Test	Значајност Sig.
Ниво репродукција	0.007
Ман-Витнијев тест независних узорака Independent-Samples Mann/Whitney U Test	Значајност Sig.
Ниво практична примена	0.041

The Mann-Whitney test showed that in the context of biology content (Table 10), seventh grade students have better results on the final compared to the initial knowledge test, both on the test as a whole and within individual levels (overloading, reproduction and practical application). The students solved the tasks given on the final test more successfully, after they had undergone training and lessons outside the classroom within the Out and about project, compared to the initial test they did before the implementation of project activities related to lessons outside the classroom began.

Табела 10. *Разлика између резултата на иницијалном и финалном тесту из биологије и свим поднивоима знања*

Ман-Витнијев тест независних узорака Independent-Samples Mann/Whitney U Test	Значајност Sig.
Ниво препознавање	0.048
Ман-Витнијев тест независних узорака Independent-Samples Mann/Whitney U Test	Значајност Sig.
Ниво репродукција	0.013
Ман-Витнијев тест независних узорака Independent-Samples Mann/Whitney U Test	Значајност Sig.
Ниво практична примена	0.003

Similar results were shown by comparing the achievements on the initial and final geography test (Table 11).

Table 11. *The difference between the results on the initial and final geography test*

Mann-Whitney U	4577.000
Wilcoxon W	11363.000
Z	-4.210
Asymp. Sig. (2-tailed)	.000

And last, but not least, are the results of the seventh grade students on the mathematics test. Values $r=0.022$ for recognition level; $r=0.011$ for the level of reproduction and $r=0.021$ for the level of practical application, indicate the fact that there is a statistically significant difference in the achievements of students on the final test compared to the initial test. Therefore, the students achieved better results and a significantly higher number of points on the final test compared to the initial knowledge test.

Conclusions

Based on the statistical analysis and the results of the initial and final knowledge tests, as well as the individual answers of the students to the questions given in the research instruments, we can include the following:

The application of the learning model outside the classroom, which was developed within the Out and about project, had a significant impact on improving the quality of students' knowledge in the field of nature and society and mathematics at the age of fourth grade, as well as in the field of physics, chemistry, biology, geography and mathematics at the age of students seventh grade. Therefore, there is a statistically significant difference in student achievements on the initial and final knowledge test, in favor of the final test and at all levels of task complexity (recognition, reproduction and practical application).

After participating in project activities and in lessons outside the classroom, the students acquired better quality and practically changeable knowledge, i.e. more functional knowledge. The biggest shift occurred at the advanced level, which implies the practical application of knowledge in new and different situations.

The students who participated in the research acquired generalized knowledge and were more successful in seeing the essential relationships between the contents of living nature (the connection between plants, animals, humans and elements of non-living nature). Also, they achieved better results on tasks that required the application of knowledge in new situations when solving unknown problems. They cited new examples of food chains by creatively connecting previous knowledge about the way animals eat and the specifics of living communities. Through abstraction and hypothetical thinking, they successfully discerned cause-and-effect relationships between links in food chains, showing a penchant for research and citing new examples and explanations. They were more successful in tasks of orientation in space and time (both on the geographical map and on the settlement plan and timeline). Solved complex, text-based math problems with multiple operations.

Recommendations

The research was done on a small sample, but the program of lessons outside the classroom that the students went through was intensive and we assume that it had an impact on the students' achievements.

It would be useful to repeat the research with a larger sample of students in order to compare the results of this research.

It would be even more important to conduct an experiment with parallel groups, experimental and control, in order to obtain statistically significant results of experimental research in Serbia, Turkey and Greece.

There is certainly room for improvement of teaching by applying and implementing lessons outside the classroom.

The collection of preparations can be used by practitioners in both schools and teacher training colleges, as a model for the implementation of lessons, but also as examples for designing similar lessons in other locations and in different environments, not only in Serbia, Greece and Turkey, but also in other countries. Saint.