Abstract. The widely explored but still unresolved question about the contribution of the application of different ways of performing experiments in the integrated natural sciences is addressed in this research. The aim was to determine the contribution of demonstration hands-on experiments (DHE) and student hands-on experiments (SHE) in relation to conventional teaching method (CTM) on the quality and durability of 3rd grade students (between 9 and 10 years of age) from primary school. The research involved 180 students, further divided into three groups: E1 (experimental group 1, where content was learned through DHE), E2 (experimental group 2, where content was learned through SHE) and C (control group, where content was learned through CTM). The results of the research point to the fact that priority should be given to the DHE and SHE over the CTM in the realization of air-related content in the 3rd grade. SHE should be used more than DHE when it comes to teaching this specific content.

Keywords: demonstration hands-on experiments, student hands-on experiments, integrated sciences, primary school, quality of knowledge

Introduction

The beginning of the application of laboratory work in scientific education is related to the introduction of demonstration experiments (experiment performed by a teacher or student with advanced experimental skills in front of the whole class), while the beginning of the application of laboratory work within school practice (curricula) is related to the introduction of student experiments (experiment performed by students independently) (Hodson, 1993; Hofstein & Lunetta, 2004). Demonstration and student experiments can be laboratory, or simple hands-on experiments (Demircioğlu & Yadigaroğlu, 2011). A laboratory experiment involves a set of procedures that teacher or students perform with the application of laboratory equipment and materials under the supervision of teachers in order to obtain certain information about studied phenomenon (Salameh El Rabadi, 2013). Simple hands-on experiments also refer to a set of procedures that teacher or students perform (with the guidance of teacher), but their application does not require expensive equipment and laboratory equipment, and they can be realized with the help of simple, inexpensive, easily accessible materials and resources (Hırça, 2013, Sadi & Cakiroglu, 2011). However, they differ from the science laboratory in the fact that they do not represent experiments through which students discover new scientific knowledge, but acquire knowledge, which is already known to science, but is new and unknown to students themselves (Ruby, 2001). They also differ from the CTM (the usual teacher’s transfer of content), because students have the opportunity to: communicate with materials, make their own observations, expand their knowledge, understand causal and consequential relations and understand the basis of the scientific research and develop experimental skills (Akan, 2015; Kibirige, Maake, & Francis-Mavhunga, 2014; Sadi & Cakiroglu, 2011).

Relevant Research

Research which have included the contribution of the application of laboratory and hands-on experiments (demonstration and student experiments) are particularly numerous in science education i.e. science subject
TEACHER-Demonstration AND STUDENT HANDS-ON EXPERIMENTS IN TEACHING INTEGRATED SCIENCES

It has been done with students from the fifth grade of primary school and onwards, because in this case laboratory work mostly refers to student activities within the school laboratory, or with the help of laboratory equipment in the classroom, when interacting with various materials in small groups, or individually for the purpose of examining a scientific phenomenon. In this field, such research has been done within the framework of cognitive and affective domains, and everything which could have been claimed as a part of the field of science subject teaching. The results of these recent research have shown that there are positive effects:

- **SHE in relation to the CTM to:** achievement, learning outcomes of low performing students, mastery of chemistry concepts, science process skills students' attitudes (Ateş & Eryilmaz, 2011; Cardak, Onder, & Dikmenli, 2007; Ojediran, Oludipe, & Ehindero, 2014; Okam & Idris-Zakari, 2017; Sadi & Cakiroğlu, 2011), student's understanding of subjects (Demircioğlu & Yadigaroğlu, 2011), students' performance (Kibirige et al., 2014), scientific thinking (Hugert, Najami, Abbasi, & Dkeidek, 2014). For example, Cardak et al. (2007) have examined the impact of student experiments within the experimental group in comparison to the traditional approach within the control group. The results of the research have shown better achievements of the students from the experimental group who studied the same contents with the application of experiments.

- **DHE in relation to the CTM to:** elimination of misconception (Sert-Çibik, Diken, & Darçin, 2008), students' performance (Udo, 2010), on student's achievement (Solà & Ojo, 2007). For example, the research by Sert-Çibik et al. (2008), has examined the influence of group work and demonstration experiments (E-experimental group) in comparison to traditional teaching (C-control group). Group work with a demonstration approach to performing experiments has positively influenced the removal of student's misconceptions, i.e. E group students have shown better results.

- **SHE in comparison to the DHE to:** the achievement of students in learning (Musasia, Ocholla, & Sakwa, 2016; Wachanga & Gwóland-Mwaangi, 2004), student performance, quality and durability of knowledge (Irinoye, Bamidele, Adetunji, & Awodele, 2015; Logar & Savec-Ferk, 2011), development of scientific skills and students attitudes towards subject and science (Bilgin, 2006; Obadović, Rančić, Cvjetićanin, & Segedinac, 2013). In addition, the student experiments have proven to be as effective as the teacher's demonstration experiments on the conceptual understanding of the content being studied (McKee-Vickie, Williamson, & Ruebush, 2007). For example, the research conducted by McKee-Vickie et al. (2007), has examined the impact of inquiry hands-on experiments within the experimental group in comparison to inquiry demonstration experiments within the control group, on the conceptual understanding the content of Reactions of Calcium. The results of the study have indicated that the students of both groups showed almost equal conceptual understanding of the content they were studying.

- **SHE and DHE in relation to the CTM to:** students' achievement (Solà & Ojo, 2007), students' performance (Udo, 2010). For example, a research conducted by Udo (2010) has studied the impact of applying guided discovery in the experimental group 1 (E1), the student-centered demonstration within the experimental group 2 (E2) and the application of the method of exposure within the control condition (C). The students from the Group E1 have shown the best achievements, followed by the students from the Group E2 and the students from the Group C.

Students who are a part of the integrated science education system (from the 1st to the 4th grade) do not go into school laboratories but learn the content of biology- physics-chemistry- ecology in an integrated form, mainly in their classrooms. In the teaching process, when the content permits, they can use certain safe laboratory equipment, or easily accessible, simple materials (with the supervision of teachers), which enable them to acquire scientific knowledge in practice. In this field, research involving the application and contribution of simple hands-on experiments (demonstration and students) using the above-mentioned equipment, has been very rare and represents an unexplored field in world scientific education. The research on the application and contribution of the application of simple hands-on experiments in integrated science education within the cognitive and affective domains are significantly less represented. The results of these research have shown positive effects:

- **SHE in relation to the CTM to:** quantum and quality of knowledge (Cvjetićanin, Segedinac, & Halaši, 2010), understanding of the subject (Cakici & Yavuz, 2010), student interest and motivation (Golubović-Ilić, 2011). For example, Cakici and Yavuz (2010) have explored the impact of a constructivist approach (with experiments) within the experimental group in relation to the traditional approach within the control group. The results of the research have indicated a much better success (understanding the content of Matter) of students from the experimental group.
• **SHE in comparison to DHE to:** quality and retention of student knowledge (Cvjetićanin & Maričić, 2017; Cvjetićanin, Obadović, & Rančić, 2015). For example, a research by Cvjetićanin et al. (2010) has studied the influence of student hands-on experiments which has been realized within the experimental group in comparison to traditional teaching which has been carried out within the control group. The results have shown a better achievement of the students of the experimental group in terms of quantum and quality of knowledge.

*Identifying the Research Problem – Focus and Aim*

Most previous studies related to the examination of the contribution of DHE in comparison to SHE to the quality and durability of student knowledge, i.e. DHE and SHE in contrast to CTM (learning through instruction, without hands-on experiments) have been carried out within individual subject teaching (mostly Physics and Chemistry). With classroom students, such research has been significantly less widely conducted in the processing of various elements of the integrated natural sciences. Based on a detailed analysis of the available literature, no studies have been found on the comparative contribution of DHE and SHE in relation to CTM in the integrated sciences. In all the analyzed studies, the application of DHE and SHE has been examined; SHE versus DHE, SHE versus CTM or SHE versus CTM. No studies have been found which have examined the difference in the contribution of DHE and SHE in relation to CTM to the quality and durability of students' knowledge at different cognitive levels (under the Anderson-Krathwohl-Bloom taxonomy).

The content of the integrated natural sciences is significantly more complex in the 3rd grade of primary school. Content related to the topic of the air has a special place. Students need to understand various concepts and properties of air (air movement, mass, volume, pressure, and so on), the influence of temperature on air, and air as an insulator. For this reason, the question arises: When will students achieve better and more lasting knowledge, if they study this content using DHE, SHE or CTM (verbal-textual method)?

The following aim thus comes into focus: Determining the contribution of DHE and SHE in relation to the CTM to the quality and durability of 3rd grade student knowledge about air-related content in the integrated sciences at all levels of the Anderson-Krathwohl-Bloom taxonomy. From the set goal, the following tasks can be identified:

1. Examine whether DHE contribute more in relation to the CTM to gaining better quality and more lasting knowledge of 3rd grade students on air at all cognitive levels?
2. Examine whether SHE contribute more in relation to the CTM to gaining better quality and more lasting knowledge of 3rd grade students on air at all cognitive levels?
3. Examine the quality and durability of 3rd grade students' knowledge of air on all cognitive levels by the way in which experiments are carried out (when the students themselves perform or when these experiments are demonstrated to them)?

*Research Methodology*

*Research Design*

The research was quasi-experimental in character and was carried out according to the experimental design with parallel groups, in the school year 2017/2018 in a time period of three months, through 8 phases (Table 1).
### Sample

The research included students from the 3rd grade (between 9 and 10 years of age) from nine classes in four primary schools ($N = 180$ students) in the municipality of Novi Sad - area of the Autonomous Province of Vojvodina (Republic of Serbia). The sample of the research was of a convenience character, i.e. the students who were available to the researcher at that particular moment, were selected. The research has been approved by the management of selected schools with the consent of: pedagogue, psychologist and teachers of selected classes. The research involved exactly 180 pupils, since these schools have nine 3rd grade classes. Three numerically uniform groups of students were formed out of these nine classes. As there were between 20 and 25 students in each class, each of the three groups consisted of 60 students. In addition to the criteria of numerical uniformity, the researchers have paid attention for the selected classes to have: approximate average grades at the end of the 2nd grade; approximate average grades from the subject *The World around Us*; and approximate previous knowledge of air-related content from the previous grades (at pre-test). Data were obtained by *one-way ANOVA analysis* and it showed that there is no statistically significant difference between these criteria ($F = .084, p > .05$).

### Research Instruments

All groups of students were given the same tests: pre-test, post-test and re-test. Since there are no standardized tests for checking students' knowledge of the contents of integrated natural sciences in the system of education of the Republic of Serbia, the researchers had themselves created tests with the help of teachers who taught in classes. The validity of the tests was provided by examination and confirmation of two teachers who have a working experience of over 10 years, and by examination and confirmation of two methodologists of integrated natural sciences - experts in this field. The values of the *Cronbach Alpha* coefficient for each test were: pre-test $\alpha = .81$, post-test $\alpha = .83$, re-test $\alpha = .86$. Anderson, Krathwohl and Bloom taxonomy (Anderson, Krathwohl, & Bloom, 2001) was used to create test item, while the manual *Smart tests: Teacher-made tests that help students learn* (Walker & Schmidt, 2004) was used for their design at cognitive levels. The pre-test was different from the post-test and re-test by the contents that were examined. Items based on air-related content from the previous classes were in the pre-test, while items based on air-related content from the 3rd grade were on the post-test and re-test. Post-test and re-test differed only in the formulation of items, while in structure and content they were very similar. In all tests, the quality of students' knowledge was tested with two items of different structures, at each cognitive level. Each item carried a higher number of points at the next - higher cognitive level.

The pre-test was designed and applied before the application of new content in order to provide the last criteria for balancing groups and determine the quality of the existing student knowledge at all cognitive levels on air-related content, and provide an updated basis for upgrading, expanding and adopting new knowledge of air-related content. Pre-test results showed researchers whether the students had adopted the basic concepts of air in the previous classes, since they are necessary for the acquisition of new knowledge about these contents. The pre-test consisted of 12 items, created according to the model of the authors (Blagdatić, Kovačević, & Jović, 2016; Kukić & Aćimović, 2016). The post-test was designed and applied after the pre-test in order to determine the quality of the newly acquired knowledge of students at all cognitive levels on air-related content. The re-test was designed and applied two months after the application of the post-test in order to determine the durability of the students' newly acquired air-related content at all cognitive levels. Post-test and re-test consisted of 12 items, created according to the model of the authors (Blagdatić, Jović, Kovačević, & Petrović, 2016; Ralić-Žeželj, 2016).
Procedure and Treatment

For the purpose of this research, the following themes were selected: Non-living nature and Materials and their use. The air-related content that was realized in the groups C, E1, E2, and processed within the aforementioned teaching topics consisting of the following teaching units: The shape and pressure of the air, Changes arising from heating and cooling of air (change temperature, volume, air flow...), Air heat insulator. By character, the selected content relates to basic knowledge in physics and chemistry, but in an integrated form. The same teacher worked in all groups (C, E1 and E2). In the groups E1 and E2, the same experiments were performed (10 experiments in total per group), and they were selected and adapted to the cognitive, mental and physical characteristics of students of the 3rd grade. The level of complexity of the experiment increased (Figure 1) with each lesson. The basic methodic rule was employed. The simplest experiments were performed first (Example 1), followed by the experiments of average complexity (Example 2) and ending with more complex experiments (Example 3). Before each class the teacher complied with the basic methodological rule and performed experiments in his cabinet in order to see if they could be applied.

Example 1.
Name of the experiment: Warm and cool air!
Necessary equipment: balloon, and empty glass bottle, deeper bowl with warm water.
Description of the experiment: Put a balloon on the bottle, merge it in warm water and keep it like that for a minute! What goes on? After that put the bottle in cold water and see what happens.
Pupil’s tasks: Does air have a shape? How does the temperature affect expanding and contracting of air?

Example 2.
Name of the experiment: Air mass
Necessary equipment: two identic balloons and a weighing scale.
Description of the experiment: put the balloons on a weighing scale. After that, remove one balloon and blow it, tie it and put on a weighing scale.
Pupil’s tasks: What happened to the balloon when it was blown and why? What happened to the weighing scale during the experiment and why? Based on this, which conclusion can you make?

Example 3.
Name of the experiment: Air insulator
Necessary equipment: three plastic bottles, jug with warm water; thermometer.
Description of the experiment: Put one glass inside another, and then pour the warm water in to reach the top of the glass. Take a third glass and pour the warm water in, to reach the top of the glass. After that, use the thermometer to measure the temperature of the water in the glass (both in regular and double glass). Measure the temperature in three-minute intervals. Enter data into the table.

<table>
<thead>
<tr>
<th>Time (minutes)</th>
<th>Water temperature in double cup (° C)</th>
<th>Water temperature in a plain cup (° C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>In the beginning (0 minutes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Student’s tasks: How did the temperature act in the regular and how in the double glass? What is between two glasses? What can you conclude based on it, what attribute does the air have? Give some examples from everyday life, where can this characteristic of air be applied.

Figure 1. Examples of experiments with student’s tasks.
Treatment in the group C

In the group C, the air-related contents were realized in a conventional way (CTM) using the verbal-textual method. The teacher taught new air-related content to pupils using heuristic approach, following the material from the textbook. Group C students did not perform experiments. After each explained air-related concept, students noted the most important information into their notebooks. After finishing teaching, the teacher would revise with students all the taught concepts.

Treatment in the group E1

Before the demonstration of each experiment in the group E1, the teacher explained the following to students: which equipment and materials are necessary for the experiment, how to demonstrate the experiment, and how to observe it. The frontal approach was used. After each experiment carried out by the teacher, a brief discussion of the results was initiated, and the students correctly recorded the views in their special notebooks (experiment notebooks). After all experiments were demonstrated, the common important conclusions about the air were made, which the teacher wrote on the board, and the students in their regular notebooks at the end of the class.

Treatment in the group E2

Group E2 members were split into heterogeneous groups, composed of three students of different levels of knowledge (good, average and low) and that did not change during classes in which air-related content was taught. Before the beginning of each lesson, during which the pupils acquired new air-related content through conducting an experiment, the teacher placed on the table of each group the necessary material (equipment) for experiments and an instructive sheet for each group member in particular. Prior to performing the experiment, the students were supposed to carefully read the instructions from the instruction sheet, and then talk in the group about what is expected of them. The instructions contained the following information: the equipment and material, the description of how to conduct the experiment and the tasks which the students were supposed to complete (based on their results and observations after the experiment was performed). When the teacher estimated that the students were familiar with the text and were ready to work, they approached the experimental phase. All groups of students performed experiments of the same level of complexity at all classes. Upon completion of the work of the group, a representative of each group presented the results that his/her group obtained from the experiment. After results of all groups were presented, a discussion was initiated in which the students corrected the wrong conclusions and made the correct ones. The students then checked their notes to see whether they wrote down the course of the experiment correctly, whether their answers and conclusions were correct and if they were not, they were corrected. After that students used their regular notebooks to write down the common conclusions about air, which came as a result of all group’s work.

Data Analysis

Cronbach alpha coefficient was used in order to determine the reliability of the tests. One-way ANOVA analysis was applied to determine if there is a statistically significant difference between the average grades of students at the end of the 2nd grade, the average grades of students from the subject The World around Us and the grades they achieved on the pre-test. One-way ANOVA analysis was also applied to compare differences between mean values of all groups (E1, E2 and C) at pre-test, post-test and re-test. To determine the statistically significant difference between the quality and the durability of the knowledge of each of the two groups separately (E1 and E2, E1 and C, E2 and C), the Scheffe post-hoc test was used. Wilcoxon test was used in order to determine the difference between pre-test and post-test and the difference between post-test and re-test in all three groups. SPSS program, version 22, was used for statistical analysis of data.

Research Results

The results of Spearman’s correlation coefficient showed a moderate correlation between the average grades of students at the end of the 2nd grade and the grades they achieved at the pre-test ($\rho = .463$ with $p < .001$), and
between the average grades of students from the subject *The World around Us* and their grades at pre-test ($\rho = .481$, $p < .001$). One-way ANOVA analysis showed that there are no statistically significant differences between the students' knowledge of C, E1 and E2 group at all cognitive levels at the pre-test ($p > .05$). Statistically significant difference is present at post-test and re-test at the level of: analysis, evaluation and synthesis (Table 2).

### Table 2. Statistically significant differences between students of C, E1 and E2 group at all cognitive levels at post-test and re-test - one-way ANOVA.

<table>
<thead>
<tr>
<th>Cognitive level</th>
<th>Group</th>
<th>Sum of Squares</th>
<th>$SD$</th>
<th>$SE$</th>
<th>$F$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>AS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-test</td>
<td></td>
<td>Between groups</td>
<td>90.989</td>
<td>2</td>
<td>46.214</td>
<td>8.097</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Within groups</td>
<td>1119.622</td>
<td>175</td>
<td>6.162</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>1221.503</td>
<td>178</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Between groups</td>
<td>397.230</td>
<td>2</td>
<td>231.122</td>
<td>26.623</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Within groups</td>
<td>1331.334</td>
<td>168</td>
<td>8.112</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>1882.656</td>
<td>171</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Between groups</td>
<td>675.655</td>
<td>2</td>
<td>354.298</td>
<td>29.103</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Within groups</td>
<td>2098.297</td>
<td>178</td>
<td>10.925</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>2773.131</td>
<td>181</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Re-test</td>
<td></td>
<td>Between groups</td>
<td>110.993</td>
<td>2</td>
<td>54959</td>
<td>12.996</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Within groups</td>
<td>752.795</td>
<td>175</td>
<td>4.189</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>859.876</td>
<td>180</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Between groups</td>
<td>492.878</td>
<td>2</td>
<td>249.022</td>
<td>32.003</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Within groups</td>
<td>1352.524</td>
<td>176</td>
<td>7.03</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>1871.124</td>
<td>179</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Between groups</td>
<td>747.205</td>
<td>2</td>
<td>373.031</td>
<td>39.025</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Within groups</td>
<td>1698.997</td>
<td>177</td>
<td>9.71</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>2447.122</td>
<td>179</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The Scheffe post-hock test showed that there is a statistically significant difference between the quality and durability of E1 and E2 group knowledge at the level of evaluation and synthesis at post-test and at the level of analysis, evaluation and synthesis at re-test (Table 3).

### Table 3. Statistically significant differences in the quality and retention of knowledge between students of C, E1 and E2 group - Scheffe post-hock test.

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>(I) Group</th>
<th>(J) Group</th>
<th>Mean Difference (I-J)</th>
<th>SE</th>
<th>$p$</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lower Bound  Upper Bound</td>
</tr>
<tr>
<td>Post-test</td>
<td>C</td>
<td>E1</td>
<td>-1.54866</td>
<td>.45838</td>
<td>.011</td>
<td>-2.3912 - .1092</td>
</tr>
<tr>
<td></td>
<td>E2</td>
<td>C</td>
<td>-1.54866</td>
<td>.45838</td>
<td>.011</td>
<td>-2.3912 - .1092</td>
</tr>
<tr>
<td></td>
<td>E1</td>
<td>E2</td>
<td>-1.42098</td>
<td>.48087</td>
<td>.630</td>
<td>-2.5665 .7025</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>E1</td>
<td>-1.68667</td>
<td>.45838</td>
<td>.001</td>
<td>-2.7979 .5397</td>
</tr>
<tr>
<td></td>
<td>E2</td>
<td>C</td>
<td>-1.68667</td>
<td>.45838</td>
<td>.001</td>
<td>-2.7979 .5397</td>
</tr>
<tr>
<td></td>
<td>E1</td>
<td>E2</td>
<td>-1.42098</td>
<td>.48087</td>
<td>.630</td>
<td>-2.5665 .7025</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The **Coefficient of variation (CV)** showed significantly lower values at the level of analysis, evaluation and synthesis in the E1 and E2 groups compared to the C group at the post-test, as well as the lower values at the level of analysis, evaluation and synthesis in the E2 group compared to E1 and C group at the re-test (Table 4).

### Table 4. CV in all groups (C, E1 and E2) at individual cognitive levels at post-test and re-test.

<table>
<thead>
<tr>
<th>Group</th>
<th>Analysis</th>
<th>Evaluation</th>
<th>Synthesis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>Post-test</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>15.73</td>
<td>17.22</td>
<td>18.75</td>
</tr>
<tr>
<td>E1</td>
<td>4.98</td>
<td>5.88</td>
<td>6.68</td>
</tr>
<tr>
<td>E2</td>
<td>3.73</td>
<td>4.34</td>
<td>5.12</td>
</tr>
<tr>
<td></td>
<td>Re-test</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>19.01</td>
<td>19.63</td>
<td>20.01</td>
</tr>
<tr>
<td>E1</td>
<td>11.95</td>
<td>12.91</td>
<td>14.08</td>
</tr>
<tr>
<td>E2</td>
<td>4.12</td>
<td>4.97</td>
<td>5.03</td>
</tr>
</tbody>
</table>
The Wilcoxon test showed that there is no statistically significant difference between the pre-test and the post-test results of the students C group at all cognitive levels. Statistically significant difference can be found in the E1 group at the analysis level (Z = -1.782, p = .026, r = .187), i.e. in the group E2 at the analysis (Z = -1.824, p = .018, r = .192), evaluation (Z = -1.763, p = .012, r = .173) and synthesis level (Z = -1.886, p = .026, r = .190). When analyzing the difference between pre-test and re-test results, it can be noted that there is a statistically significant difference only at the analysis level in the E1 group (Z = -1.801, p = .039, r = .188).

When comparing students' success in each group (E1, E2, and K) at the same cognitive level at the pre-test, post-test and re-test, it can be noticed that there is a statistically significant difference in group E1 and group E2 in the number of students who have correctly solved the items at the cognitive level of: application (E1: \(F = 10.212, p = .012, E2: F = 101.031, p = .001\)), evaluation (E1: \(F = 1.92, p = .0001, E2: F = 6.989, p = .001\)) and synthesis (E1: \(F = 11.292, p = .011, E2: F = 12.783, p = .001\)) on these tests. This difference is not present in the C group (\(p > .05\)).

### Discussion

The results of the research showed that the high average grades of students at the end of the 2nd grade, the high average grades in the subject of The World around Us, as well as revising the previous knowledge of the air-related content, is not a reliable criterion for the students to achieve good pre-test results. The poor knowledge of students in groups C, E1 and E2 at pre-test is most likely the result of the previous learning method through which students acquired these contents, as well as the process of forgetting and interfering with other content (Rather, 2010). In conversation with teachers of students who participated in the research, as well as with the students themselves, it was concluded that the air-related content, as well as other contents of natural sciences in the 1st and 2nd grade, were realized through the application of the CTM, without use of experiments. In addition, the results showed that there was no difference in the quality of knowledge of students of all three groups at all cognitive levels at the pre-test. At the level of analysis, evaluation and synthesis, students C, E1 and E2 group showed lower knowledge. The results obtained on the pre-test are similar to the results of numerous research in which the efficiency of the application of CTM has been examined and compared with the laboratory method on the quality of students' knowledge in science subject teaching and integrated science education (Cvjetićanin & Maričić, 2017; Golubović-Ilić, 2011).

Students of all three groups achieved better post-test results, which can be noticed on the basis of a comparative analysis of the total number of points achieved on the pre-test and post-test. Statistically significant difference between pre-test and post-test results of students is noticed within the E1 and E2 groups. However, it is not statistically significant in the group C, which leads to the conclusion that CTM cannot enable students to acquire knowledge at higher cognitive levels (Ivić, Pešikan, & Antić, 2001). The obtained results are consistent with the results of most of the previous research on the contribution of experiments (DHE and SHE) to the quality of students' knowledge of the contents of natural sciences in relation to CTM, or some teaching method in science subject teaching and integrated science education. Thus, for example, in a research by Musasia et al. (2016) students of control and experimental group who studied physics contents using traditional approach and practical (experimental) activities have achieved better results at post-test than at pre-test. The same results were obtained in the research conducted by Taul and Erol (2008), where a difference in the quality of student knowledge of magnetism has been studied, when students were taught in a conventional, lecturing way and with cooperative learning. In contrast to the aforementioned, research conducted by Uside, Barchok, and Abura (2013) found that between the results of the pre-test and the post-test, no statistically significant difference was noticed in the students' knowledge of physical content, given they studied using experiments and in a traditional way. Students of all three groups achieved similar knowledge at the level of knowledge, understanding and application. Statistically significant difference between the quality of students' knowledge was noticed at the level of analysis, evaluation and synthesis in groups C and E1 and groups C and E2. The difference in success between students in group E1 and E2 was noticed only at the level of evaluation and synthesis, i.e. students in the E2 group were more successful than E1 students in the correct solution of two or one item at these levels. In addition, it can be noted that DHE significantly influenced the quality of E1 students' knowledge of the level of analysis and slightly at the level of evaluation and synthesis, when comparing their results from the pre-test and post-test. Similar results have also been found in the research Sert-Çibik et al. (2008) in which demonstration experiments significantly influenced students' success in relation to the conventional approach. However, the results of some research disagree with the results of this and other research. Thus, for example, research by Cvjetićanin et al. (2015) has shown that demonstration experiments contributed to the quality of student knowledge about the materials...
at the level of analysis, as there was difference in knowledge at this level between students to whom experiments were demonstrated and students who conducted same experiments independently. The results of the research have shown that students of E1 and E2 groups have higher uniformity of knowledge at the level of: analysis, evaluation and synthesis, than C group students, whereas this uniformity is higher in the E2 group than in the E1 group at the level of evaluation and synthesis. The obtained results point to the fact that SHE contribute more to the better quality of student knowledge of 3rd grade on air-related content when compared to DHE. The results of this research are similar to the results of the research (Wachanga & Gowland-Mwaangi, 2004) in which both experimental groups of students (1EG and 3EG), who studied chemical contents with the application of experiments with a cooperative approach, have achieved considerably better results than the students of both control groups (2CG and 4CG), who have been taught the same contents with the application of regular learning methods. In a research by Rendler and Hulde (2007), students who performed hands-on experiments have gained more durable knowledge from students to whom these experiments have been demonstrated. Contrary to what was stated, in the Logar and Savec-Ferk (2011) research, results have shown that demonstration experiments had a greater effect on students' knowledge quality than students' hands-on experiments.

The results of the research showed that there was no difference in the C and E2 groups between the post-test and the re-test results in terms of the quality of students' knowledge at every cognitive level. Statistically significant difference is noticed in the group E1 but only at the level of the analysis. At the lower cognitive levels (knowledge, understanding and application), students of all three groups achieved similar knowledge. Students of the E2 group were more successful at the level of analysis, evaluation and synthesis, i.e. they showed more durable knowledge than students from the C and E1 groups. None of the students in the group E1 at the re-test did both items at the analysis level, while only 11.66% answered correctly to one item. When comparing students' knowledge of air on all cognitive levels at post-test and re-test, it is noted that students in C, E1 and E2 groups achieved lower success at re-test from success they achieved at the post-test. These results are in correlation with research results (Cvjetičanin & Maričič, 2017) in which the students of the control and experimental group displayed lower knowledge at the re-test in relation to the post-test. Bearing in mind the effect of spontaneous and active forgetting, this was to be expected (Glynn, Britton, & Yeany, 2012; Robbins, Schwartz, & Wasserman, 2001). The students did not revise the air-related content between the post-test and the re-test i.e. they studied other contents during that time period, which affected the air-related content (Sternberg & Zhang, 2001). The obtained results are similar to the results of numerous research. For example, in the research Cvjetičanin et al. (2015), E group students (student experiments) achieved more lasting knowledge in relation to the students from the K group (demonstration experiments) at higher cognitive levels. The results of the Badeleh (2011) research are similar, where the students of the experimental group achieved more lasting knowledge of chemistry contents from the students of the control group at all cognitive levels. In addition, the results of the research have shown that there is a statistically significant difference (as opposed to post-test) between the knowledge of students E1 and E2 groups at the level of analysis, evaluation and synthesis. The students in group E2 showed greater uniformity in their knowledge at these cognitive levels, as compared to the students in groups C and E1.

When comparing students' success in each group, especially, on pre-test, post-test, and re-tests at the same level of knowledge, it can be noted that there is no statistically significant difference in the number of students who correctly solved the items at the first two cognitive levels. Statistically significant difference is noticed within the E1 and E2 groups at the cognitive level of application, analysis, evaluation, and synthesis, while it does not exist within C group. These data confirm the fact that simple hands-on experiments (SHE and DHE) contribute more to a higher quality and durability of students' knowledge than CTM.

Conclusions

This research confirms the fact that students achieve better knowledge using hands-on experiments than when they study the same content using the CTM. In this study, unlike in previous (available) research, the difference in the contribution of DHE and SHE comparing to the CTM in terms of both the quality and durability of student knowledge in the 3rd grade of primary school education on air at different cognitive levels has been examined for the first time. The majority of students who studied using the CTM achieved the highest quality and durability of knowledge at the cognitive level of application. The means of conducting the experiment affects the quality and durability of students' knowledge of the topic of the air. DHE significantly contributes to the quality of students' knowledge at the level of analysis and partly at the level of evaluation, while their impact on the durability of knowledge at these levels decreases over time. Unlike DHE, SHE affects the quality of knowledge at higher cognitive levels, including those of
analysis, evaluation and synthesis. Their contribution does not decrease over time; that is, these experiments equally influence the durability of students’ knowledge. Therefore, it is suggested that when processing this content, preference should be given to hands-on experiments in relation to CTM, i.e. SHE as opposed to DHE.

Acknowledgements

Funding: This research was done within the following projects:
- Ministry of Science Montenegro – HERIC project [Grant number: 01-2864].

References


