

# THE IMPACT OF PROFESSIONAL DEVELOPMENT IN INQUIRY-BASED METHODS ON SCIENCE TEACHERS' CLASSROOM PRACTICE

**Paweł Bernard,  
Karol Dudek-Różycki**

## Introduction

Fundamentals of education system reform and a shift towards active, inquiry-based methods in the US can be found in the "A Nation at Risk" report (Gardner et al., 1983). This document called for revisions of the United States educational system, focusing primarily on teaching science and mathematics (Johnson, 2007). As a result, a new science curriculum was introduced by the National Research Council (1996), and an inquiry was established as an essential strategy for teaching the sciences; it was also argued that students should learn to focus on authentic questions generated from their experience. In Europe, a shift towards inquiry-based methods was announced almost 10 years later. The basis for the change was constituted by the European Union policy (EU, 2006) promoting the development of a knowledge society, to which the scientific education was expected to make a significant contribution. Moreover, the Rocard's report (Rocard, Csermely, Jorde, Walberg-Henriksson, & Hemmo, 2007) strongly advocated inquiry-based methods and suggested that the use of inquiry in science education can increase students' interest in the sciences and boost the number of graduates in the sciences, technical and mathematical studies. As a result, "Supporting and coordinating actions on innovation in the classroom: Dissemination and use of inquiry-based teaching methods on a large scale in Europe" was one of the key elements in the "Science in Society" area in the European Union's Seventh Framework Programme (FP7). Projects in this area were in progress from 2008 to 2017. Two major groups of the projects can be distinguished: I — those focused on a general introduction of Inquiry-Based Science Education (IBSE) and development of teaching materials (i.e., S-TEAM, Fibonacci Pathway, INQUIRE, SECURE, PROFILES, Primas); and II — those introducing inquiry in relation to specific conditions (for example, the ESTABLISH project) — IBSE, and building links with industry, SAILS, AssistMe (focusing on the assessment of students learning through inquiry) or IRRESISTIBLE — incorporating the idea of Responsible Research and Innovation into inquiry-based student activities (SCIENTIX, 2013). However, the implementation of the IBSE in Europe had a different character than in the US. A curriculum unambiguously defining the inquiry as a method that should be prominent in teaching the sciences has been adopted in the US, while in the EU, every country has its own curriculum, and inquiry plays a different role in each (McLoughlin, Finlayson, Van Kampen, & McCabe, 2013). Additionally, every country has its own history and tradition of education. As a result, at the beginning of realisation of the FP7 projects,



JOURNAL  
OF BALTIC  
SCIENCE  
EDUCATION

ISSN 1648-3898 /Print/  
ISSN 2538-7138 /Online/

**Abstract.** *Inquiry-based methods have become very popular in science education all over the world. In Europe, they were strongly advocated in years 2007-2013 within the 7th EU Framework Programme projects, and in that time, in many countries teacher educational programmes (TEP) in inquiry were developed and implemented. However, there has been limited information on how effective those trainings are, and how teachers bring theory into practice. Therefore, the aim of this research was to determine the impact of training in inquiry-based methods on the teachers' professional practices. The training consisted of two stages: I – one-week-long core training, and II – one-year-long extended support. Ninety-two science teachers participated in the research. It had a quantitative character and was based on a self-evaluation sheet that was completed three times: before the training, a month after the training, and a year after the training. The research results revealed current Polish science teachers' practices related to facilitating learning by inquiry and identified the inquiry's features that were immediately adopted by teachers, those that were adopted after a longer period, and finally, those that were not implemented at all. These results can be especially useful to educators who design and run TEPs in inquiry-based teaching and learning.*

**Keywords:** *inquiry-based learning, in-service teachers, professional development, science education.*

**Paweł Bernard, Karol Dudek-Różycki**  
Jagiellonian University, Poland



countries in which the IBSE was used for many years where educators and teachers knew and used this method, as well as countries in which this methodology was completely new may be found in Europe. While participating in the realisation of the aforementioned projects, the partners had to mutually determine how they understood inquiry, what its characteristic features were, and which inquiry elements they wanted to develop/promote. Many of those projects used the definition of IBSE proposed by Linn, Davis and Bell (2004), who stated that inquiry is the: '*...intentional process of diagnosing problems, critiquing experiments, and distinguishing alternatives, planning investigations, researching conjectures, searching for information, constructing models, debating with peers, and forming coherent arguments. (p. 4)*'. This definition is in line with the National Research Council (1996) guides, where students engaged in inquiry '*...describe objects and events, ask questions, construct explanations, test those explanations against current scientific knowledge, and communicate their ideas to others. (p. 2)*'. Apart from the definition, invaluable contributions from Bybee (2000; 2002) should be noted, where among others, three descriptions of inquiry for teachers were provided and the 5E inquiry cycle was described. Those elements were used in almost all of the 7FP EU projects.

Observing the history of the global IBSE implementation, one might think that both the process itself and its effects have been thoroughly studied already. However, despite the fact that education experts have extensively described their experience connected with teachers' training on the application of inquiry and have published the results of completed case studies, the extent to which training in the IBSE influences teachers' practices and students' achievement is still unclear (Capps, Crawford, & Conostas, 2012; Ermeling, 2010). In this research, an attempt was made to examine the effect of the teacher training in the IBSE on their professional practices. The training encompassed by the research was carried out within the framework of the SAILS FP7 project. This project was realised in 12 European countries from 2012 to 2015, and during its realisation, more than 2500 of pre-service and in-service teachers participated in the trainings (Jönsson et al., 2015). The aim of the project was to prepare the teachers to teach through inquiry as well as have the confidence and competence in the assessment of their students' learning.

The described research was carried out among in-service teachers participating in the project in Poland. In the Polish curriculum, the inquiry-based methods are not listed among the required/recommended teaching methods unambiguously, however, advised teacher/student actions and skills to be developed are in line with the competencies that are taught with IBSE (Polish Core Curriculum, 2008). The designed and implemented teacher education programme (TEP) was based on the framework developed within the SAILS project. The research carried out among the participants of the training was meant to answer a general question: "What is the effect of the training on the actual teachers' school practices?" The authors of the TEP and training instructors had experience in training teachers in inquiry, which was acquired during the realisation of the previous FP7 project – ESTABLISH. Since the effectiveness of the training can be influenced by teachers' beliefs and attitudes towards the trained method (Arce, Bodner, & Hutchinson, 2014; Luft, 2001; Windschitl, 2004), basic studies in this area were carried out among Polish teachers earlier (Bernard, Maciejowska, Krzeczowska, & Odrowąż, 2015; Bernard, Maciejowska, Odrowąż, Dudek, & Geoghegan, 2012), and they revealed a generally positive teachers' attitude towards the inquiry-based methods.

### Teacher Training

Design of the teacher's professional development programme is a major challenge. Many factors have to be considered, and even subtle aspects can determine the success of the training and its impact on the teacher's classroom practice (Taitelbaum, Mamluk-Naaman, Carmeli, & Hofstein, 2008). Inspiration while designing presented below programme was Kegan's model of development (Kegan, 1994), which states that when a person changes, he doesn't just change '*... what he knows, but the way he knows... (p. 17)*'; moreover, working with adult learners we are '*... not merely asking them to take on new skills... [but] 'asking them to change the whole way they understand themselves, their world, and the relationship between the two. (p. 275)*'. Attempting to reach this goal, the ways teachers learn may be more similar to the way students learn than it was previously recognised (Lieberman & Miller, 1992). Loucks-Horsley, Hewson, Love and Stiles (1998) claimed that science and mathematics teachers '*...need to experience for themselves the science and mathematics learning they will want their students to do. (p.13)*'. Furthermore, Harrison, Howard and Matthews (2016) suggested that teachers need to experience any reform of student's teaching as learners themselves because durable changes in pedagogy only occur when the teachers understand what their learners experience in their classroom. Darling-Hammond and McLaughlin (1995) claimed that the teachers learn by doing, reading and reflecting (as students do); by collaborating with other teachers; by looking closely at students and their work; and by sharing what they see. To understand the process deeply, teachers must learn about, see, and experience successful learning-centred teaching. Thus, there is no doubt that the teachers' professional development (PD) programme



should actively engage participants and treat teachers as learners rather than as information gatherers. Research has shown that such professional development programmes are effective in changing teachers' knowledge and practice (Brooks & Brooks, 1993; Bybee, 1993; Layman, Ochoa, & Heikkinen, 1996; Sparks & Hirsh, 1997). Several sets of teachers' PD programmes characteristics can be found in the literature, three of which are presented in Table 1.

**Table 1**

*Selected sets of effective professional development programme characteristics*

Effective professional development programme		
<i>According to Garet, Porter, Desimone, Birman and Yoon (2001) summarised by Capps et al., (2012):</i>	<i>According to Loucks-Horsley et al., (1998) summarised by Shepardson and Harbor (2004):</i>	<i>According to Posnanski (2002):</i>
<ul style="list-style-type: none"> <li>• focuses on content knowledge;</li> <li>• provides opportunities for active learning;</li> <li>• connects to or is coherent with other activities;</li> <li>• engages teachers in reform-based PD;</li> <li>• promotes collective participation of teachers;</li> <li>• provides an adequate amount of time.</li> </ul>	<ul style="list-style-type: none"> <li>• provides a well-defined image of effective classroom learning and teaching;</li> <li>• is based on experiences that provide teachers with the opportunity to build knowledge and skills;</li> <li>• is based on activities that model the strategies teachers will use with students;</li> <li>• includes the establishment of a learning community;</li> <li>• involves teachers in leadership roles;</li> <li>• links to other parts of the education system;</li> <li>• ensures a continuous assessment to make improvements;</li> <li>• ensures a positive impact on the teacher's effectiveness, student's learning, leadership, and the school community.</li> </ul>	<ul style="list-style-type: none"> <li>• is a long-term programme;</li> <li>• includes a group of teamed teachers from individual schools or school districts;</li> <li>• includes activities that model the theory or rationale behind the design and implementation of curricula and innovative teaching strategies;</li> <li>• includes local curricular resources embedded within the programme activities;</li> <li>• uses a constructivist approach with participant experimentation with instructional strategies and curricula/activity design;</li> <li>• includes opportunities for reflection (self-reflection and obtaining data on the effectiveness of activity implementation in their classrooms as well from their students) to promote the identification and modification of self-efficacy beliefs and teaching behaviours.</li> </ul>

Apart from active learning and experiencing the new method, the element that appears to be crucial for the teachers' PD is the opportunity to discuss the changes they might make in their practice and to understand what is expected from them (Cordingley, Bell, Thomason, & Firth, 2005; Griffin, 1983; Harrison et al., 2016; Harrison, Hofstein, Eylon, & Simon, 2008; Mule, 2006). In this context, a Learning Community becomes an important element. Garet et al. (2001) argued that the fact that teachers were communicating about practice and collective participation (e.g., teacher teams from the same school, grade, subject) influenced the degree to which teachers changed their practice. Luft (2001) stated that professional development opportunities should contain follow-up experiences with multiple opportunities for interaction. Additionally, Vescio, Ross and Adams (2008) indicated that the research on the relationship between PD based on a Learning Community and student's achievements clearly demonstrated a strong positive connection.

While developing a PD programme in the IBSE, it should be considered that success in inquiry science teaching requires that a teacher understand the nature of scientific inquiry and inquiry-based learning. There are three aspects of inquiry that were identified by Anderson (2002): how science is done (scientific inquiry), how science is taught (inquiry teaching), and how science is learned (inquiry learning). Therefore, the teachers must understand how scientists develop concepts and gain knowledge, and how to model these processes with students (e.g., observe, collect, and analyse data) so they could obtain own scientific understandings. Additionally, the teachers have to be able to recognise multiple forms of inquiry and deconstruct inquiry into essential parts (Blanchard et al., 2009; Wenning, 2005). Therefore, they must gain essential inquiry skills: formulate scientifically oriented questions and hypotheses, plan and carry out investigations, record and process data, formulate explanations based on evidence, evaluate their explanations, and communicate precisely and effectively (Chin & Osborne, 2008; Gyllenpalm, Wickman, & Holmgren, 2010; Jeanpierre, Oberhauser, & Freeman, 2005; National Research Council, 1996; Taitelbaum, Mamluk-Naaman, Carmeli, & Hofstein, 2008). Teachers can gain these skills and experience scientific inquiry through participation in authentic research. This approach had a positive impact not only on teachers' abilities to use inquiry-based teaching techniques but also on their content knowledge (Bazler, 1991; Caton, Brewer, & Brown, 2000; Loucks-Horsley et al., 1998; Luft, 2001; Westerlund, García, Koke, Taylor, & Mason, 2002). Unfortunately, including



teachers in real scientific research can be problematic (Abd-El-Khalick et al., 2004; Lustick, 2009). Capps et al. (2012) analysed 22 PD programmes in the IBSE, and only five included elements of an authentic inquiry experience, and in most cases, teachers participated only in small parts of ongoing research and were not able to investigate their own inquiry questions. Therefore, instead of an authentic inquiry experience, a model-based inquiry can be used. In this approach, the educators have to provide general tasks/problem areas that are interesting and challenging for the teachers so they can engage in their own investigations and experience the entire inquiry process from the perspective of the learners. Moreover, it is preferable that the problems used be adapted to the school curriculum and used by the teachers in the classroom.

### *Teacher Education Programme*

Within the aforementioned SAILS project, a general framework of the teacher education programme was developed, but each partner was able to customise the form and the length of the training to his or her own needs (Jöhnson, Lundström, Finlayson, McLaughlin, & McCabe, 2014). In every case, the training consisted of two stages: I – core training and II – extended support in the form of a Community of Practice (CoP) (Coakley, 2013). In Poland, the core training included an intensive 5-day training, encompassing 33 hours of contact classes, and the extended support included participation in the CoP for at least one year.

#### *The Day-By-Day Core Training Programme*

**Day 1.** The training started with an introductory lecture, discussing the fundamentals of constructivism and IBSE methodology, as well as depicting the IBSE elements contained in the valid curriculum. Next, the self-evaluation tool was presented to the teachers, enabling them to independently discuss the extent of their use of the IBSE (which IBSE aspect they are using). Then, the teachers participated in workshops intended mainly to set forth the scientific inquiry cycle. To this end, the teachers performed a simple hands-on investigation, in which they were to answer the following general question: “Which infant diapers are the best?” After carrying out the exercise, the teachers analysed the undertaken actions and independently developed the operational pathway for the scientific research.

**Day 2.** The day began with a workshop on research planning — aspects of proper formulations of inquiry questions, hypotheses and research procedure plans. The next classes had a laboratory character and utilised a model-based inquiry. The teachers were presented with general problems that they were to face, and the teachers independently formulated the inquiry questions and hypotheses, as well as planning and carrying out their experiments. The teachers worked in groups of 3-4 persons; they were grouped independently but were directed to create teams that specialising in teaching of one school science subject on the selected education level. Therefore, the educators selected the appropriate general issues for every group. These issues were chosen from a material package developed within the framework of the project (Finlayson et al., 2015a, 2015b). During the laboratory classes, the teachers were tasked with obtaining results that were processed by the groups during subsequent classes that had a workshop character. The final product of each investigation was constituted by a laboratory report.

**Day 3.** The training began with another laboratory class, during which the teachers carried out investigations as on the previous day but with more advanced issues with a set of measurement sensors. As on the previous day, workshop classes, during which the teachers processed the results of measurements, took place after the laboratory classes. The last item of the agenda of this day consisted of a workshop, during which the teachers practised adaptation of the investigations learned to the school conditions. Its aim was to learn about various forms of inquiry. To this end, teachers had to prepare instructions for a specific exercise in various forms according to the classification by Wenning (2005).

**Day 4.** The day was dedicated to the problems with the evaluation of students learning by inquiry. The training began with a lecture on “Fundamentals of formative assessment”, presenting the theoretical foundations of this assessment form and connecting its use with the IBSE and the valid curriculum. Then, during workshop classes, the teachers developed assessment tools and strategies for exercises that they had carried out on previous days. They used the developed tools for a peer evaluation based on the laboratory reports prepared earlier.

**Day 5.** This day was dedicated to microprojects, during which the teachers were tasked with planning inquiry-based classes, which could be carried out in their schools. They were to choose the issue, prepare instructions according to



the chosen IBSE type, and develop evaluation tools and strategies. In the end, the teachers presented their projects to each other and discussed their pros and cons. In the end, each of the teachers was tasked with writing a letter to themselves with plans for the introduction of the IBSE in their schools. The teachers put these letters in envelopes, which were closed, addressed and given to the educators. After approximately six months (after the summer break), the letters were sent back to them, which was to remind them of the goals they chose.

In summary, during the core training, the teachers participated in 4 hours of lectures, 12 hours of laboratory classes, and 17 hours of workshops.

### *Summary of Extended Support*

Stage II of the training covered extended support that was realised by participation in the Community of Practice. The community gathered educators and teachers from various training groups on the dedicated social media platform. On the platform, didactic materials from the core training were shared. The educators initiated and moderated discussions, encouraged teachers to run inquiry-based activities, and published new didactic materials, which could be pilot-implemented by the teachers, etc. The teachers had the opportunity to communicate with each other and with the educators, ask questions, and work together on lesson plans and instructions. After a year, the teachers were invited to participate in a national conference, where they had an opportunity to present their experiences and develop the didactic materials and results of case studies conducted during the pilot implementations of the didactic materials, tools, teaching and assessment strategies. Those with the best presentations were invited to participate in an international conference of teachers participating in the project (Melia, McLaughlin, McCabe, & Finlayson, 2015).

## **Research Methodology**

### *General Background*

The training in IBSE for in-service Polish science teachers within the framework of the SAILS project was carried out in 2013, 2014, and 2015. The first group (2013) pioneered the realisation of the developed training programme, and after the training was completed, adjustments to the programme and the materials used were introduced. Therefore, the 2013 group was excluded from the research. In the following years, the programme was not changed further. The training was voluntary and free of charge, the teachers had full participation costs paid, together with accommodation and board. Teachers applied for the training independently, and acceptance was on a first-come, first-served basis. Before the registration date, information about the recruitment, dates and participation terms were publicly available for a month.

### *Participants*

Participation in the training did not automatically translate to participation in the research; the participation was voluntary and anonymous. Every participant was aware of the data to be collected, the goal of their collection, and the mode of their processing, according to the Jagiellonian University ethical standards. The participants were able to renounce their participation in the research at any stage. Teachers who started the training in the two following years participated in the research: 2014 ( $N = 60$ ) and 2015 ( $N = 46$ ), total  $N = 106$ . The group comprised 93 females and 13 males, including 33 chemistry teachers, 39 biology teachers, and 34 physics teachers (in cases where the teachers were certified to teach two or more subjects, they declared one main subject, most often in accordance with their master's degree). Only the data obtained from the teachers who participated in all stages of the research and returned all completed forms were used in the analysis. Therefore, the final research group included 92 persons, 57 in 2014 and 35 in 2015; subgroups according to the taught subject: biology – 31; chemistry – 35; and physics – 26. The gender of the responders who had their filled forms analysed is not known because of the anonymous character of the research.

### *Instrument and Procedures*

The research was based on the Self-Reflection Tool for Teachers developed within the framework of the Fibonacci FP7 Project (Harlen, 2012). It contains a list of 38 indicators divided into 7 categories to judge the imple-



mentation of inquiry-based teaching through a self-analysis of classroom practices. The categories were grouped in three sections:

Section A: The Teacher's Role

- (1) Building on the students' ideas (3 questions)
- (2) Supporting the students' own investigations (7 questions)

Section B: Students' Activities

- (3) Guiding analysis and conclusions (7 questions)
- (4) Carrying out investigations (9 questions)
- (5) Working with others (5 questions)

Section C: Student's records

- (6) Students documenting their own work (1 question)
- (7) Aspects of student's written records (6 questions)

The indicators, or criteria, are expressed as questions that teachers can ask themselves concerning a sequence of science activities that were intended to enable students to learn through inquiry. Every indicator is accompanied by examples of good practices intended as an aid for the teacher, facilitating proper understanding of the question. Further columns contain answers (a choice of YES, NO, NA) and empty cells for recording notes when a teacher needed to supplement the answer with an additional comment. Every positive (affirmative) answer meant that the responder proceeded according to the guidelines contained in the work cycle in the IBSE methodology, so the higher the number of positive answers was, the closer the methods used by them for the IBSE were. It should be noted here that in the original approach, the questionnaire used was for a teacher's self-evaluation for specific classes. In the case of this research, the self-evaluation pertained to the teacher's practice in general. The change significantly affected the interpretation of one of the available answers – NA. In the original approach, this option should be selected by a teacher when the situation is improper in the context of the classes, e.g., the indicators in the D category — "Carrying out investigations" pertain to carrying out experiments/observation, but only formulation of inquiry questions and hypotheses, and preparation of the research plan are scheduled during the classes. In the general approach, in which the aforementioned questionnaire was used in the research described, the "NA" option should be selected only when a given situation does not occur during the classes, e.g., because of logistical limitations. The teachers were also asked to comment briefly on the selection of this option in the "Notes" column. One should remember that the teachers participating in the research worked at various education levels, so general statements have been included in the questions requiring proper interpretation by the teachers. For instance, the term "prediction" has been used in the indicators, but the teachers working at the upper secondary school level should equate this term with the hypothesis. To facilitate understanding and interpretation of the questionnaire by the participants, it was thoroughly discussed by the educators before commencing the research, and the responders had instructions describing the form and goal of the questionnaire available in detail. This instruction manual was also available to the responders in the subsequent stages of the research whenever the questionnaire was completed.

The research was divided into three stages; at every stage, the teachers completed the same questionnaire sheet: stage I – directly before the training, stage II – shortly after the core training (after 3-4 weeks), and stage III – approximately one year after the core training. The questionnaire was completed anonymously; however, to ensure the possibility of comparing the answers given by the teachers in the subsequent research stages, the teachers signed the form with a pseudonym. Thus, the persons coding and analysing the results had no access to the personal data of the responders. Because of the distant date of stage III of the research (approx. 1 year after the stage I), at stage I, the teachers recorded their pseudonyms on sheets of paper that were then placed by them in closed envelopes. The teachers were provided with these envelopes together with the stage III questionnaire, enabling those who did not remember their pseudonyms to check their selections.

### *Data Analysis*

The completed questionnaires were coded in a spreadsheet. The following coding system was used: YES – 1, NA – 0, and NO – -1. Then, the numbers of YES answers were counted for every survey participant for the whole test, for each category of questions, and for every question individually, and they were recalculated as percentages. The procedure at the subsequent research stages was analogical.



The first step of the analysis consisted of checking the equivalence of the groups (2014 and 2015) participating in the research. To this end, a series of comparative analyses of variances (ANOVA) (Howell, 2002) were carried out for the two groups, both for the average results of all research stages together and for every stage separately. The ANOVA necessary conditions of application (King, Rosopa, & Minium, 2011) were examined using a Shapiro-Wilk test (Shapiro & Wilk, 1965) for a normal distribution condition, and a Levene test (Levene, 1960) for homogeneity of variances condition. In cases when the normal distribution condition was not met, an analysis with a nonparametric test was carried out in addition using the *U* test (Mann & Whitney, 1947).

Further analyses aimed to check whether the professional practices of the teachers before and after the training differ and whether the effect is maintained over time (1 year after the training) based on the variance analysis for the repeated measurements. Before this analysis, the variance sphericity assumption was checked with Mauchly's test (Mauchly, 1940). In cases when the sphericity condition was not met, the results were verified using the lower epsilon correction ( $\epsilon$ ). The ANOVA analyses for the repeated measurements were carried out for the whole questionnaire, for each of the 7 question categories, and for every question separately. Additionally, post hoc tests were carried out using the HSD Tukey's test. The significance level was defined as  $\alpha = .05$  for this research. The statistical analysis was performed using the 'Statistica' software package, version 13.

## Research Results

The average results of the survey achieved by the groups participating in the research are presented in Table 2.

**Table 2**

*Average results for the whole questionnaire at subsequent survey stages, divided into groups*

Group	Stage	Average score [%]	SD	CI	
				-.95	.95
2014	All	74.7	1.9	70.9	78.4
	I	67.1	2.4	62.4	71.8
	II	76.6	2.7	71.3	82.0
	III	80.2	2.2	75.9	84.5
2015	All	78.5	2.4	73.7	83.3
	I	72.0	3.0	66.0	78.0
	II	80.1	3.4	73.3	86.9
	III	83.3	2.7	77.9	88.8

The differences between the results achieved by the training participants in two comparable groups were analysed. The comparative test results and conditions for use of the applied statistical methods are presented in Table 3.

**Table 3**

*Comparison of results achieved by the 2014 and 2015 groups by statistical tests*

Test	<i>p</i>			
	All stages	Stage I	Stage II	Stage III
Shapiro-Wilk	.00552	.07876	<.00001	.00002
Levene	.93583	.24321	.94207	.36845
ANOVA	.21869	.20533	.43026	.37316
Mann-Whitney	.21240	.26128	.26581	.37047



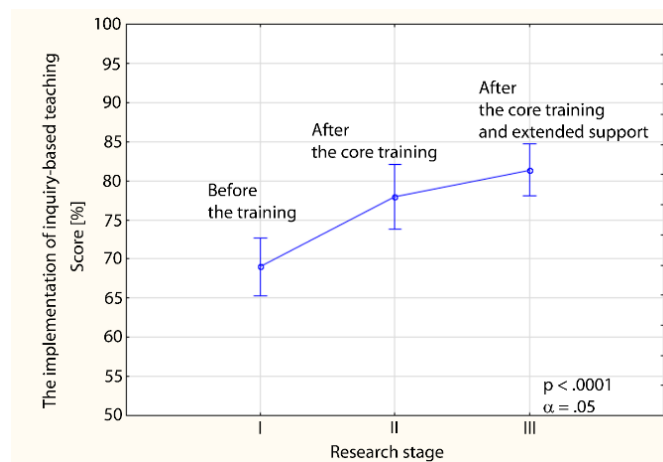
The results of the Levene test indicated that the conditions of the variation homogeneity of the analysed results were met in all cases. However, the results of the Shapiro-Wilk test indicated that the distribution of the results was not consistent with the normal distribution. Hence, an additional comparison using a nonparametric test (Mann-Whitney test) was carried out apart from the ANOVA analysis. The results obtained by both methods were convergent.

Comparing the results of the 2014 and 2015 groups, both completed, and those from the individual research stages, a shift of the average results achieved by the 2015 group towards higher values was observed. However, with reference to the assumed significance level of  $\alpha = .05$ , the statistical tests carried out indicated that the differences were statistically not significant. Hence, further analyses were carried out for all survey participants altogether without dividing them into groups.

Another analysis aimed to check whether the average survey results differ among the subsequent research stages. A comparison of the average results carried out using the variation analysis with repeated measurements for the whole questionnaire is shown in Table 4 ('All questions', first row). The results of the Mauchly's test and the post hoc tests (Tukey's HSD) are also presented in Table 4. The obtained results indicated meeting of the variation sphericity condition and that the observed increase in the values of the average results in subsequent survey stages was statistically significant. The post hoc test results showed that the difference between stages I and II of the research was statistically significant and the further increase in the results average between stages II and III was not statistically significant.

**Figure 1**

*The average result for the whole questionnaire at subsequent survey stages, with the 95% confidence interval marked*



Apart from the analysis of the results for the whole questionnaire, the analogous results for each of the categories separately are shown in Table 4 (a group of questions 1-7). It should be noted that the sphericity condition was not met in categories 6 and 7. Hence, the ANOVA results were verified using the epsilon lower correction, and the result of this test is reported in parentheses under the ANOVA result.





**Table 4**

The results for the whole questionnaire and for the individual survey areas at the subsequent survey stages, and statistical comparison of the results together with the post hoc tests

Group of questions	Survey stage	Average score [%]	SD	CI		Mauchly's test	p		
				-95%	95%		ANOVA (€)	Tukey vs. I	Tukey vs. II
All questions	I	68.99	1.87	65.28	72.71	.2034	<.0001	<.0001	.1999
	II	77.95	2.10	73.77	82.12				
	III	81.38	1.69	78.03	84.73				
1. Building on the students' ideas	I	88.04	2.03	84.00	92.08	.1074	.0110	.0286	.9876
	II	94.20	1.97	90.29	98.11				
	III	94.57	1.57	91.45	97.68				
2. Supporting the students' own investigations	I	69.88	2.46	64.98	74.77	.5168	<.0001	.0002	.7006
	II	80.75	2.39	75.99	85.50				
	III	82.92	1.94	79.06	86.77				
3. Guiding analysis and conclusions	I	66.46	2.72	61.06	71.86	.1689	<.0001	<.0001	.1543
	II	79.50	2.40	74.73	84.27				
	III	84.63	1.85	80.95	88.30				
4. Carrying out investigations	I	63.16	2.56	58.08	68.25	.9954	<.0001	.0382	.0160
	II	70.77	2.90	65.02	76.53				
	III	79.35	2.31	74.75	83.94				
5. Working with others	I	68.04	2.87	62.35	73.74	.0775	.0047	.0160	.9728
	II	76.09	2.80	70.52	81.66				
	III	76.74	2.82	71.14	82.34				
6. Students documenting their own work	I	84.78	3.77	77.30	92.26	.0001	.1963	.9571	.3358
	II	85.87	3.65	78.62	93.12				
	III	91.30	2.95	85.44	97.17				
7. Aspects of student's written records	I	68.30	3.15	62.05	74.54	<.0001	.0477	.0524	.9168
	II	75.72	3.30	69.16	82.29				
	III	74.46	3.10	68.31	80.61				

The obtained results indicated statistically significant differences for areas 1-5 (marked with bold). For area 6, the change in the average results was statistically not significant, which was indicated by the ANOVA results and the result of the test using the epsilon correction. The results for the last category in the ANOVA test indicated a statistically significant change, but the fact that the sphericity condition was not met and the result of the test using the correction showed that the results were falsely positive and that the corrected results should be considered firm, which indicated a lack of statistically significant differences. Analysing the post hoc tests results, one may see that in all cases for which a statistically significant difference between the results was found, the change was connected to the differences between stages I and II and stages I and III of the research. Only in the case of area 4 a statistically significant difference between stages II and III was observed.

The next step consisted of a significance analysis of the change in the average result at subsequent research stages for every question separately. The goal of this analysis was to indicate specific actions leading to a change in the classroom practice of the teachers. A procedure analogous to the analysis of question groups was applied. The results are collected in Table 5.



**Table 5**

The results for every question separately at the subsequent survey stages, and statistical comparison of the results together with the post hoc tests (\*Ps means students/pupils)

Question	Survey stage	Average score [%]	SD	CI		p			
				-95%	95%	Mauchly's test	ANOVA (€)	Tukey vs. I	Tukey vs. II
1a. Did you ask questions to reveal and show interest in Ps ideas?	I	90.22	3.11	84.03	96.40				
	II	93.48	2.59	88.34	98.62	.0844	.1051	.5340	
	III	96.74	1.86	93.04	100.44			.0827	.5340
1b. Did you help Ps to express their ideas clearly?	I	90.22	3.11	84.03	96.40				
	II	95.65	2.14	91.41	99.90	.0236	.2308	.2961	
	III	95.65	2.14	91.41	99.90		(.2272)	.2961	1.0000
1c. Did you give Ps positive feedback on how to review or take their ideas further?	I	83.70	3.87	76.00	91.39				
	II	93.48	2.59	88.34	98.62	.0009	.0535	.0527	
	III	91.30	2.95	85.44	97.17		(.0879)	.1673	.8634
2a. Did you encourage Ps to ask questions?	I	59.78	5.14	49.57	69.99				
	II	79.35	4.24	70.92	87.78	.0781	.0015	.0018	
	III	76.09	4.47	67.20	84.97			.0120	.8356
2b. Did you help them formulate productive (investigable) questions?	I	55.43	5.21	45.09	65.78				
	II	76.09	4.47	67.20	84.97	.4095	.0005	.0017	
	III	76.09	4.47	67.20	84.97			.0017	1.0000
2c. Did you ask them to make predictions?	I	94.57	2.38	89.84	99.29				
	II	94.57	2.38	89.84	99.29	.0047	.9267	1.0000	
	III	93.48	2.59	88.34	98.62		(.7832)	.9390	.9390
2d. Did you involve them in planning investigations?	I	68.48	4.87	58.80	78.15				
	II	72.83	4.66	63.56	82.09	.6552	.0328	.7445	
	III	83.70	3.87	76.00	91.39			.0281	.1598
2e. Did you encourage them to include fair testing where appropriate?	I	53.26	5.23	42.87	63.65				
	II	66.30	4.95	56.46	76.15	.8292	.0655	.0859	
	III	65.22	4.99	55.30	75.13			.1268	.9830
2f. Did you ask them to check their results or observations?	I	72.83	4.66	63.56	82.09				
	II	89.13	3.26	82.65	95.61	.0848	.0001	.0013	
	III	91.30	2.95	85.44	97.17			.0002	.8867
2g. Did you help them to keep notes and record results systematically?	I	84.78	3.77	77.30	92.26				
	II	86.96	3.53	79.94	93.97	.3615	.0467	.8575	
	III	94.57	2.38	89.84	99.29			.0459	.1539
3a. Did you ask Ps to provide some conclusions from their work?	I	89.13	3.26	82.65	95.61				
	II	91.30	2.95	85.44	97.17	.0324	.3288	.8231	
	III	94.57	2.38	89.84	99.29		(.2929)	.2975	.6456
3b. Did you ask Ps to check that their conclusions were consistent with their results?	I	65.22	4.99	55.30	75.13				
	II	89.13	3.26	82.65	95.61	.0157	<.0001	<.0001	
	III	88.04	3.40	81.29	94.80		(.0003)	<.0001	.9747



3c. Did you ask Ps to compare their conclusions with their predictions?	I	81.52	4.07	73.44	89.60				
	II	84.78	3.77	77.30	92.26	.2010	.0386	.7756	
	III	93.48	2.59	88.34	98.62			.0341	.1659
3d. Did you ask Ps to think of reasons or explanations for what they found?	I	67.39	4.91	57.63	77.15				
	II	81.52	4.07	73.44	89.60	.3933	.0006	.0225	
	III	88.04	3.40	81.29	94.80			.0003	.4417
3e. Did you ask Ps to identify possible sources of error?	I	60.87	5.12	50.71	71.03				
	II	73.91	4.60	64.77	83.06	.4316	<.0001	.0209	
	III	84.78	3.77	77.30	92.26			<.0001	.0675
3f. Did you ask Ps to identify further questions?	I	48.91	5.24	38.50	59.32				
	II	59.78	5.14	49.57	69.99	.4712	.0868	.2308	
	III	63.04	5.06	52.99	73.09			.0847	.8758
3g. Did you encourage Ps to reflect on what they found and how they found it?	I	52.17	5.24	41.77	62.58				
	II	76.09	4.47	67.20	84.97	.0844	<.0001	.0001	
	III	80.43	4.16	72.17	88.70			<.0001	.7248
4a. Did Ps work on questions which they identified as their own, even though introduced by you?	I	55.43	5.21	45.09	65.78				
	II	70.65	4.77	61.17	80.13	.8201	.0011	.0375	
	III	78.26	4.32	69.67	86.85			.0007	.4370
4b. Did Ps make predictions based on their ideas?	I	52.17	5.24	41.77	62.58				
	II	60.87	5.12	50.71	71.03	.4496	.0001	.3501	
	III	79.35	4.24	70.92	87.78			.0001	.0093
4c. Did Ps take part in planning an investigation?	I	68.48	4.87	58.80	78.15				
	II	68.48	4.87	58.80	78.15	.1390	.0002	1.0000	
	III	86.96	3.53	79.94	93.97			.0009	.0009
4d. Did Ps include 'fair testing' in their plan if appropriate?	I	45.65	5.22	35.28	56.02				
	II	55.43	5.21	45.09	65.78	.4233	.2114	.2753	
	III	55.43	5.21	45.09	65.78			.2753	1.0000
4e. Did Ps carry out an investigation themselves?	I	73.91	4.60	64.77	83.06				
	II	84.78	3.77	77.30	92.26	.7815	.0614	.0980	
	III	84.78	3.77	77.30	92.26			.0980	1.0000
4f. Did Ps gather data using methods and sources appropriate to the inquiry question?	I	70.65	4.77	61.17	80.13				
	II	80.43	4.16	72.17	88.70	.3789	.0314	.2112	
	III	85.87	3.65	78.62	93.12			.0239	.6176
4g. Did the data gathered enable Ps to test their predictions?	I	67.39	4.91	57.63	77.15				
	II	71.74	4.72	62.36	81.12	.9098	.0512	.7407	
	III	81.52	4.07	73.44	89.60			.0434	.2203
4h. Did Ps consider their results in relation to the inquiry question?	I	72.83	4.66	63.56	82.09				
	II	73.91	4.60	64.77	83.06	.5050	.1909	.9813	
	III	82.61	3.97	74.72	90.50			.2177	.2993
4i. Did Ps propose explanations for their results?	I	61.96	5.09	51.85	72.07				
	II	70.65	4.77	61.17	80.13	.2768	.0277	.3662	
	III	79.35	4.24	70.92	87.78			.0187	.3662



5a. Did Ps collaborate with others during group work?	I	90.22	3.11	84.03	96.40				
	II	94.57	2.38	89.84	99.29	.4485	.5070	.5005	
	III	91.30	2.95	85.44	97.17			.9576	.6772
5b. Did Ps engage in class or group discussions of their investigations and explanations?	I	65.22	4.99	55.30	75.13				
	II	76.09	4.47	67.20	84.97	.3867	.1353	.1386	
	III	73.91	4.60	64.77	83.06			.2814	.9235
5c. Did Ps report their work in some form to the whole class?	I	70.65	4.77	61.17	80.13				
	II	81.52	4.07	73.44	89.60	.7430	.0303	.0720	
	III	82.61	3.97	74.72	90.50			.0417	.9738
5d. Did Ps listen to each other during reporting?	I	69.57	4.82	59.98	79.15				
	II	70.65	4.77	61.17	80.13	.4718	.0531	.9780	
	III	81.52	4.07	73.44	89.60			.0697	.1102
5e. Did Ps respond to each other during reporting?	I	44.57	5.21	34.22	54.91				
	II	57.61	5.18	47.32	67.90	.7700	.0693	.0653	
	III	54.35	5.22	43.98	64.72			.2140	.8418
6a. Did Ps make some record of what they did and found?	I	84.78	3.77	77.30	92.26				
	II	85.87	3.65	78.62	93.12	.0001	.1963	.9571	
	III	91.30	2.95	85.44	97.17		(.2032)	.2083	.3358
7a. Did Ps include in their record a clear statement of the inquiry question or problem?	I	73.91	4.60	64.77	83.06				
	II	73.91	4.60	64.77	83.06	.1260	1.0000	1.0000	
	III	73.91	4.60	64.77	83.06			1.0000	1.0000
7b. Did Ps' records indicate what data were collected and how they were collected?	I	81.52	4.07	73.44	89.60				
	II	80.43	4.16	72.17	88.70	.8604	.9676	.9731	
	III	80.43	4.16	72.17	88.70			.9731	1.0000
7c. Did Ps record observations and collect data in a systematic way?	I	67.39	4.91	57.63	77.15				
	II	70.65	4.77	61.17	80.13	.0001	.8372	.8278	
	III	68.48	4.87	58.80	78.15		(.6742)	.9792	.9194
7d. Did Ps indicate in their records whether results agreed with their predictions?	I	55.43	5.21	45.09	65.78				
	II	66.30	4.95	56.46	76.15	.1659	.1425	.1193	
	III	61.96	5.09	51.85	72.07			.4634	.7101
7e. Did Ps state their conclusions in their record?	I	70.65	4.77	61.17	80.13				
	II	80.43	4.16	72.17	88.70	.5200	.1476	.1467	
	III	78.26	4.32	69.67	86.85			.3121	.9090
7f. Did Ps make personal notes during their work?	I	60.87	5.12	50.71	71.03				
	II	82.61	3.97	74.72	90.50	.0634	<.0001	.0003	
	III	83.70	3.87	76.00	91.39			.0001	.9789

Among the analysed questions, 18 indicated significant differences between the subsequent research stages. The post hoc tests showed that in 9 cases, statistically significant difference between stages I and II was observed, and in 7 cases, difference between stages I and III was observed. Additionally, in 2 cases, when the difference between stages I and III was evident, a difference was also statistically revealed between stages II and III. It is noteworthy that no clear falling tendency of the average of the answers to questions in the subsequent survey stages was observed after the training. In general, the results for the individual questions may be divided into 5 groups:



### A. Statistically significant changes

- A1. Change between stages I and II and stages I and III – change occurring directly after the training and stable in time. Questions: 2a, 2b, 2f, 3b, 3d, 3e, 3g, 4a, and 7f.
- A2. Change only between stages I and III – the change occurred after a longer period; increasing trend was constant, but statistically significant differences were between extreme stages of the research. Questions: 2d, 2g, 3c, 4f, 4g, 4i, and 5c.
- A3. The change between stages I and III and stages II and III – a small change directly after the training but significant in a longer period. Questions: 4b and 4c.

### B. No statistically significant changes

- B1. The changes were not statistically significant, and the result in stage I was high (>85%). Questions: 1a, 1b, 2c, 3a, 5a, and 6a.
- B2. No changes or the changes were not statistically significant. Questions: 1c, 2e, 3f, 4d, 4e, 4h, 5b, 5d, 5e, 7a, 7b, 7c, 7d, and 7e.

### *Analysis of the Results*

First, it should be noted that the analysis of the combined results (whole questionnaire) indicate that the training has affected the professional practices of the teachers, resulting in increase of IBSE aspects applied during lessons. Moreover, this effect is stable over time, which is indicated by the results of the research stage III, carried out after one year after the core training. A constant increase in the range of IBSE implementation is evident in the subsequent stages, the observed difference is statistically significant between stages I and II and stages I and III, however, the increase between stages II and III is not statistically significant. Thus, the teachers are able to apply the new work methods relatively quickly and consequently use them in the future. Further analyses aimed to identify the areas of the teacher's operation in which the IBSE elements are used and specific actions undertaken by the teachers.

Breaking up overall results into categories and further dividing the categories into specific questions detailed the areas and specific teacher actions that have changed significantly as well as the dynamics of these changes. The actions of the teachers in area 1, "Building on students' ideas" change directly after the training, and the observed change is statistically significant. Unfortunately, it is not possible to indicate which of the actions has had the strongest impact on the observed change because the independent analysis of questions from this category does not indicate any statistically significant differences for any of the questions. However, it should be noted that this area has been characterised by the highest average of affirmative answers in stage I (before the training). More than 90% of the teachers have ascertained that in their practice, they ask questions to reveal and show interest in student's ideas and that they help them to express their ideas clearly. Slightly fewer teachers have ascertained that they give students (Ps) positive feedback on how to review or take their ideas further. In relation to this action, the value of the change before and after the training is closest to be a statistically significant. It is likely that the high average percentage of answers in this area both before and after the training prevent meeting the variation sphericity condition for 2 of 3 questions in this category.

In category 2, "Supporting student's own investigations", a general result at a lower level of approx. 70% in stage I was observed. The change observed in next stages is similar to area 1; however, if the questions in this area are analysed separately, it can be noted for 5 of 7 questions, statistically significant differences have been observed in the subsequent research stages. However, the course of the changes is not identical for all questions. Within a short time after the training, significantly more teachers started encouraging students to ask questions (Q 2a) as well as formulate inquiry questions (Q 2b) and check their results and observations (Q 2f). Engaging students in experiment planning (Q 2d) and engaging themselves in the process of systematic creation of notes by students (Q 2g) required more time from the teachers to be introduced. On the other hand, no statistically significant differences were found for question (Q 2c) pertaining to making predictions. However, this question may be included in the B1 group, i.e., a constant high result, because the result amounted to approx. 95% both before and after the training. For the question pertaining to encouraging students to carry out fair testing (Q 2e), a 12% increase is evident between stages I and II; however, this difference is not statistically significant at the adopted significance level.

The course of the changes for area 3, "Guiding analysis and conclusions", is also analogous with areas 1 and 2, but the extent of the change is larger (~20%) than in the previous cases. Statistically significant changes were



already found at stage II for questions on whether the teacher was asking Ps to check that their conclusions were consistent with their results (Q 3b), to consider reasons or explanations for what they found (Q 3d), to identify possible sources of error (Q 3e) and to encourage them to reflect on what they found and how they found it (Q 3g). For questions pertaining to asking Ps to compare their conclusions with their predictions (Q 3c), the change was evident not earlier than between stages I and III. In the case of two questions for which no statistically significant differences were found, the course of the changes is similar to the case of area 2. Answers to questions 3a, "Did you ask Ps to provide some conclusions from their work?" were ~90% affirmative before the training, which is why the observed increase by ~5% was not statistically significant, which was similar to question 2c. A change in the results for question 3f, "Did you ask Ps to identify further questions?", is analogous to 2e, where the effect found is too small to be statistically significant.

Category 4, "Carrying out investigations", includes the highest number of questions ( $n = 9$ ). It is the only area for which a statistically significant increase in all subsequent research stages is observed, which is not surprising because this area included questions on the independent experimental work of students and the introduction of modifications to experiments carried out independently requires time. Apart from time, experience gained by the students in subsequent classes realised by inquiry may be an important factor contributing to the constant increase of results. One could conclude that a significant change between stages I – II and II – III in this category means that teachers significantly changed their practice directly after training and applied further significant changes in time. A subsequent analysis of separated questions revealed that such an approach is untrue, and the result is a consequence of several other effects. The teachers were able to introduce student's work on questions, which they identified relatively quickly (Q 4a), and only in this case is a difference evident between stages I and II of the research. More time was required to include predictions based on students' ideas (Q 4b) and planning an investigation (Q 4c) by students into practice. It should be noted that in these cases, there were no differences shortly after training between I and II survey stages, but a statistically significant change occurred between stages II and III. These are the only cases of such a type in the whole research. In other cases, when a long time was necessary for change implementation, the change had a more linear character, and significant differences were noticed between stages I and III. Such a situation refers to students gathering data using methods and sources appropriate for the inquiry question (Q 4f); the data gathered enabled Ps to test their predictions (Q 4g), and the students proposed explanations for their results (Q 4i). Additionally, question 4g, for which the analysis of variance returned a value close to the limit value  $\alpha$  and the post hoc test indicated a difference between stages I and III, is an interesting case. For questions 4d, "Did Ps include 'fair testing' in their plan if appropriate?" and 4e, "Did Ps carry out an investigation themselves?", an increase in the results is evident between stages I and II; however, this increase is not statistically significant. The results for question 4h, "Did Ps consider their results in relation to the inquiry question?", change slightly differently because the change is noticeable between stages II and III. However, the extent of the change is too small to be statistically significant. Therefore, even though statistically significant changes were between stages I – II and II – III for the entire category, such a change was not noticed for the individual questions.

Modernisation of work in the group, which was the subject of the research in area 5, proved to be a major challenge for the teachers. This area included 5 questions. The analysis indicated statistically significant differences for this area; however, while analysing the questions individually, a statistically significant difference (between stages I and III) was evident only for one, Q 5c – "Did Ps report their work in some form to the whole class?". For two questions (Q 5a and 5d), the value of the analysis of variance was close to the value of  $\alpha$ . Question 5e, "Did Ps respond to each other during reporting?", is an interesting case. This question obtained the lowest number of affirmative answers (approx. 45%) in stage I. An approx. 14% increase in the use of this element is evident between stages I and II, followed by a decrease by 4% in stage III. Thus, in the first moment, the teachers attempted to adopt this inquiry element in their practices but abandoned it later. However, it should be noted that the observed changes are not statistically significant.

The last two categories pertained to the creation of notes by the students. Area 6 contained only one question, so the results of this question are identical to the result for the whole area. Additionally, it should be noted that this question is general and was specified in category 7. For question 6a, an ~7% increase in the number of affirmative answers between stages II and III is evident, but this difference is not statistically significant. However, it should be noted that even before the training, 85% of the teachers answered affirmatively that the students are recording what they did and found.

The last category, area 7, included 6 questions. Statistically significant differences between stages I – II and I – III were observed for only one question, 7f: "Did Ps make personal notes during their work?" The results for questions



7a "Did Ps include in their record a clear statement of the inquiry question or problem?" and 7b "Did Ps' records indicate what data were collected and how they were collected?" are interesting because practically no changes are observed for these questions in any of the research stages, and the obtained values are constant at the level of 75% and 80%, respectively. This means that during the training, the teachers did not change or attempt to change their practices in this area. The results for the other questions of category 7 (Q 7c, 7d and 7e) are developing differently. For all these questions, an increase between stages I and II is evident, followed by a decrease between stages II and III, approximately at the initial level. For these questions, it may be concluded that the teachers have attempted to introduce changes in their practice but have abandoned them in a longer time period.

The results may also be analysed from the point of view of development of specific skills. Considering the skills chronologically, according to a typical inquiry cycle, it should be started with an analysis pertaining to the formulation of inquiry questions. Four questions are related to this aspect directly: 2a, 2b, 4a, and 3f, where the teachers were to answer whether they were encouraging students to ask questions and formulate inquiry questions and whether the students were working based on these questions. The analysis of the answers indicates that as a result of the training, the teachers changed their practice to a significant degree. For all these questions, an ~20% change is observed, and in almost all cases, it is a statistically significant change. This change was found directly after the training or between stages, I – II, and the change of practice was permanent, which was indicated by the differences between stages I and III. The question pertaining to the formulation of further inquiry questions was the only question for which no statistically significant changes were observed. Therefore, the teachers changed their practice in this scope in the investigation preparation stage, and more rarely, they introduced development of the skill of inquiry questions formulation to their practice after a completed inquiry and the so-called closure of the scientific inquiry cycle.

In analysing the results in the context of making predictions, the answers of two questions (Q 2c and 4b) may be compared, in which the teachers were to indicate whether they asked students to make predictions and whether they are making predictions for their ideas. Considering the first aspect — making predictions — the lack of a statistically significant change is observed, but it should be noted that this question may be classified in the B1 group — high-end results. Thus, no change was found, but practically all teachers used this element in their professional practices before the training. In this area, the teachers enhanced their professional practices with predictions by the students based on their ideas. The change in this aspect is statistically significant and occurs in a time interval, or between stages II and III.

Next, questions pertaining to experiment planning (Q 2d, 2e, 4c, 4d) may be analysed. With regard to this skill, two aspects may be distinguished: 1 – pertaining to engagement of students in experiment planning (Q 2d and 4c) and 2 – pertaining to application of fair testing in experiment plans. In the first aspect, a statistical change is evident for both questions; consequently, the change occurred between survey stages II and III. No statistically significant changes were observed for area 2. Moreover, the observed changes in the average results for both questions are analogous, an ~10% increase is observed between stages I – II, and the average value is constant between stages II – III.

Analysing the realisation of an exercise, an increase in affirmative answers indicated a larger share of independent execution of experiments by students (Q 4e) between survey stages I and II and a constant level of the result between stages II and III, although, the observed changes are not statistically significant. The same skillset includes collection of data, which serves to answer the inquiry question or verify the formulated predictions (Q 4f and 4g). For both questions, an increase in the average value is observed in the subsequent research stages. However, the general ANOVA test indicates a lack of statistically significant changes, but it is worth mentioning that in both cases, the probability coefficient value for the post hoc tests between stages II and III amounts to  $p < \alpha$  in one case and  $p \sim \alpha$  in the other one.

Considering the development of skills connected with drawing conclusions (Q 3a, 3g, 4h 3d, 4i), it may be noted that statistically significant changes occurred for numerous elements in this scope. Admittedly, no change was found in the first question from this group (Q 3a), but this question may be classified in the B1 group or constant high-end results. Then, it may be seen that the teachers were encouraging students to reflect on what they found and how they found it (Q 3g) and to consider reasons or explanations for what they found more frequently (Q 3d, 4i). Only in case of the question pertaining to the relationship between the results and the inquiry question (Q 4h) the change is not statistically significant. Moreover, the observed increase in the results occurred between stages II and III.

In the area of the skills connected with evaluation of the experiments, statistically significant changes were also found (Q 2f, 3b, 3e, and 3g). Noteworthy the changes occurred directly after the core training (between stages I and II), and the results in stage III were similar to these in stage II.

Work in groups and taking notes are the next skills which may be considered; however, in these cases, the skills



coincide with question categories (5, 6 and 7), so a separate analysis is not necessary.

## Discussion

The goal of the undertaken training was to expand the range of methods used by the teachers with inquiry-based methods, not a complete change of their teaching styles. Application of the inquiry elements is recommended by the Polish core curriculum, but as opposed to the USA (National Research Council, 1996), the inquiry is not expected to be a leading/prevaling method. Thus, the aim of the training was to provide the teachers with the knowledge and skills necessary to use various aspects and types of inquiry. Implementation of inquiry elements in traditional instructions by the teachers is not an atypical behaviour. Such actions were observed, among others, by Johnson (2007) and Lotter, Harwood, & Bonner (2006), who also analysed the impact of the IBSE training on the professional practices of the teachers.

It is not a surprise that participation in the training affected the practice of the teachers significantly, since similar results have been reported earlier (Blanchard, Southerland, & Granger, 2009; Capps et al., 2012; Ermeling, 2010; Jeanpierre et al., 2005; Luft, 2001; Radford, 1998; Shepardson & Harbor, 2004; Taitelbaum et al., 2008). Luft (2001), Shepardson and Harbor (2004) carried out quantitative studies which indicated that the most likely change in the practices of the teachers pertains to the following areas: teacher as a guide, cooperative learning, inquiry questions and designing, and conducting investigations. Therefore, the results obtained in this research coincide to a significant degree with previous results. However, the detailed analysis carried out allowed for a more precise description of changes in teachers' and students' actions. Moreover, presented research yielded information on the dynamics of these changes which had not been reported earlier.

The described research was based on a self-reflection tool for teachers. Tools of such a type have their limitations. Frechtling, Sharp, Carey and Vanden-Kieman (1995) argued that the teacher self-report data alone may not reflect what is happening in the classroom. In the context of PD in the inquiry, it was confirmed by Lee, Hart, Cuevas and Enders (2004), who used a mixed approach, and it was found that the self-reported data conflicted with classroom observations. On the other hand, Jeanpierre et al., (2005) confirmed the consistency of the teacher self-report data with observations. In the case of the presented research, extent of the training (number of participants and timescale) did not allow to confront the obtained data with observations. It should be also remembered that the limited number of pre- and post-classroom observations may not accurately represent the day-to-day nature of a teacher's practice (Capps et al., 2012). Another limitation of the presented research consisted of the fact that the applied tool practically omitted issues related to the assessment of the students. It was due to the fact that the SAILS project, within the framework of which the described training and research were carried out, was oriented on the development of methods for evaluation of students working by inquiry, therefore, this aspect was considered a separate project task. The developed assessment strategies and didactic materials were implemented and evaluated by the teachers, the results of those operations were described as separate case studies (Bernard, Dudek, & Orwat, 2019; Finlayson et al., 2015a, 2015b).

## Conclusions and Implications

The research revealed current Polish science teachers' practices related to facilitating learning by inquiry. The teachers claim to practise with students making predictions, drawing conclusions, independent making notes, they ask questions to reveal student's ideas and to help them express their ideas clearly, they organize work in groups during lessons even without prior training in IBSE. This set of practices can be treated as a starting point for the IBSE training. However, the fact that teachers use certain elements of inquiry doesn't mean that they are doing it correctly, according to the IBSE methodology. Therefore, these actions should be still discussed during the training. On the other hand, the fact that teachers already use these elements suggests, that they would keep doing it after the training. This fact was confirmed by the obtained results, none of the practices was abandoned after the training. It can be also expected that new knowledge and skills acquired during the training provide higher value and deeper meaning of these practices. Diagnosed changes in teachers' actions suggest that practically immediately after the training they incorporate a number of new elements of inquiry into their practice (a difference between research stages I and II). This group includes actions connected with formulating research inquiries, predictions and reflections on the obtained results by students, as well as evaluation of investigations by them. Moreover, inquiry elements the introduction of which requires more time were identified. This group of practices includes involving





students in experiment planning, students' work based on their own ideas, independent data collection, making notes and presenting results. Therefore, it can be concluded that the long-term training system consisting of core training and extended support is well-justified. Basing on the obtained results, constituent elements of the extended support can be planned, along with didactic materials and tasks formulated by tutors, to stimulate teachers' actions at adequate phases of the training. After the core training, tutors may want to first focus on these aspects of inquiry that were earlier used by teachers, and those that can be incorporated easily. Such a bottom-up strategy, with a number of easily reachable goals, can be a positive motivation for teachers. Subsequently, the aspects of inquiry-based learning whose incorporation require more time should be enhanced.

From the point of view of persons constructing and running the PD programme, the most important group of questions describes practices that were not adopted by the teachers during the research time, or the change was not statistically significant, and the results in research stage I have not been high (the B2 questions group). These encompass planning and use of fair testing, identification of further research questions, carrying out an investigation by students themselves and elements of group work, including discussions of the results and various aspects of written records. These aspects should be extensively discussed during the core training to ensure their meaning and deep-seated understanding. Moreover, it is advised to intensively enhance classroom application of these aspects during the extended support later phase, along with other elements of inquiry that require more time to be incorporated.

### Acknowledgements

The research was based on the teacher training within the SAILS FP7 project (Strategies for Assessment of Inquiry Learning in Science). The SAILS project has received funding from the European Commission Seventh Framework Programme [FP7/2007-2013] under grant agreement number 289085 and support from the Polish Ministry of Science and Higher Education [2887/7PR/2013/2], funds for science in the years 2013-2015 for co-financed international projects.

### References

- Abd-El-Khalick, F., BouJaoude, S., Duschl, R., Lederman, N. G., Mamlok-Naaman, R., Hofstein, A., Niaz, M., Treagust, D., Tuan, H. (2004). Inquiry in science education: International perspectives. *Science Education*, 88(3), 397-419. <https://doi.org/10.1002/sce.10118>
- Anderson, R. D. (2002). Reforming science teaching: What research says about inquiry. *Journal of Science Teacher Education*, 13(1), 1-12. <https://doi.org/10.1023/A:1015171124982>
- Arce, J., Bodner, G. M., & Hutchinson, K. (2014). A study of the impact of inquiry-based professional development experiences on the beliefs of intermediate science teachers about 'best practices' for classroom teaching. *International Journal of Education in Mathematics, Science and Technology*, 2(2), 85-95.
- Avalos, B. (2011). Teacher professional development in teaching and teacher education over ten years. *Teaching and Teacher Education*, 27(1), 10-20. <https://doi.org/10.1016/j.tate.2010.08.007>
- Bazler, J. A. (1991). A middle school teacher summer research project. *School Science and Mathematics*, 91(7), 322-324. <https://doi.org/10.1111/j.1949-8594.1991.tb12110.x>
- Bernard, P., Dudek, K., & Orwat, K. (2019). Integration of inquiry-based instruction with formative assessment: The case of experienced chemistry teachers. *Journal of Baltic Science Education*, 18(2), 184-196. <https://doi.org/10.33225/jbse/19.18.184>
- Bernard, P., Maciejowska, I., Krzeczowska, M., & Odrowąż, E. (2015). Influence of in-service teacher training on their opinions about IBSE. *Procedia – Social and Behavioral Sciences*, 177, 88-99. <https://doi.org/10.1016/j.sbspro.2015.02.343>
- Bernard, P., Maciejowska, I., Odrowąż, E., Dudek, K., & Geoghegan, R. (2012). Introduction of inquiry-based science education into polish science curriculum – general findings of teachers' attitude. *Chemistry-Didactics-Ecology-Metrology*, 17(1-2), 49-59. <https://doi.org/10.1023/A:1015171124982>
- Blanchard, M. R., Southerland, S. A., & Granger, E. M. (2009). No silver bullet for inquiry: Making sense of teacher change following an inquiry-based research experience for teachers. *Science Education*, 93(2), 322-360. <https://doi.org/10.1002/sce.20298>
- Brooks, J., & Brooks, M. (1993). *The case for constructivist classrooms*. Association for Supervision and Curriculum Development.
- Bybee, R. W. (1993). *Reforming science education: Social perspectives and personal reflections*. Teacher's College Press.
- Bybee, R. W. (2000). Teaching science as inquiry. In J. Minstrell & E. H. Van Zee (Eds.), *Inquiring into inquiry learning and teaching in science* (pp. 20-26). American Association for the Advancement of Science.
- Bybee, R. W. (2002). *Learning science and the science of learning*. NSTA Press.
- Capps, D. K., Crawford, B. A., & Constanas, M. A. (2012). A review of empirical literature on inquiry professional development: Alignment with best practices and a critique of the findings. *Journal of Science Teacher Education*, 23(3), 291-318. <https://doi.org/10.1007/s10972-012-9275-2>
- Caton, E., Brewer, C., & Brown, F. (2000). Building teacher-scientist partnerships: Teaching about energy through inquiry. *School Science and Mathematics*, 100(1), 7-15.



- Chin, C., & Osborne, J. (2008). Students' questions: A potential resource for teaching and learning science. *Studies in Science Education*, 44(1), 1-39. <https://doi.org/10.1080/03057260701828101>
- Coakley, P. (2013). *Report on the organisation of the Community of Practice (CoP) in each participating country*. SAILS project. <http://sails-project.eu/sites/default/files/outcomes/d5-2.pdf>
- Cordingley, P., Bell, M., Thomason, S., & Firth, A. (2005). The impact of collaborative continuing professional development (CPD) on classroom teaching and learning. Review: How do collaborative and sustained CPD and sustained but not collaborative CPD affect teaching and learning? In *Research evidence in education library*. EPPI-Centre, Institute of Education, University of London.
- Darling-Hammond, L., & McLaughlin, M. W. (1995). Policies that support professional development in an era of reform. *Phi Delta Kappan*, 76(8), 597-604.
- Ermelg, B. A. (2010). Tracing the effects of teacher inquiry on classroom practice. *Teaching and Teacher Education*, 26(3), 377-388. <https://doi.org/10.1016/j.tate.2009.02.019>
- EU (2006). Recommendation of the European Parliament and of the Council. *Official Journal of the European Union*, L394, 10-16.
- Finlayson, O., McLaughlin, E., Coyle, E., McCabe, D., Lovatt, J., & Van Kampen, P. (Eds.). (2015a). *SAILS Inquiry and Assessment Units Volume 1*. SAILS project. [http://sails-project.eu/sites/default/files/outcomes/SAILS\\_units\\_volume-1.pdf](http://sails-project.eu/sites/default/files/outcomes/SAILS_units_volume-1.pdf)
- Finlayson, O., McLaughlin, E., Coyle, E., McCabe, D., Lovatt, J., & Van Kampen, P. (Eds.). (2015b). *SAILS Inquiry and Assessment Units Volume 2*. SAILS project. [http://sails-project.eu/sites/default/files/outcomes/SAILS\\_units\\_volume-2.pdf](http://sails-project.eu/sites/default/files/outcomes/SAILS_units_volume-2.pdf)
- Frechtling, J., Sharp, L., Carey, N., & Vanden-Kiernan, N. (1995). *Teacher enhancement programs: A perspective on the last four decades*. National Science Foundation.
- Gardner, D., Larsen, Y., Baker, W., Campbell, A., Crosby, E., Foster, Jr., C., Wallace, R. (1983). *A nation at risk: The imperative for educational reform. A report to the Nation and the Secretary of Education United States Department of Education by The National Commission on excellence in education*. U.S. Government Printing Office.
- Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, 38(4), 915-945. <https://doi.org/10.3102/00028312038004915>
- Griffin, G. (1983). Introduction: The work of staff development. In G. Griffin (Ed.), *Staff development, eighty-second yearbook of the national society for the study of education*. University of Chicago Press.
- Gyllenpalm, J., Wickman, P., & Holmgren, S. (2010). Teachers' language on scientific inquiry: Methods of teaching or methods of inquiry? *International Journal of Science Education*, 32(9), 1151-1172. <https://doi.org/10.1080/09500690902977457>
- Harlen, W. (2012). *Tools for enhancing inquiry in science education*. (S. Borda Carulla, Ed.). Fibonacci Project. [www.fondation-lamap.org/sites/default/files/upload/media/minisites/action\\_internationale/1-tools\\_for\\_enhancing\\_inquiry\\_in\\_science\\_education.pdf](http://www.fondation-lamap.org/sites/default/files/upload/media/minisites/action_internationale/1-tools_for_enhancing_inquiry_in_science_education.pdf)
- Harrison, C., Hofstein, A., Eylon, B.-S., & Simon, S. (2008). Evidence-based professional development of science teachers in two countries. *International Journal of Science Education*, 30(5), 577-591. <https://doi.org/10.1080/09500690701854832>
- Harrison, C., Howard, S., & Matthews, B. (2016). Crafting a teacher education program to support inquiry-learning in science: the SAILS project. In *Electronic Proceedings of the ESERA 2015 Conference. Science education research: Engaging learners for a sustainable future* (Vol. In-service science teacher education, continued professional development). University of Helsinki.
- Howell, D. (2002). *Statistical methods for psychology* (5th ed.). Thomson Learning.
- Jeanpierre, B., Oberhauser, K., & Freeman, C. (2005). Characteristics of professional development that effect change in secondary science teachers' classroom practices. *Journal of Research in Science Teaching*, 42(6), 668-690. <https://doi.org/10.1002/tea.20069>
- Johnson, C. C. (2007). Whole-school collaborative sustained professional development and science teacher change: Signs of progress. *Journal of Science Teacher Education*, 18(4), 629-661. <https://doi.org/10.1007/s10972-007-9043-x>
- Jöhnson, A., Lundström, M., Finlayson, O., McLaughlin, E., & McCabe, D. (2014). *Report on IBSE teacher education and assessment programme – stage 1*. SAILS project. <http://sails-project.eu/sites/default/files/outcomes/d4-2.pdf>
- Jönsson, A., Lundström, M., Finlayson, O., McLoughlin, E., Lovatt, J., Van Kampen, P., & McCabe, D. (2015). *Report on IBSE teacher education and assessment programme – stage 2*. SAILS project. <http://sails-project.eu/sites/default/files/outcomes/d4-3.pdf>
- Kegan, R. (1994). *In over our heads: The mental demands of modern life*. Harvard University Press.
- King, B. M., Rosopa, P. J., & Minium, E. W. (2011). *Statistical reasoning in the behavioral sciences* (6th ed.). John Wiley & Sons.
- Layman, J. W., Ochoa, G., & Heikkinen, H. (1996). *Inquiry and learning*. The College Entrance Examination Board.
- Lee, O., Hart, J. E., Cuevas, P., & Enders, C. (2004). Professional development in inquiry-based science for elementary teachers of diverse student groups. *Journal of Research in Science Teaching*, 41(10), 1021-1043. <https://doi.org/10.1002/tea.20037>
- Levene, H. (1960). Robust tests for equality of variances. In *Contributions to Probability and Statistics: Essays in honor of Harold Hotelling*. (pp. 278-292). Stanford University Press.
- Liberman, A., & Miller, L. (1992). Teacher development in professional practice schools. In *Professional practice schools: Linking teacher education and school reform* (pp. 105-123). Teachers College Press.
- Linn, M. C., Davis, E. A., & Bell, P. (2004). *Internet environments for science education*. Lawrence Erlbaum Associates.
- Lotter, C., Harwood, W. S., & Bonner, J. J. (2006). Overcoming a learning bottleneck: Inquiry professional development for secondary science teachers. *Journal of Science Teacher Education*, 17(3), 185-216. <https://doi.org/10.1007/s10972-005-9002-3>
- Loucks-Horsley, S., Hewson, P. W., Love, N., & Stiles, K. E. (1998). *Designing professional development for teachers of science and mathematics*. Corwin Press.



- Luft, J. A. (2001). Changing inquiry practices and beliefs: The impact of an inquiry-based professional development programme on beginning and experienced secondary science teachers. *International Journal of Science Education*, 23(5), 517-534.
- Lustick, D. (2009). The failure of inquiry: Preparing science teachers with an authentic investigation. *Journal of Science Teacher Education*, 20(6), 583-604. <https://doi.org/10.1007/s10972-009-9149-4>
- Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, 18(1), 50-60. <https://doi.org/10.1214/aoms/1177730491>
- Mauchly, J. W. (1940). Significance test for sphericity of a normal  $n$ -variate distribution. *The Annals of Mathematical Statistics*, 11(2), 204-209. <https://doi.org/10.1214/aoms/1177731915>
- McLoughlin, E., Finlayson, O., Van Kampen, P., & McCabe, D. (2013). *Report on how IBSE is involved in national curricula and assessment in the participating countries*. SAILS project. <http://sails-project.eu/sites/default/files/outcomes/d1-2.pdf>
- Melia, M., McLaughlin, E., McCabe, D., & Finlayson, O. (2015). *Report on the activities of the Community of Practice (CoP) II*. SAILS project. <http://sails-project.eu/sites/default/files/outcomes/d5-5.pdf>
- Mule, L. (2006). Preservice teachers' inquiry in a professional development school context: Implications for the practicum. *Teaching and Teacher Education*, 22(2), 205-218. <https://doi.org/10.1016/j.tate.2005.09.011>
- National Research Council. (1996). *National science education standards*. National Academy Press.
- Polish Core Curriculum. (2008). Act of the Polish Parliament. Regulation of the Minister of Education. Dz.U. 2008 Nr 4, poz. 17. Podstawa programowa z komentarzami, tom 5. Edukacja przyrodnicza w szkole podstawowej, gimnazjum i liceum.
- Posnanski, T. J. (2002). Professional development programs for elementary science teachers: An analysis of teacher self-efficacy beliefs and a professional development model. *Journal of Science Teacher Education*, 13(3), 189-220.
- Radford, D. (1998). Transferring theory into practice: A model for professional development for science education reform. *Journal of Research in Science Teaching*, 35(1), 73-88. [https://doi.org/10.1002/\(SICI\)1098-2736\(199801\)35:1<73::AID-TEA5>3.0.CO;2-K](https://doi.org/10.1002/(SICI)1098-2736(199801)35:1<73::AID-TEA5>3.0.CO;2-K)
- Rocard, M., Csermely, P., Jorde, D., Walberg-Henriksson, H., & Hemmo, V. (2007). *Science education now: A renewed pedagogy for the future of Europe*. European Communities.
- SCIENTIX. (2013). The community for science education in Europe. <http://www.scientix.eu/web/guest/projects>
- Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3-4), 591-611. <https://doi.org/10.1093/biomet/52.3-4.591>
- Shepardson, D. P., & Harbor, J. (2004). ENVISION: The effectiveness of a dual-level professional development model for changing teacher practice. *Environmental Education Research*, 10(4), 471-492. <https://doi.org/10.1080/1350462042000291010>
- Sparks, D., & Hirsh, S. (1997). *A new vision for staff development*. Association for Supervision and Curriculum Development. National Staff Development Council.
- Taitelbaum, D., Mamlok-Naaman, R., Carmeli, M., & Hofstein, A. (2008). Evidence for teachers' change while participating in a continuous professional development programme and implementing the inquiry approach in the chemistry laboratory. *International Journal of Science Education*, 30(5), 593-617. <https://doi.org/10.1080/09500690701854840>
- Vescio, V., Ross, D., & Adams, A. (2008). A review of research on the impact of professional learning communities on teaching practice and student learning. *Teaching and Teacher Education*, 24(1), 80-91. <https://doi.org/10.1016/j.tate.2007.01.004>
- Wenning, C. J. (2005). Levels of inquiry: Hierarchies of pedagogical practices and inquiry processes. *Journal of Physics Teacher Education Online*, 2(3), 3-12.
- Windschitl, M. (2004). Folk theories of "inquiry:" How preservice teachers reproduce the discourse and practices of a theoretical scientific method. *Journal of Research in Science Teaching*, 41(5), 481-512. <https://doi.org/10.1002/tea.20010>

Received: November 12, 2019

Accepted: March 10, 2020

Cite as: Bernard, P., & Dudek-Różycki, K. (2020). The impact of professional development in inquiry-based methods on science teachers' classroom practice. *Journal of Baltic Science Education*, 19(2), 201-219. <https://doi.org/10.33225/jbse/20.19.201>

**Paweł Bernard**  
(Corresponding author)

PhD, Assistant Professor, Jagiellonian University, Department of Chemical Education, 2 Gronostajowa Str. 30-387 Kraków, Poland.  
E-mail: pawel.bernard@uj.edu.pl  
Website: <http://www.zdch.uj.edu.pl/bernard>  
ORCID: [orcid.org/0000-0002-8618-3447](https://orcid.org/0000-0002-8618-3447)

**Karol Dudek-Różycki**

PhD, Assistant, Jagiellonian University, Department of Chemical Education, 2 Gronostajowa Str. 30-387 Kraków, Poland.  
E-mail: karol.dudek@uj.edu.pl  
ORCID: [orcid.org/0000-0003-3518-2089](https://orcid.org/0000-0003-3518-2089)

