Abstract. The aim of this research is to establish the contribution of demonstration and student-led experiments on the quality of students’ knowledge about the air-related content in the third grade of primary school. This research included 120 students from Serbia, divided into two groups: control (C) and experimental (E). In group C the experiments were demonstrated by the teacher, whereas those same experiments were independently conducted by the students’ inside smaller groups in group E. The quality of the students’ knowledge was examined after the application of the experiments with the post-test, while the quality of the duration of knowledge was examined with the retest. The students in the group E acquired the knowledge of the higher quality and more durable knowledge about the air-related content then the students’ from the group C on the cognitive levels: analysis, evaluation and synthesis. Therefore, the priority should be given to the student-led experiments over the demonstration, when processed air-related content in the third grade of primary school, because their application contributes to the increase of the students’ knowledge quality.

Keywords: demonstration experiments, primary school, student knowledge quality, student-led experiments, third grade.

The Contribution of Demonstration and Student-led Experiments on the Students’ Knowledge Quality in the Third Grade of Primary School

Stanko Cvjetićanin, Mirjana Maričić

Introduction

Children start exploring from the earliest age, but first organized knowledge about nature, they learn in an integrated form during their preschool education, and then in primary school. In accordance with the educational system of the Republic of Serbia, students in the first four grades of primary school study natural sciences in an integrated form through mandatory subjects: The World around Us (first and second grade) and Nature and Society (third and fourth grade). They can also choose from optional subjects (intended to expand the knowledge) such as: Guardians of the Nature and Hands in the dough (taught in all four grades).

Students in the first four grades of primary school become familiar with studying nature through different methods and processes, but acquire scientific approach much better through direct practical experience (Eshach, 2006) when they: measure, touch and manipulate different materials, draw, make charts and write data in (Ateş & Eryılmaz, 2011, Karamustafaoğlu, 2011). Research activities have a central and distinct role in natural sciences teaching which can be neither efficient nor relevant without their use in the process (Katcha & Wushishi, 2015, Hart, Mulhall, Berry, Loughran & Gunstone, 2000). They usually imply the use of the problem approach and the usage of certain scientific skills, but are mostly based on conducting experiments (Koray & Serdar Köksal, 2009). Laboratory work (research activities) is carefully organized in teaching practice and presented to students through the laboratory-experimental method, combined with different teaching methods in the classroom and outside of it. Being subjected to it students learn independently (more than through other methods), look, conduct experiments and perform tasks which additionally activate them to participate in the teaching process (Cvjetićanin, 2009, Özmen, Demircioğlu, Burhan, Nasieriazar & Demircioğlu, 2012). This kind of activity, primarily based on noticing, gathering and analyzing data through experiments can provide students with a deeper understanding of nature and enable developing their problem-solving skills and research abilities. Furthermore it can teach them how to make generalizations about the most important scientific elements,
to acquire it and thus form corresponding scientific knowledge (Tamir, 1977). Over the last couple of years there have been a large number of experimental and observational research on the scientific laboratory work in schools as well as at universities (Hofstein & Lunetta, 1982, Hofstein & Lunetta, 2003). Research on their performance in subject teaching at primary, high schools and universities are especially notable. These research have mostly dealt with the contribution of laboratory work on students (students at universities and in schools) knowledge (McKee-Vickie, Williamson & Rubebush, 2007, Myers & Dyer, 2006); achievement; attitude towards nature (Huggerat, Najami, Abbasi & Dkeidek, 2014, Ogundiwon, Asaaju, Adegoke & Ojo, 2015) interest to study certain subjects and achieving better performance (Gendjova, 2007, Salameh El-Rabadi, 2013), i.e. (that is, in other words) having greater success in mastering specific content: “Reproduction, growth and development in living things” (Cardak, Onder & Dikmenli, 2007); effect on developing scientific and cognitive skills (Odebunmi & Balogun, 1991, Khan & Iqbal, 2011); eliminating misconceptions (Sert-Cibik & Diken, 2008); understanding certain content and forming alternative concepts (Ozmen, Demircioğlu, Burhan, Nasiriazar & Demircioğlu, 2012); increasing motivation; active thinking; finding a suitable way of work (Stavreva-Veselinovska, Koleva-Gudeva & Djokic, 2011); developing creative and logic thinking (Koray & Serdar Köksal, 2009). Results of listed works have indicated a positive contribution and application of laboratory work and conducting experiments, not only in courses, but also in classes in primary and high school. Authors Ogundiwon, Asaaju, Adegoke & Ojo, (2015) researched the effect of application of group research laboratory work on students’ performance in biology (high school) and the results showed that there is a significant statistical difference in students of the experimental group who learned through the aforementioned strategy. Salameh El-Rabadi (2013) conducted a research on the effect of laboratory experiments on students’ achievement in the 10th grade physics classes, and the results indicated that students of the experimental group, who have studied through experiments, had a greater success when compared to the control group. Apart from the mentioned research, it is also important to mention the contribution of those who dealt with the efficiency of hands-on activities and research based learning as key aspects of laboratory-experimental work. These research studied their influence on students’ achievement (Sadi & Cakiroglu, 2011); quality and duration of the acquired knowledge (Logar & Savec-Ferk, 2011); acquiring and developing scientific skills and attitudes towards science and specific subject content: biology, chemistry and physics (simple electric circuit) (Ateš & Eryilmaz, 2011, Ergül, Şimşekli, Calış, Özdolek, Göcmençelebi & Şanlı, 2011); change of misconceptions (Unal, 2008); students’ interest and activity (Holstermann, Grube & Bögeholz, 2010, Maxwell, Lambeth & Cox, 2015); conceptual understanding of the content and skill improvement (Şimşek & Kabapinar, 2010, Turpin & Cage, 2004). Results have shown a positive contribution of the application of hands-on activities and IBL on students’ achievement (Bilgin, 2006), but have also indicated that teachers/instructors can organize experiments, which include cheap and available materials in order to improve students’ skills (Hirça, 2013). Authors Sadi & Cakiroglu, (2011) researched the influence of hands-on activities on students’ achievement and their attitudes towards science, as well as their results in post-SAT test which have shown higher students’ achievement in members of the experimental group who have studied through hands-on activities, whereas there was no difference in students’ attitudes in both groups.

Research based on the application and contribution of laboratory-experimental work in the first four grades of primary school are considerably rare. Few research have studied the influence of conducting experiments on students’ understanding of the presented content (Cakici & Yavuz, 2010), efficiency and attitudes of students towards its application, as well as the interest and motivation of students to learn the subject in question (Golubović-Ilić, 2011), the influence on both quantity and quality of students’ knowledge (Cjjetčanin, Obadović & Rančić, 2015, Cjjetčanin, Segedinac & Halaši, 2010). The research in question proved positive contribution of application of the laboratory method, but have also laid the ground for further research related to this phenomenon (Cjjetčanin, Obadović & Rančić, 2015). If the focus is set to a single segment i.e. a single category in the experiment division, it can be noticed the lack of research on their separate contribution such as: the contribution of application of basic, comparative, model, long-term experiments etc. One of the most important experiment divisions for this research is the one based on who conducts the experiment. These experiments can be either demonstration (performed by a teacher, professor or a better trained student in front of the entire class) or student-led (performed by students: individually, in pairs or in smaller groups) (Cjjetčanin & Segedinac, 2007). The efficiency of student-led experiments compared to traditional (lecture) teaching methods was researched by Golubović-Ilić and Cjjetčanin, Segedinac & Halaši, while a comparative analysis of application of demonstration and student-led experiments was presented by Cjjetčanin, Obadović & Rančić. The research of Golubović-Ilić (2011) noted a positive contribution of student-led experiments on achievement, attitudes, interest and motivation of third grade students when compared to traditional teaching approach applied in the control group. The research of Cjjetčanin, Segedinac & Halaši, (2010)
studied the influence of student-led experiments on the quantity and quality of students' knowledge in fourth grade, and the results have shown better achievement for students of the experimental group when compared to the control group in which students learned through traditional methods. The research of Cvjetićanin, Obadović & Rančić, (2015) (which is of the highest relevance for this research) dealt with the efficiency of student-led and demonstration experiments on the quality of students' knowledge in the fourth grade and the results proved that the students’ from the experimental group displayed better knowledge than students from the control group who only saw the experiments demonstrated.

**Focus and Aim of Research**

Based on the analysis of the available research conducted on the contribution of the laboratory-experimental method in teaching natural sciences in the first four grades of primary school, it can be seen that there is a small number of comparative analysis done on the contribution of demonstration and student-led experiments on the quality of students' knowledge in realization of specific natural sciences content in different grades. All content is vertically connected through the first four grades of the primary school and becomes gradually more complex, which changes the aims and tasks of teaching. Students are taught air-related content in all four grades, but the matter does not become complex before third grade, so it is very suitable for this type of research. Reasons to do so are the following: students already have some knowledge of the air from previous grades and the complexity of the content is suitable for application, making it suitable for determining the contribution of the methods which are being assessed (laboratory-experimental method in this work).

The question is: Does the way of conducting experiment (by demonstration, or when students' perform them) contribute to the quality of students’ knowledge about air-related content in the third grade of primary school? This question follows the basic aim of the research which is: Establishing the contribution of demonstration and student-led experiments on the quality of students’ knowledge about the air-related content, in third grade of primary school. The aim is followed by the next goals:

1. Determining if there are statistically significant differences in the quality of knowledge in students who acquired air-related content through demonstration experiments when compared to those who acquired it through independently conducted experiments;
2. Analyzing and establishing the contribution of demonstration i.e. student-led experiments on the quality of acquired knowledge about the air-related content in different cognitive levels (knowledge, understanding, application, analysis, evaluation and synthesis);
3. Determining if there are statistically significant differences in quality of duration of students’ knowledge for the students who acquired air-related content through demonstration experiments when compared to those who acquired it through independently conducted experiments.

**Methodology of Research**

**General Background and Design**

Demonstration and student-led experiments are fully developed experimentation models in all methodology of teaching sciences (Cvjetićanin, 2009, De Zan, 2005) in the first four grades of primary school. Although this is a well-known fact they are rarely applied in practice, and even less scientifically assessed for their contribution and significance. For the needs of this research 16 experiments about air-related content were chosen (for four teaching units). The same 16 experiments were prepared for the both groups of students (C and E), but so that each group conducts it in a different way in order to determine their contribution. Their suitability was checked and confirmed by the teachers who were a part of the research, as well as by the experienced methodologists in the area of teaching sciences. The experiments' model in this research was tested through the experimental design as the best way for establishing their contribution.

An experimental research has been conducted. The main aim of the experimental research (research with an experimental design) is proving the cause-effect relation i.e. causal reasoning between the variables (Fraenkel & Wallen, 2003). This kind of research has an experimental factor (cause, independent variable) which is applied in the educational process in order to cause a change as a consequence of the experimental factor (dependent variable) (Knežević-Florić, Ninković, 2012). The research used the draft of the parallel-group experiment which implies
two ways (which are compared) realized at the same time with at least one experimental and one control group in order to establish their contribution. The research was conducted in a three and a half months' time span (fourteen weeks) of 2015/2016 (from the beginning of March to mid-June), and was realized through the following phases:

1. Preparatory phase: includes the analysis of pedagogical documentation in order to establish the GPA (grade point average) in the end of the previous grade, as well as their GPA in the subject *The World around Us* in the end of the second grade. Experiments which had been applied in realization of new air-related knowledge were identified in this phase. Items for the pre-test were chosen in the end of this phase, in order to establish the quality of the existing air-related knowledge. This phase was introduced three weeks before the realization of phase two and lasted for the entire period;

2. Phase of revising the previously acquired air-related knowledge (acquired by the students in the previous two grades). This phase lasted for a single school lesson and was realized after the pedagogical documentation had been reviewed;

3. Phase of measuring the initial state (quality of the existing air-related knowledge): includes measuring the dependent variable through a pre-test. This phase lasted for a single school lesson and was realized in the first lesson which followed revising previously acquired knowledge;

4. Group-creation phase (C and E): in this phase students were split into two groups which were made equal in the following criteria: GPA in the end of the second grade; GPA in the subject *The World around Us* in the end of the second grade; results (knowledge quality) acquired in pre-test and the number of students (C group n=60, E group n=60). This phase was realized after the pre-test had been finished, and it lasted for a week;

5. Phase of the experimental factor realization: introduction of the chosen treatments—experiments into the experimental (16 student-led experiments) and the control group (16 demonstration experiments) which are then compared. This phase lasted for two weeks (four school lessons). At each school lesson is treated one unit, and within each teaching unit were performed four experiments;

6. The newly-acquired knowledge revision phase (acquired by students through experiments). This phase lasted for a single school lesson and was realized in the very next class after the experimental factor ended. The teacher went through all new air-related content with students. Doing that, special attention was put to gradually going from one level to another (a higher one) in asking questions. After the revision was realized, items for the post-test and retest were chosen;

7. Phase of the measuring the final state (quality of the newly acquired air-related knowledge): which includes measuring the dependent variable through a post-test. This phase was realized in the class (lasting for a single school lesson) which immediately followed the lesson in which the newly acquired air-related knowledge was revised;

8. Phase of the measuring the quality of knowledge duration: which included measuring the dependent variable through a retention test – retest. This phase was realized a month after the post-test and lasted for a single school lesson;

9. Final phase: comparing the state of the dependent variable (quality of students' knowledge) based on the results of the pre-test, post-test and the retest. This phase was realized after the retest was realized, and it lasted for a week.

The basic aim of the experimental design is to establish the difference between the state before introducing the experimental factor (initial knowledge – in this case results of the pre-test) and the state following its application (final state – in this case results of the post-test and retest). The difference (between the knowledge quality before and after introducing the treatment) is seen as the effect of the experimental factor (in this case the contribution of the treatment i.e. demonstration and student-led experiments) (Knežević-Florić, Ninković, 2003).

The independent variable was: application of experiments (demonstration and student-led) in processing air-related content in the third grade of primary school. The dependent variable was: the third grade students' quality knowledge about air-related content on the six cognitive levels: knowledge, understanding, application, analysis, evaluation and synthesis.

**Sample**

Participants of the research were 120 third grade students from six classes of two primary schools in the area of the Autonomous Province of Vojvodina (Republic of Serbia). The students were split into two groups: C (group in
which experiments were demonstrated by the teacher) and E (group in which students independently conducted experiments based on written instructions given by the teacher).

Instrument and Procedures

After the data of the students’ GPA in the end of the second grade was gathered through the analysis of the pedagogical documentation (in group C it is 4.03 and in group E it is 4.08), and the students’ GPA in the subject *The World around Us* in the end of the second grade (in group C it is 4.27 and in group E it is 4.15), it was started with choosing items for further measuring. This research used pre-test, post-test and retest.

**Pre-test**

The pre-test is a combination of items from four sources, and was comprised of 12 items in total. Items were taken from the authors: Kukić & Acimović (2016), Stokanović & Lukić (2016), Tadijin (2006) and Životić (2016). Items were designed to measure the quality of air-related knowledge acquired in previous grades at all cognitive levels. An example of items at the levels of understanding and synthesis in the pre-test is given below (Figure 1).

**LEVEL OF UNDERSTANDING**

1: Wind is the movement of warm and cold air.

**LEVEL OF SYNTHESIS**

2: Look closely at the picture and answer the question.

What should you do so the bee could survive in the jar?

**Post-test**

The post-test is a combination of items from six sources and was comprised of 12 items in total. Items were taken from the authors: Andjelić, Erić & Vićentijević (2010), Blagdanić, Jović, Kovačević & Petrović (2016), Marinković & Marković (2011), Matanović, Vlahović, Joksimović & Djurdjević (2015), Munitlak, Šiki-Erski & Holond (2016) and Ralić-Žeželj (2016). Items were designed to measure the quality of the newly acquired air-related knowledge at all
cognitive levels. An example of items at the levels of understanding and synthesis in the post-test is given below (Figure 2).

LEVEL OF UNDERSTANDING
1. Look at the picture on which are shown the experiments. What do we prove with these experiments. Circle the letter in front of the correct answers.

- a) The air is all around us
- b) The water is good solvent
- c) The air has the shape of the space in which it is located

- a) The air expands on heating
- b) The air is moving in nature
- c) The air is all around us

LEVEL OF SYNTHESIS
2. Why bird and fox in the winter is not cold?

______________________________________     __________________________________________

______________________________________     __________________________________________

Figure 2: An example of items at the levels of understanding and synthesis in the post-test (Andjelić, Erić & Vićentijević, 2010, Marinković & Marković, 2011).

The retest is a combination of items from six sources and was comprised of 12 items in total. Items were taken from the authors: Andjelić, Erić & Vićentijević (2010), Blagdanić, Jović, Kovačević & Petrović (2016), Marinković & Marković (2011), Matanović, Vlahović, Joksimović & Djurdjević (2015), Munitlak, Šikl-Erski & Holond (2016) and Ralić-Žeželj (2016). Items were designed to measure the quality of duration of the newly acquired air-related knowledge at all cognitive levels. Based on the structure and content these items were similar to those given in the post-test, they measured the same newly acquired air-related knowledge through items which were only formulated in a different way. An example of items at the levels of understanding and synthesis in the post-test is given below (Figure 3).
LEVEL OF UNDERSTANDING

1. Look at the picture on which are shown the experiments. What do we prove with this experiments. Circle the letter in front of the correct answers.

a) The air expands on heating
b) The air takes up space
c) The air is dissolved in the water

LEVEL OF SYNTHESIS

2. Is it cold for eskimos in the igloo?

Why?

_____________________________________________________________________________________________
_____________________________________________________________________________________________
_____________________________________________________________________________________________
_____________________________________________________________________________________________
_____________________________________________________________________________________________
_____________________________________________________________________________________________
_____________________________________________________________________________________________
_____________________________________________________________________________________________
_____________________________________________________________________________________________
_____________________________________________________________________________________________
_____________________________________________________________________________________________
_____________________________________________________________________________________________

Figure 3: An example of items at the levels of understanding and synthesis in the retest (Andjelić, Erić & Vićentijević, 2010, Marinković & Marković, 2011).

All Tests

The items in tests were created based on the Anderson, Krathwohl & Bloom taxonomy which is a revised Bloom taxonomy (Anderson, Krathwohl & Bloom, 2001). In choosing items (from the previously mentioned sources) which would follow all cognitive levels, it was used the collection Smart tests: Teacher-made tests that help students learn (Walker, Schmidt, 2004), which gives the basics of each cognitive level with item examples to suit them. Validity of the tests was ensured by assessment, discussion and approval of several teachers with considerable experience (over ten years), as well as assessment, discussion and approval of several experts on teaching natural sciences, as experts in this area. Reliability of the test was calculated using the Cronbach’s alpha coefficient (alpha coefficient) which was .86 for all tests. According to this coefficient, all values which are above .70 are acceptable and considered reliable. In all tests (pre-test, post-test and retest) the quality of students’ knowledge was measured through two items for each cognitive level, with the items different in structure. In valuation of the items the applied rule was for items of higher cognitive level to bear more points. At the cognitive level of knowledge the students were supposed to recall new air-related knowledge, to detect and reproduce the information, idea and the principle similarly to what they learned. In items at the level of understanding the students were asked to notice and connect
main properties of air, understand the most important ideas, explain and interpret the learnt content and different air-related processes (such as movement and pressure), as well as to make logical conclusions based on available information on air i.e. to spot causal relations in air behavior (such as the influence of warmth on air weight). The items at the level of application were constructed to make students solve problems by using the newly acquired air-related content in learning context or in a new situation, followed by application of knowledge about air as an isolator, the importance of clear air for living beings, protection of air from pollution etc. At the level of analysis students were demanded to differentiate between important and insignificant information on the impact external factors have on air (on its movement, pressure, weight...), parse the given information to establish the parts on all acquired air-related knowledge, causal relations, causes and effects of different factors on air and its influence on living beings. At the level of evaluation students were supposed to compare and discover similarities and differences between air, oxygen and CO₂, as well as the differences between gas and liquid matter of state i.e. gas and solid matter of state. They were supposed to predict air behavior based on images showing different seasons, to discover and explain how living beings would act in given conditions. In items on the synthesis level students were supposed to creatively, or divergently use the acquired air-related knowledge to create a new concept on mutual relationship of air, water and soil. They were also asked to make generalizations based on information of the influence of different substances (materials) on air quality and how it would affect beings, environment etc.

Treatment

The chosen content for establishing the contribution of the way of conducting the experiments in the third grade of primary school on the quality of students' knowledge was Gasses as a matter of state-air. Content realized in both groups C and E is comprised of the following teaching units:
1. Air pressure and movement (changes in state and weight);
2. Changes which occur due to heating and cooling of the air (changes in temperature, weight, movement...);
3. Air as an isolator;
4. The properties of air that are important for the living world and human activity.

The Treatment in the Control Group

The teacher in group C gave explanations to students before each experiment: students were informed which equipment and materials were necessary for the experiment, how the experiment was going to be conducted as well as how they should observe. This involved frontal instructions. After the experiment was demonstrated and conclusions reached, students wrote down the process of the experiment in a separate notebook (experiments notebook). The teacher wrote important conclusions on the board during the discussion so students can write them down in their notebooks in the end.

The Treatment in the Experimental Group

In group E the teacher made heterogeneous groups of three students which differed in the initial knowledge: good, average and bad. Members of the group were the same in all experiments. Before each lesson in which the experiments were to be conducted the teacher would put the required equipment and materials on every desk. Each member of the group would get a paper with experiment instructions. Students in group E were asked to carefully read the instructions before the experiment, and repeat aloud to each other which tasks have to be conducted and how to conduct the experiment. The paper had the information about the needed equipment and material, how the experiment has to be conducted and the questions which have to be answered based on the results. After the teacher makes sure that all students know what has to be done, the students proceed to conducting experiments. All groups conducted experiments of same complexity. The level of complexity increased with each lesson. Inside their groups, students tried to answer the questions based on the results they reached. After groups finish, a student from each group presents the results in front of the class. When the presentation is done, the teacher encourages students to discuss, correct the wrong conclusions and make the right ones. Students check if they wrote down the correct procedure into their notebooks and if their conclusions were correct. After that, they rewrite into their notebooks conclusions written on the board, reached as the joint effort of all groups.

An example of experiments (for both groups) for some of teaching units is given below (Figure 4).
Experiment for the unit: Changes which occur due to heating and cooling of the air

**Necessary materials and supplies:**
- bowl with water
- slot
- glass bottle
- balloon

**Description of the experiment:**
Take the bowl with water and put it on the heated slot. When the water is heated, put in the glass bottle, on which outfall is pulled deflated balloon. Observe what happens with the balloon.

**Think and answer:**
- Why the balloon inflated?
- Based on the knowledge about the behavior of solids to heat, try to explain the behavior of air.
- Provide examples of everyday life, where you can see the phenomenon described.
- Note down the answers in your experiment notebook.

Experiment for the unit: The properties of air that are important for the living world and human activity

**Necessary materials and supplies**
- two candles
- glass cup
- matches

**Description of the experiment:**
Put on the table two candles and lit them up, using the matches. After that, one of the candles cover with the glass cup. Now inhale and hold your breath for a while. What do you feel? Can you stay long in this condition?

**Think and answer:**
- When the candle has been lit, which process has happened?
- When the candle has been covered, what has happened and why?
- What is needed from the air so the candle could burn?
- Why is air important for all living beings?
- Note down the answers in your experiment notebook.

Figure 4: An example of experiments for some of the teaching units.

---

Data Analysis

The acquired data were used for:

1. Balancing the groups: analyzing the GPA in the end of the second grade, students’ GPA in the subject *The World around Us* in the end of the second grade, and results (knowledge quality) of the pre-test in order to establish:
   - If there is a statistically significant difference between the students of groups C and E in the GPA in the end of the second grade and their GPA in the subject *The World around Us* in the end of the second grade, by establishing the following values: arithmetic mean (M (X)), standard deviation (SD) and the independent t-test;
   - Spearman's correlation coefficient was applied to establish the level of correlation between the students’ GPA in the end of the second grade and their GPA in the subject *The World around Us* in the end of the second grade and the results of the pre-test. The quality of students' knowledge about air was analyzed from two aspects – marks and the scored number of points in pre-test;

2. Analysis of the quality of newly acquired air-related content displayed by the students in the post-test in order to establish:
   - The contribution of demonstration and student-led experiments on the quality of students’ knowledge i.e. the difference in students’ knowledge between groups C and E, after the experimental part of the
research was analyzed by calculating and comparing the GPA (of marks) and number of scored points in post-test (for all cognitive levels). If there is a statistically significant difference in results i.e. quality of students’ knowledge in groups C and E in the post-test, was acquired through the non-parametric Mann-Whitney test and the independent t-test, whereas the Kolmogorov–Smirnov normality test was used to establish if the acquired data correspond to the normal distribution;

3. Analysis of the quality of duration of the newly acquired air-related content displayed by the students in the retest in order to establish:
   • The contribution of demonstration and student-led experiments on the quality of duration of students’ knowledge i.e. the difference in students’ knowledge between groups C and E after the experimental part of the research was analyzed by calculating and comparing the GPA (of marks) and the number of scored points in the retest (for all cognitive levels). If there is a statistically significant difference in results i.e. quality of duration of students’ knowledge in groups C and E in the retest, was acquired through the non-parametric Mann-Whitney test and the independent t-test, whereas the Kolmogorov–Smirnov normality test was used to establish if the acquired data correspond to the normal distribution;

4. Comparative analysis of the quality of students’ knowledge about air-related content (all tests):
   • One-way ANOVA was used to establish if there is statistically significant difference in the quality of students’ knowledge in C and E group at cognitive levels in pre-test and post-test as well as on the retest.

**Results of Research**

**Balancing the Groups**

Based on the values of the t-test it was established that there was no statistical significance in GPA difference between the groups C and E in the end of the second grade (t=.598, p=.613). Furthermore, there was no statistically significant difference in students’ GPA in the subject *The World around Us* in the end of the second grade (t=.603, p=.592). The third parameter for balancing groups was the quality of students’ knowledge in pre-test of both groups (marks and the number of scored points).

The quality of students’ knowledge in C and E group on the pre-test at different cognitive levels is shown in the table (Table 1).

<table>
<thead>
<tr>
<th>Cognitive level</th>
<th>Group</th>
<th>M (X̄)</th>
<th>SD</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge</td>
<td>E</td>
<td>3.4532</td>
<td>1.787</td>
<td>.6621</td>
<td>.4130</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>3.633</td>
<td>1.705</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Understanding</td>
<td>E</td>
<td>4.496</td>
<td>.537</td>
<td>.7050</td>
<td>.4770</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>4.414</td>
<td>.496</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Application</td>
<td>E</td>
<td>5.338</td>
<td>1.151</td>
<td>.9230</td>
<td>.5930</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>5.903</td>
<td>1.097</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analysis</td>
<td>E</td>
<td>8.322</td>
<td>2.216</td>
<td>.5850</td>
<td>.5650</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>8.407</td>
<td>2.086</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evaluation</td>
<td>E</td>
<td>0.947</td>
<td>1.421</td>
<td>.9110</td>
<td>.9030</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>0.901</td>
<td>1.381</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Synthesis</td>
<td>E</td>
<td>0.383</td>
<td>2.783</td>
<td>.4130</td>
<td>.6110</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>0.414</td>
<td>1.951</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The results acquired in the pre-test show that the students in groups C and E share similar knowledge about air-related content in previous grades in all cognitive levels (t-test significance is above .05 for every cognitive level) (Table 1). Neither C nor E group students solved the syntheses item correctly. Bad results were also present at both analysis and evaluation level.

There should be correlation between the GPA of students in the end of second grade, GPA in the subject *The World around Us* in the end of second grade and the quality of their knowledge on the pre-test. Correlation level was established based on value of the Spearman's correlation coefficient. Results have shown a moderate connection between the students' GPA in the end of the second grade and marks they got in the pre-test ($\rho=497$ with $p<.001$). The same situation emerges in correlation with the total number of scored points in testing ($\rho=528$ with $p<0.01$). Similar results appeared in establishing correlation between the students’ GPA in the subject *The World around Us* in the end of the second grade and marks they got in the pre-test ($\rho=512$ with $p<.001$), as well as between their GPA in the subject *The World around Us* in the end of the second grade and the total number of points they scored in the pre-test ($\rho=518$ with $p<.001$). Based on the intensity of the correlation coefficient of researched phenomena (variables), it can be spot moderate correlation which was unexpected (correlation intensity significantly lower than expected).

**Analysis of the Quality of the Students’ New Knowledge of the Air-Related Content**

The contribution of the way of conducting an experiment to the quality of students' knowledge about air in C and E group on the post-test at different cognitive levels is shown in the table (Table 2).

<table>
<thead>
<tr>
<th>Cognitive level</th>
<th>Group</th>
<th>M (X)</th>
<th>SD</th>
<th>CV (%)</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge</td>
<td>E</td>
<td>2.889</td>
<td>.86</td>
<td>20.218</td>
<td>.7970</td>
<td>.0730</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>2.668</td>
<td>.94</td>
<td>21.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Understanding</td>
<td>E</td>
<td>3.792</td>
<td>.98</td>
<td>21.22</td>
<td>.8550</td>
<td>.4030</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>3.688</td>
<td>1.31</td>
<td>23.62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Application</td>
<td>E</td>
<td>5.087</td>
<td>6.102</td>
<td>18.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>4.581</td>
<td>5.755</td>
<td>19.56</td>
<td>.6810</td>
<td>.3870</td>
</tr>
<tr>
<td>Analysis</td>
<td>E</td>
<td>9.225</td>
<td>1.88</td>
<td>5.23</td>
<td>.9060</td>
<td>.0630</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>5.141</td>
<td>2.09</td>
<td>6.38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evaluation</td>
<td>E</td>
<td>5.344</td>
<td>1.64</td>
<td>4.98</td>
<td>4.976</td>
<td>.0001</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>2.298</td>
<td>7.23</td>
<td>19.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Synthesis</td>
<td>E</td>
<td>3.771</td>
<td>2.14</td>
<td>4.62</td>
<td>8.865</td>
<td>.0001</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>0.987</td>
<td>9.12</td>
<td>15.97</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The results obtained from post-test show that the students from the E group performed better, meaning they acquired the knowledge of higher quality when compared to the students in the C group.

After analyzing the quality of students' knowledge of both groups at some cognitive levels it is possible to say that C and E group students shared similar results in: knowledge ($t=.7970$, $p=.0730$), understanding ($t=.8550$, $p=.4030$), application ($t=.6810$, $p=.3870$) and analysis ($t=.9060$, $p=.0630$). The E group students (Table 2) performed better than C group students in evaluation ($t=4.976$, $p=.0001$) and synthesis ($t=8.865$, $p=.0001$). If this is compared to the number of students in C group who successfully completed pre-test and post-test items at the level of analysis, it can be seen that this number is higher at the post-test. On the pre-test, none of C group students managed...
to successfully complete both items at the analysis level and only 8.33% have completed only one, while 34.77% completed both items on the post-test and 43.42% completed only one. The progress of C group at the evaluation level is higher than on the pre-test. Although none of C group students completed both items at the evaluation level, 26.64% completed only one. At the synthesis level 6.66% did only one and the rest did not do anything at this level. The E group outperformed C group in solving both items with 53.42% and only one with 41.76% at the evaluation level. Their success was by far the best at the level of synthesis. Although only 13.33% managed to complete both items, 39.84% did only one.

In order to determine if the existing differences (in favor of students from E group) are statistically significant, it’s tested the hypothesis of the normal data distribution.

The obtained values of the Kolmogorov-Smirnov test of normality of the students in the C and E group at the post-test and the retest are shown in the table (Table 3).

### Table 3. Results of students C and E group on the Kolmogorov–Smirnov normality test.

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Kolmogorov-Smirnov Z</th>
<th>Asymp. Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Post-test</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>60</td>
<td>.9020</td>
<td>.0001</td>
</tr>
<tr>
<td>C</td>
<td>60</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Retest</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>60</td>
<td>.8740</td>
<td>.0001</td>
</tr>
<tr>
<td>C</td>
<td>60</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Based on the obtained value of the Kolmogorov–Smirnov normality test (Z=.9020, p=.0001) it was noted that data does not have normal distribution and that statistical significance of the difference (Table 3) between students in groups C and E in post-test should be confirmed through the Mann-Whitney non-parametric test.

Statistical significance of difference in the quality of students’ knowledge between groups C and E in post-test is acquired through the Mann-Whitney non-parametric test, and it is shown in the table (Table 4).

### Table 4. Difference in the quality of students’ knowledge in C and E group (post-test) on the Mann-Whitney non-parametric test.

<table>
<thead>
<tr>
<th>The total number of points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mann-Whitney U</td>
</tr>
<tr>
<td>Wilcoxon W</td>
</tr>
<tr>
<td>Z</td>
</tr>
<tr>
<td>Asymp. Sig. (2-tailed)</td>
</tr>
</tbody>
</table>

The significance obtained through the Mann-Whitney test (p=.0001) shows that there is a statistically significant difference in results of the post-test in groups C and E (Table 4).

**Analysis of the Quality of Duration of the Students’ New Knowledge of the Air-Related Content**

Quality of duration of the students’ knowledge in both groups on the retest at different cognitive levels is shown in the table (Table 5).
Table 5. Differences in the quality of duration of students’ knowledge in C and E group on the retest at the same cognitive levels.

<table>
<thead>
<tr>
<th>Cognitive level</th>
<th>Group</th>
<th>M ((\bar{X}))</th>
<th>SD</th>
<th>CV (%)</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge</td>
<td>E</td>
<td>2.921</td>
<td>.94</td>
<td>22.72</td>
<td>.7570</td>
<td>.7240</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>2.813</td>
<td>.81</td>
<td>25.58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Understanding</td>
<td>E</td>
<td>3.926</td>
<td>1.42</td>
<td>18.67</td>
<td>.8860</td>
<td>.3460</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>3.711</td>
<td>1.35</td>
<td>20.56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Application</td>
<td>E</td>
<td>4.988</td>
<td>.910</td>
<td>16.79</td>
<td>.9720</td>
<td>.1140</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>3.962</td>
<td>1.44</td>
<td>18.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analysis</td>
<td>E</td>
<td>9.022</td>
<td>1.43</td>
<td>4.79</td>
<td>1.957</td>
<td>.0020</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>5.121</td>
<td>7.11</td>
<td>18.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evaluation</td>
<td>E</td>
<td>5.022</td>
<td>1.29</td>
<td>5.12</td>
<td>5.226</td>
<td>.0001</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>2.003</td>
<td>6.98</td>
<td>39.96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Synthesis</td>
<td>E</td>
<td>3.887</td>
<td>2.18</td>
<td>4.95</td>
<td>5.885</td>
<td>.0001</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>0.793</td>
<td>8.81</td>
<td>15.89</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Results showed that E group acquired a more durable knowledge about air-related content than group C (Table 5) i.e. demonstration experiments did not contribute to durability (quality of duration) of the students’ knowledge comparing to student-led experiments.

Unlike on the post-test students of both groups displayed similar amount of knowledge at the first three cognitive levels. They were equal in knowledge (t=.7570, p=.7240), understanding (t=.8860, p=.3460) and application (t=.9720, p=.1140). Significant difference was at the levels of analysis (t=1.957, p=.0020), evaluation (t=5.226, p=.0001) and synthesis (t=5.885, p=.0001). Although the post-test showed no statistically significant difference in results at the cognitive level of analysis, on the retest it appeared as significant. Two items at the level of analysis (on the retest) were correctly solved by 13.22% students in group C, and by 44.94% in group E. One correctly solved item in the group C was acquired by 6.66% of the students, while in group E it was 39.96%. None of the students in group C managed to solve both items at the level of evaluation, while 11.66% did only one. In group E 23.33% of the students answered both items at the level of evaluation, and only 15% answered only one. At the level of synthesis none of the students in group C managed to solve any of the items. In group E 18.33% of the students solved one of the items at the level of synthesis but no one did both on the retest.

Based on the obtained value of the Kolmogorov–Smirnov normality test (Z=.8740, p=.0001) it was noted that data does not have normal distribution and that statistical significance of the difference (Table 3) between students in groups C and E in retest should be confirmed through the Mann-Whitney non-parametric test.

Statistical significance of difference in the quality of duration of students’ knowledge between groups C and E in the retest is acquired through the Mann-Whitney non-parametric test, and it is shown in the table (Table 6).

Table 6. Difference in the quality of duration of students’ knowledge in C and E group (retest) on the Mann-Whitney non-parametric test.

<table>
<thead>
<tr>
<th>The total number of points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mann-Whitney U</td>
</tr>
<tr>
<td>Wilcoxon W</td>
</tr>
<tr>
<td>Z</td>
</tr>
<tr>
<td>Asymp. Sig. (2-tailed)</td>
</tr>
</tbody>
</table>
Statistical significance obtained through the Mann-Whitney test ($p=.0001$) shows that there is a significant statistical difference between the results of the retest in groups C and E (Table 6).

**Comparative Analysis of the Quality of Students’ Knowledge (ANOVA Analysis)**

Based on the comparison of the quality of students' knowledge in each group (C and E) at the same cognitive levels in the pre-test, post-test and the retest it was concluded that there is no significant difference in number of students who solved items at the level of knowledge and understanding (value of $F$-test at both levels is higher than .005). Statistically significant difference between the groups in the post-test and the retest was noted at levels of application ($E: F=98.883, p=.0001; C: F=9.145, p=.0001$), analysis ($E: F=7.221, p=.001; C: F=1.87 p=.0001$), evaluation ($E: F=7.425, p=.001; C: F=9.112, p=.0001$) and synthesis ($E: F=10.986, p=.001; C: F=10.561, p=.0001$).

**Discussion**

Application of the laboratory-experimental method in teaching sciences in the first four grades of primary school contributes to the quality of students' knowledge. This is confirmed through research conducted by Cvjetičanin, Obadović & Rančić, (2015) and Cvjetičanin, Segedinac & Halaši, (2010). Based on the available research done on the contribution of the laboratory-experimental method in teaching sciences in the first four grades of primary school, it can be concluded that these were mostly done as comparative analyses of student-led experiments and the traditional (lecture) method (Cvjetičanin, Segedinac & Halaši, 2010, Golubović-Ilić, 2011). There is a small number of comparative research done on the contribution of demonstration and student-led experiments that they have on the quality of students’ knowledge in the first four grades of primary school. This is agreed upon by authors Cvjetičanin, Obadović & Rančić (2015). Since the research done in this area is rare and there is no research done about the contribution of demonstration and student-led experiments in air-related content for the third grade students, this research was conducted as a scientific contribution to the mentioned thematics.

Previous to the application of the experimental factor, both groups had been balanced in the aforementioned criteria. Students from both groups (C and E) had shown similar knowledge in the pre-test, with worse results in higher cognitive levels (analysis, evaluation and synthesis). The reason for decreased results at higher cognitive levels is probably due to the way in which the air-related content was acquired in previous levels, as well as the process of forgetting the previously learnt. In discussion with teachers, it was concluded that they have realized air-related content in classroom (previous classes) through verbal and written (textual) methods (the traditional, lecture approach), without using the laboratory-experimental method. This starting point has its justification in results of numerous research (in the comparison of the efficiency of application of the traditional, lecture approach with laboratory-experimental method), which have shown higher achievement and improvement in students who gained knowledge through the experimental-laboratory method, in subject (Cardak, Onder & Dikmenli, 2007, Odubunmi & Balogun, 1991, Ogundiwin, Asaaju, Adegoke & Ojo, 2015), as well as in the first four grades of primary school (Cvjetičanin, Segedinac & Halaši, 2010, Golubović-Ilić, 2011).

The students from both groups shared the similar knowledge on the levels: knowledge, understanding, application and analysis, while on the higher cognitive levels (evaluation and synthesis) students from the group E were more successful, they adopted the knowledge of the higher quality then the students from the group C. When the results of the C group students from the pre-test and the post-test are compared, it can be noted that the demonstration of experiments had a significant contribution on the increase of the quality of the students’ knowledge at the level of analysis and partly at the level of evaluation, while only slightly affected the quality of the students' knowledge at the level of synthesis. Research conducted by Cvjetičanin, Obadović & Rančić, (2015) acquired results which (unlike the listed data) shows that demonstration experiments had less contribution on the quality of knowledge in group C, because there was a statistically significant difference even at the level of analysis in favor of students who conducted experiments independently, which was not the case in this research. Research of McKee-Vickie, Williamson & Ruebush (2007), that was aimed to prove the influence of demonstration experiments on students' knowledge, has confirmed the positive contribution of demonstration and hands on experiments, i.e. a higher level of knowledge (conceptual understanding of the content) after the intervention in both groups. They have also shown that, when compared to hands-on experiments, demonstration experiments did not decrease students’ (biology students') conceptual understanding of the phenomena in question. The significance obtained through the Mann-Whitney test shows that there is a statistically significant difference in results of the post-test.
in groups C and E, in favor of the students from the group E, who have acquired the knowledge of higher quality.
The obtained results demonstrate the fact that the student-led experiments contribute more to the quality of the
knowledge than the demonstration experiments. Contrary to this conclusion, the research of Logar & Savec-Ferk
(2011) showed that students who learned through demonstration experiments acquired higher knowledge quality
when compared to those who conducted the experiments independently.

On the retest, the students from the C and E group shared the similar knowledge on the levels: knowledge,
understanding and application, while the students from the group E were more successful on the higher cogni-
tive levels (analysis, evaluation and synthesis) i.e. they adopted the more durable knowledge than the students’
from the group C. The obtained results were confirmed by the research of the following authors: Cvjetićanin,
Obadović & Rančić, (2015), in which both groups displayed similar knowledge in lower cognitive levels, with
the students who conducted student-led experiments being notably better in levels of analysis, evaluation and
synthesis when compared to the students who only saw those experiments demonstrated. When the results of
the students’ knowledge about the air-related content are compared at all cognitive levels based on the post-test
and retest, it can be seen that the students from both groups have acquired the knowledge of the lower quality
on the retest, then the knowledge acquired on the post-test. This was expected due to spontaneous forgetting
(Robbins, Schwartz & Wasserman, 2001). Students did not revise previously learnt between the post-test and the
retest, but did study different subject and thus, their air-related knowledge was disrupted (Stemberg & Zhang,
2001). Statistical significance obtained through the Mann-Whitney test shows that there is a significant statistical
difference between the results of the retest in groups C and E in favor of the students from the group E, who have
acquired the knowledge of higher durability.

Based on the comparison of the quality of students’ knowledge in each group (C and E) at the same cognitive
levels in the pre-test, post-test and the retest, it was concluded that there is no significant difference in number
of students who solved items at the levels of knowledge and understanding. Statistically significant difference
between the groups in the post-test and the retest was noted at levels of: application, analysis, evaluation and
synthesis. These results point to the fact that the application of experiments in both groups led to acquiring new
and perfecting existing knowledge (McKee-Vickie, Williamson & Ruebush, 2007) about air-related content. Similar
results were obtained by authors Cvjetićanin, Obadović & Rančić (2015).

Conclusions

The way of conducting experiments (demonstration or student-led) contributes to the quality of the third
grade students’ knowledge about the air-related content. Demonstration experiments have a significant contribu-
tion to the increase of the quality of students’ knowledge about air at the level of analysis, and partially at the level
of evaluation. However, their contribution to the increase in knowledge duration decreases by time at all levels
i.e. the students were not as successful at these levels while solving items during the retest. Unlike demonstration
experiments, student-led experiments contribute to the increase and duration of students’ knowledge quality at
the levels of analysis, evaluation and synthesis. More than demonstration experiments, student-led experiments
contribute to the students’ capability to: analyze, estimate, compare, rearrange, formulate, organize, etc. i.e. they
more contribute to the acquisition of the quality of students’ knowledge. The reason for higher contribution of
student-led experiments to the quality of the students’ knowledge about the air can be found in the facts: that
this type of experiment gives the students a chance to independently research, they follow what happens more
actively, in smaller groups they discuss mutually and obtain the conclusions of the results together, unlike those
students who merely watch the experiments during their demonstration by the teacher. These reasons clearly
point to the fact that realization of air-related content in the third grade of primary school (as well as in realization
of other natural sciences content) should be primarily done through student-led experiments. In this manner, in
addition to acquiring more quality knowledge, students develop higher experimental skills, get acquainted with
stages of scientific research and learn how to observe natural processes and phenomena through experiments.

Bearing in mind a relatively small amount of research done in this area, there is a need for more research
conducted with larger samples and related to the different content of the integrated science education in order to
establish the contribution of demonstration and student-led experiments on the quality of students’ knowledge
in realization of different content in teaching integrated natural sciences as a part of primary school.
References


Received: May 17, 2017

Accepted: August 05, 2017

Stanko Cvjetićanin (Corresponding author)
PhD, Professor, Faculty of Education, University of Novi Sad, Podgorička 4, 25 000 Sombor, Serbia.
E-mail: tozchemy@eunet.rs
Website: http://www.pef.uns.ac.rs/

Mirjana Maričić
MSc, PhD Student, Faculty of Education, University of Novi Sad, Podgorička 4, 25 000 Sombor, Serbia.
E-mail: mimiunt@hotmail.com
Website: http://www.pef.uns.ac.rs/