Development of inquiry skills at upper secondary level

Zuzana Ješková, Veronika Jurková, Stanislav Lukáč, Lubomír Šnajder and Ján Guniš

Faculty of Science, Pavol Jozef Šafárik University in Košice, Šrobárova 1, 040 01 Košice, Slovakia

Abstract. Development of inquiry skills is still in the main focus of science educators in Slovakia. Large effort has been recently put into the implementation of inquiry-based science education. In this research, the current level of development of selected inquiry skills has been monitored and analysed. For this purpose, a multiple-choice test of selected inquiry skills was designed. The tested skills involve formulating hypothesis, designing experiment, transforming results into tables and graphs, determining relationships between variables based on tables and graphs and identifying possible sources of errors. The test was implemented across several thousands of upper secondary school students. The inquiry skills test results revealed significant lack of their development. The paper describes the test design as well as it analyses and interprets the test results.

1. Introduction
In Slovakia there is a long-time effort to implement inquiry-based science education (IBSE). The current science education curriculum running from 2008 focuses not only on the development of knowledge but it strongly emphasizes development of inquiry skills. However, how the IBSE approach, if implemented, influences the inquiry skills development, is still an open question. There were several studies carried out recently in order to show the current level of development of these skills. The results of these studies show insufficient level of their development in Slovakia [1, 2, 3] as well as in other countries [4, 5, 6, 7]. Development of inquiry skills is in the focus of the large National project IT Academy (www.itakademia.sk) which tries to attract more students towards science and informatics career through an intentional implementation of inquiry-based science education (IBSE) enhanced by digital technologies and teacher training in this field. In the framework of the IT Academy project that involves several thousands of upper secondary school students from different parts of Slovakia, the following research questions were formulated and answered: What is the current level of development of selected inquiry skills of upper secondary school students? What are the differences between selected groups of students? What are the differences in the level of selected inquiry skills development?

2. Methods

2.1. Study design
This research builds on previous similar research studies [1, 2, 3] focusing on assessing inquiry skills of upper secondary school students. The previous research carried out in Slovakia in 2015/16 was based on testing applying a testing instrument implemented across several hundreds of students. This time we wanted to conduct even more consistent and in-depth research across several thousands of students involved in the IT Academy project. For this purpose, the existing test of inquiry skills used in previous research [3] has been modified. The test was implemented in the classes where teaching was conducted...
without any external pedagogical intervention. After data collection, the test results were analysed and interpreted.

2.2. Research instrument
The multiple-choice test has been developed to monitor the level of selected inquiry skills. This test was updated from its previous version presented and described in [3] in detail. The previous test version involved multiple-choice test items as well as several open-ended test items and was developed on the basis of analysis of the existing testing instruments [7, 8, 9] and inquiry skills taxonomies [10, 11, 12]. Since we planned a large sample of students answering the test, we needed an online version that is possible to evaluate electronically. That’s why the previous test version was modified slightly. Firstly, we reduced the number of tested skills. The tested skills finally involve the skills of formulating hypothesis, designing experiment, transforming results into tables and graphs, determining relationships between variables based on tables and graphs and identifying possible sources of errors. At least two test items were assigned for each of the selected skill. Some test items were situated in other than physics context in order to create more content-independent items. Most test items are multiple-choice items with one or two correct answers. In two questions students select from a set of variables to identify appropriate variables and their relationship. The test finally involves 14 test items. Two of them are divided into two parts following the same item context. The test items are distributed over different inquiry skills and subject matter as presented in table 1 to be answered within 45 minutes’ time limit. The students´ answers were evaluated on the basis of agreed criteria. They were assigned a value of 0-1 point which means that students could achieve a test score in the range of 0-14 points. For multiple-choice items with one correct answer students got 1 point for the correct choice, but in all other cases (even when the correct answer was combined with another option) they got 0 points. Multiple-choice items with two correct answers were scored with 1 point only in the case of choosing just two right answers and 0.5 points were assigned in the case of choosing just one answer that was right. In all other cases the score of 0 points was assigned. The test was prepared online that made the data collection and evaluation easier.

Table 1. Distribution of test items among different inquiry skills and subject matter.

| Number of test items | Subject                          |
|----------------------|----------------------------------|
| Formulate hypothesis to be tested | 2 chemistry, physics            |
| Design experiment (identify variables and their relationship) | 4 physics, physics, informatics, chemistry |
| Transform data to standard forms (i.e. tables or graphs) | 2 physics, informatics, mathematics |
| Determine relationship between variables (based on tables) | 2 physics, informatics, mathematics |
| Determine relationship between variables (based on graphs) | 2 physics, mathematics |
| Determine accuracy (identify possible sources of errors) | 2 physics, biology |

2.3. Research sample
The test was implemented in the period of March 2018 – June 2019 as a pre-test in the classes that come from schools involved in the IT Academy project. The research sample involves finally 6675 upper
secondary school students. They come from grammar (high) schools (74.5%) and vocational (technical) schools (25.5%). These types of schools were chosen for the study because they offer the highest level secondary education in Slovakia and students leaving these schools generally continue to study at University. The students aged 15-18 involved in the study belonged to 1st, 2nd, 3rd or 4th grade classes. They were taught by teachers who did not undergo any special training. The sample distribution can be seen in table 2 and figure 1.

### Table 2. Sample distribution.

| Gender | Grade | Type of school |
|--------|-------|----------------|
| Boys   | Girls | 1st | 2nd | 3rd | 4th | grammar | vocational |
| 3259   | 3416  | 2829| 1990| 1407| 445 | 4970    | 1705       |
| 48.9   | 51.1  | 42.4| 29.8| 21.2| 6.7 | 74.5    | 25.5       |

![Distribution of research sample across gender, grade and type of school.](image)

2.4. **Data analysis**

The basic descriptive statistics was used to summarize the main features of the sample distribution. The correlation between different set of data was calculated with the help of Pearson correlation coefficient. For testing differences in the groups (gender, school, grade), the two sample Mann-Whitney test was used. Kruskal Wallis test and Bonferroni post hoc test were applied for testing differences between the groups grade. In all mentioned parts of the data analysis normality of differences or residuals was being confirmed by the Kolmogorov-Smirnov test. The statistical analysis was realized in SPSS software.

3. **Results and discussion**

3.1. **Quantitative analysis**

The test results and its analysis revealed low level of development of tested skills with average score less than 31% (table 3). This is in correspondence with the mean score of 32.5% achieved in the previous research [3]. Statistical testing shows that the score distribution does not come from the normal distribution (p-value of normality test is less than 0.05, so we reject null hypothesis and conclude that the data are not from normal distribution), neither for the whole group nor for the sample groups (girls and boys, grammar schools/ vocational schools, different grade groups). As a result, the comparative statistics based on Mann-Whitney and Kruskal Wallis test was applied showing the following results. There is a significant difference between grammar schools and vocational schools’ results in favor of grammar schools. Boys achieved better results than girls. Also results of different grades were calculated. The null hypothesis: Mean ranks of two (or more) groups are equal was rejected (p-values
for the tests are less than 0.05). This means significant differences between grades, i.e. increasing grade indicates growing test gain.

**Table 3.** Basic statistics for the whole sample and for sample groups and comparisons between different sample groups.

| Skill group                  | Inquiry skill                                      | Number of test items | Type of school | Subject matter | Mean score individual items (in %) | Mean score groups of items (in %) | Gender | Test of normality (p-value) |
|------------------------------|----------------------------------------------------|----------------------|----------------|----------------|------------------------------------|------------------------------------|--------|-----------------------------|
| Conception, planning and design | Formulate hypothesis to be tested                  | 2                    | 1              | Ph             | 22.47                              | 20.46                              | p=0.000 | Z=-7.288                    |
|                              | Design experiment (identify variables and their relationship) | 4                    | 14             | Ch             | 18.45                              |                                    |         |                |
|                              |                                                    |                      | 2-3            | Ph             | 32.96                              | 32.10                              | p=0.000 | χ²(3)= 247.388             |
|                              |                                                    |                      | 5              | I              | 32.87                              |                                    |         |                |
|                              |                                                    |                      | 12             | Ph             | 34.01                              |                                    |         |                |
|                              |                                                    |                      | 15-16          | Ch             | 28.55                              |                                    |         |                |
| Analysis and interpretation   | Transform data to standard forms (i.e. tables or graphs) | 2                    | 6              | Ph/M           | 15.22                              | 14.85                              | p=0.201 | Z=-1.279                    |
|                              |                                                    |                      | 10             | I/M            | 14.49                              |                                    |         |                |
|                              | Determine relationship between variables (based on tables) | 2                    | 7              | Ph/M           | 15.58                              | 26.69                              | p=0.000 | Z=-28.902                   |
|                              |                                                    |                      | 11             | I/M            | 37.81                              |                                    |         |                |
|                              | Determine relationship between variables (based on graphs) | 2                    | 8              | Ph/M           | 23.42                              | 39.32                              | p=0.000 | Z=-37.219                   |
|                              |                                                    |                      | 13             | M              | 55.22                              |                                    |         |                |
|                              | Determine accuracy (identify possible sources of errors) | 2                    | 4              | Ph             | 26.63                              | 24.84                              | p=0.000 | Z=-4.553                    |
|                              |                                                    |                      | 9              | Bio            | 23.04                              |                                    |         |                |

The best results were achieved in the skill of determining the relationship based on graphs (39.32%) and design experiment (identify variables and their relationships, 32.10%) and the lowest results in the skill.
of formulating hypothesis to be tested (20.46%) and in the skill of transforming data to standard forms (14.85%). Both last poorest skills seem to be the worst regardless the gender, type of school and even the grade. Results for different skills and comparisons for different groups are shown in figures 2-5. It can be seen, that results in all particular skills for different groups (boys and girls, grammar and vocational schools, grades) are in accordance with the overall results.

The best results in all skill groups were achieved by the 4th grade students. The differences between the test score for different skills are statistically significant that was proved with the help of Friedman test ($p=0.000$, $\chi^2(5) = 4043.824$) and Bonferroni post hoc test except from the following pairs of skills: determine accuracy and formulate hypothesis, formulate hypothesis and determine relationships based on tables, design experiment and determine relationships based on graphs.

Nevertheless, the results indicate that students are likely involved more in the activities connected with identifying variables for the experimental design and also interpretation of data presented in graphs, however the skills to formulate hypothesis or transform data to graphs is much less developed. In figures 2, 3, 4, 5 the test results for different skills and the comparison of mean score with regard to gender, grade and type of school can be seen.
We have also tested differences between test items aimed at measuring the same skills, e.g. the four test items for designing experiment. Since the data did not follow the normal distribution, the differences were tested with the help of non-parametrical pair tests. For this particular skill, the Friedman test showed significant statistical differences between the four test items, while the Bonferroni post hoc test proved that differences are between all pairs of items. Statistically significant differences between the results of test items testing the same skill was shown for all tested skills except from the skill to transform data to standard forms (15.22%, 14.49%). These results indicate that the context of the test item and the way how the test item is formulated plays a role in the achieved test score.

Finally, in table 5 the results for two large skill groups Conception, planning and design and Analysis and interpretation are shown. The mean score for these two skill groups achieved the values of 26.22% and 26.28%. Statistical testing based on non-parametrical Wilcoxon Signed Rank pair test proved no statistical difference between these two groups, i.e. they are statistically equal.

| Skill group | Mean score (in %) | Wilcoxon Signed Rank test |
|-------------|------------------|---------------------------|
| Conception  | 26.22            | Z=-0.254                  |
| Analysis    | 26.28            | p=0.799                   |

3.2. Qualitative analysis

In the following section the examples of specific test items are analysed and discussed in detail with regard to most common revealed difficulties and misconceptions.

3.2.1. Skill to design experiment (identify variables). This test item (number 5 in the list, table 6) is connected with the informatics concepts. Nevertheless, it is testing the skill to identify independent and dependent variable (number of colours and image file size in this case) and control variables (resolution in dpi, image size in mm²) that is kept constant. Other variables (picture and file type) are irrelevant.

Table 6. Example of test item focusing on designing experiment (identifying variables).

| picture | File type | Image size (mm²) | Number of colours | Resolution (dpi) |
|---------|-----------|------------------|-------------------|------------------|
| school  | type 1    | 30 000           | 2⁸                | 200              |
| teacher | type 1    | 30 000           | 2¹⁶               | 300              |
| students| type 2    | 60 000           | 2⁸                | 300              |
| class   | type 2    | 30 000           | 2¹⁶               | 200              |

Supposing that neither the displayed picture nor the file type influences the image file size, decide which two image files students should compare.

a) Picture of school and teacher because the image size (mm²) and the file type should stay the same. 8.20%

b) Picture of school and students because the number of colours should stay the same. 14.48%

c) Picture of teacher and students because the resolution of images should stay the same. 17.97%

d) Picture of students and class because the file type should stay the same. 5.42%

e) Picture of school and class because the resolution of images and the image size (mm²) should stay the same. 32.87%
In this test item students achieved one of the highest score that is almost 33%. That means that almost 33% of students were able to correctly identify the independent, dependent, control or irrelevant variable. We also analysed other selected answers. Assuming that students are deciding about the answer on the basis of the introductory text and the analysis of information in the table, we can conclude about which variables students were able to recognize properly. The results of the analysis are summarized in table 7.

In the answers e) and c) students were able to identify all the variables involving independent, dependent and irrelevant variables. In the correct answer e) both control variables were identified properly, while in the answer c) almost 18% of students did not realize that it is not enough to keep the resolution constant. Surprisingly, 14.5% of students selected the answer b) that describes the independent variable (number of colours) as a variable that is kept constant even though we want to know how the number of colours affects the file size. That means that students were not able to distinguish between the independent and control variable in this case. The answers a) and d) states that comparing certain pictures the file type should stay the same. Even though it is explicitly stated that this variable does not play any role, altogether 13.62% of students selected these two answers. Moreover, the answer a) states that the dependent variable, i.e. the file size is kept constant. These results indicate that students have also reading comprehension problems that can significantly affect the achieved results.

| Answer | Control variable | Independent variable | Irrelevant variable | Dependent variable |
|--------|------------------|----------------------|---------------------|--------------------|
| a) (8.20%) | ✗              | ✓                   | ✗                   | ✗                  |
| b) (14.48%) | ✗              | ✗                   | ✓                   | ✓                  |
| c) (17.97%) | ✗              | ✓                   | ✓                   | ✓                  |
| d) (5.42%)  | ✗              | ✓                   | ✗                   | ✓                  |
| e) (32.87%) | ✓              | ✓                   | ✓                   | ✓                  |

3.2.2. Skill to transform data to standard forms (graphs). This test item (number 6 in the list) presents data of measuring resistance vs. temperature relationship. The test item was focused dominantly on transforming data from table to graph. They achieved the score of 15% that is one of the lowest score in the test.

The correct answer b) is highlighted in the box. In this test item students were expected firstly to understand the assignment and select variables that correspond with the investigated relationship. Reading the text with understanding is an inevitable assumption for choosing appropriate data columns from the table. This seems to be a problem for about 30.32% of students (incorrect answers a) and c)) that selected and analysed temperature vs. time or resistance vs. time instead of resistance vs. temperature relationship. Secondly, the selected table data should be transformed into the graph. We expected that students sketch the graph on the paper. To draw the graph first was quite important since the data in the last column were not dramatically different so that it was not so easy to draw the graph based on a quick glance. More than 36% of students selected correct variables on the horizontal and the vertical axis and they also identified the increasing tendency in the measured data (answers b) and e)). The further analysis required to compare equal changes of temperature with the changes of the corresponding resistance resulting in the conclusion that this relationship is linear. Finally, only 15.22% of students were able to draw the correct conclusion selecting the only correct answer b). More than 7.2% of students combined the correct answer b) with another incorrect answer, mostly d) (about 3%).

3.2.3. Skill to formulate hypothesis. In this test item (table 6, number 14 in the list) students were expected to select a correct hypothesis to test the research question if the temperature of water affects
the amount of sugar or salt dissolved in it. This test item had two correct answers as presented in table 9. Students were given 1 point only for selecting both correct answers. They were given 0.5 point in case of selecting only one correct answer, while if they combined the correct with another incorrect answer, no points were given in this case.

Table 8. Example of test item focusing on transforming data to standard forms (graphs).

Martin was exploring how electric resistance of copper wire changes with gradual increase of temperature. After certain time intervals he recorded time, temperature and resistance into the table (the table contains rounded values of temperatures and resistances). Based on the table select the graph representing the investigated relationship.

| Time (s) | Temperature (°C) | Resistance (Ω) |
|---------|-----------------|-------------|
| 10      | 30              | 16.0        |
| 15      | 35              | 16.9        |
| 35      | 40              | 17.8        |
| 40      | 50              | 19.6        |
| 50      | 70              | 23.2        |

In table 9 the score for the individual test answers as well as for the most common combinations of two answers can be seen. The mean score achieved in this test item was only 18,45% (1*8.01+0.5*17.81+0.5*2.74). Students’ common mistake involve selecting just one answer instead of two correct answers (e.g. answer c)).

Table 9. Example of test item focusing on formulating hypothesis.

Students wanted to find out if the temperature of water affects the amount of sugar or salt that would dissolve in it. They put equal amount of water of different temperature into two beakers. Then they add some sugar or salt and dissolved as much as they could by stirring. What is the hypothesis being tested?

| a) How temperature of water affects how sugar (salt) is dissolved in it? | 11.19% |
| b) The greater the amount of stirring, the greater the amount of sugar (salt) dissolved. | 3.10% |
| c) **The higher the temperature, the greater the amount of sugar (salt) dissolved.** | 17.81% |
| d) More sugar than salt is dissolved in 20°C water. | 1.20% |
| e) **The temperature of water affects the dissolving of sugar more than that of salt.** | 2.74% |

| a) and c) | 18.89% |
| a) and e) | 3.45% |
| c) and e) | 8.01% |
Common misconception in this case is also confusion between the hypothesis and the question. This is the case of almost 11.2% of students selecting answer a). 22.34% of students combine the answer a) with the correct answer c) or e). The results indicate that students have significant difficulties with understanding what hypothesis means.

4. Conclusion
In this study we focused on monitoring selected inquiry skills development, namely formulating hypothesis, designing experiment, transforming results into tables and graphs, determining relationships between variables based on tables and graphs and identifying possible sources of errors. More than 6600 students involved in the study answered the test of inquiry skills. Students achieved the mean score about 30%. This result is in correspondence with the previous research results [3] in which similar score was achieved on the sample of several hundreds of students. The boys achieved better results than girls, students of grammar schools were better than students of vocational schools and the test score increases with the grade. Even the highest score gained by the 4th grade students (36.43%) cannot be considered a positive result. Detailed analysis revealed the following facts. Best results were achieved in the skills to design experiment (identify variables and relationships) and to determine relationship between variables based on graphs. On the other hand, the results achieved for the skill to determine relationship between variables based on tables is much less. The worst results are connected with the skills to transform data to standard forms (graphs) and to formulate hypothesis. However, we are also aware of the limitation of this study related to the way how these results were obtained. The ability of multiple-choice testing instrument to reveal students’ thinking more deeply is limited and other research instruments, e.g. more open questions, interviews with students or even assignments requiring labwork could be helpful. On the other hand, the designed testing instrument was agreed to be a good option with regard to the huge number of students participating in the study. Nevertheless, the results gained in the study still indicate that students are unlikely involved in inquiry activities on a regular basis that is the inevitable assumption of inquiry skills development. This is a strong signal to schools and science teachers. Currently, the large national project IT Academy is focused on development of teaching and learning materials based on inquiry approach with strong emphasize on the development of inquiry skills. At the same time science teachers’ educational programme was designed and implemented. The innovative teaching and learning materials are trailed at schools by more than 900 science teachers. They are presented online and currently they are available for 204 schools all over Slovakia. At the end of the project (end of 2020) the level of inquiry skills development will be monitored again in order to compare the initial and final results. The further research in this field is expected in order to show the effect of these efforts.

Acknowledgment
This paper was made in the framework of the national project IT academy – Education for 21st century that is supported by European social fund and European regional development fund under the Operational Programme Human resources. This paper was also supported by the Scientific Grant Agency of the Ministry of Education of Slovak Republic and the Slovak Academy of Sciences – grant VEGA 1/0265/17 “Formative Evaluation in Teaching Sciences, Mathematics and Informatics”.

5. References
[1] Ješková Z, Lukáč S, Hančová M, Šnajder L, Guniš J, Balogová B and Kireš M 2016 Efficacy of inquiry-based learning in mathematics, physics and informatics in relation to the development of students’ inquiry skills J. of Baltic Sci. Educ. 15 pp 559-74
[2] Ješková Z, Lukáč S, Šnajder L, Guniš J, Balogová B and Kireš M 2016 Hodnotenie bádateľských zrušníc žiakov gymnázia Scientia in educatione 7 pp 48–70
[3] Ješková Z, Balogová B and Kireš M 2018 Assessing inquiry skills of upper secondary school students Proc. Int. Conf. on Physics Education Research-based proposals for improving physics teaching and learning – focus on laboratory August 30-September 3, 2016 Krakow ed D Sokolowska (Krakow: Jagiellonian University) pp 175–81
[4] Kask K and Rannikmäe M 2006 Estonian teachers’ readiness to promote inquiry skills among students J. of Baltic Sci. Educ. 5 pp 5–16
[5] Hart C, Mulhall P, Berry A, Loughran J and Gunstone R 2000 What is the purpose of this experiment? Or can students learn something from doing experiments? J. of Res.in Sci. Teach. 37 pp 655–75
[6] Rollnick M, Zwane S, Staskun M, Lotz S and Green G 2001 Improving pre-laboratory Preparation of first year university chemistry students Int. J. of Sci. Educ. 23 pp 1053–71
[7] Wenning C J 2007 Assessing inquiry skills as a component of scientific inquiry Journal of Physics Teacher Education Online 5
[8] Burns J C, Okey J R and Wise K C 1985 Development of an integrated process skill test: TIPSII J. of Res.in Sci. Teach. 22 pp 169–77
[9] Gormally C, Brickman P and Lutz M 2012 Developing a test of scientific literacy skills (TOSLS): Measuring undergraduates´ evaluation of scientific information and arguments Cell Biology Education – Life Science Education 11 pp 364–77
[10] Tamir P and Lunetta V N 1981 Inquiry-related tasks in high school science laboratory Science Education 65 pp 477–84
[11] Fradd S H, Lee O, Sutman F X and Saxton M K 2001 Promoting science literacy with English language learners through instructional materials development: A case study Bilingual Research Journal 25 pp 417–39
[12] Van den Berg E 2013 The PCK of laboratory teaching: Turning manipulation of equipment into manipulation of ideas Scientia in educatione 4 pp 74–92