JOURNAL OF BALTIC SCIENCE EDUCATION
ISSN 1648-3898 /Print/
ISSN 2538-7138 /Online/

INTEGRATION OF INQUIRY-BASED INSTRUCTION WITH FORMATIVE ASSESSMENT: THE CASE OF EXPERIENCED CHEMISTRY TEACHERS

Paweł Bernard,
Karol Dudek-Różycki,
Kinga Orwat

Abstract. *Inquiry-based teaching and learning methods are more and more widely used in the teaching natural sciences all over the world. The elements of inquiry are included in many national sciences' core curricula and indicate that students should be actively involved in acquiring knowledge. This fact changes the teacher's role in the classroom and the structure of everyday lessons. Independent inquiry helps students to develop various elements of scientific education, gain specific skills and competencies. Consequently, the assessment methods used so far are no longer sufficient and they should be expanded with new strategies, tools, and criteria. In recent years, there were many projects, in which teachers were trained in using both inquiry-based instructions and formative assessment. So far, unfortunately, there has been limited information on how teachers bring theory into practice and combine the new teaching and assessment methods, and how they adopt available teaching materials during everyday classes. This case study explored how experienced chemistry teachers, who have broad content, pedagogical content, and inquiry methodology knowledge, integrate inquiry-based teaching with formative assessment for the first time. The results of the study revealed the main limiting factors and teachers' dilemmas.*
Keywords: *formative assessment, inquiry-based learning, in-service science teachers, qualitative research, teacher professional development.*

Paweł Bernard, Karol Dudek-Różycki, Kinga Orwat
Jagiellonian University, Poland

Introduction

Inquiry-Based Science Education (IBSE) can be defined as *'…an intentional process of problem diagnosing, carrying out a critical analysis of experiments and searching for alternative solutions, planning research, testing hypotheses, searching for information, constructing models, discussions with colleagues and formulating coherent arguments'* (Linn, Davis, & Bell, 2004); or *as a driving strategy for learning activities, which lead to the pupil´s own active construction of the studied piece of knowledge* (Škoda, Doulík, Bílek, & Šimonová, 2015). It has been strongly advocated in the United States by the National Research Council (1996), in Europe by the 7th EU Framework Programme (Rocard, Csermely, Jorde, Walberg-Henriksson, & Hemmo, 2007), and in other regions. In contrast to the traditional teaching methods, the teacher plays here the role of a tutor and lesson moderator (Anderson, 2002; Windschitl, 2004). The students play the leading role and their activity is the most important. In 2008-2015, many teachers had an opportunity to participate in 7FP pan-European projects focused on teacher training and the introduction of inquiry-based methods into school practice. Moreover, in many projects, various educational materials have been produced to support teachers. So far, unfortunately, there has been limited information on how teachers bring theory into practice, how they use and adopt the teaching materials obtained, and what are the main limitations and problems while introducing IBSE and its various aspects.

Skills Developed Through IBSE

The implementation of inquiry-based education in practice can be described by various models (Atkin & Karplus, 1962; Martin, Sexton, & Gerlovich, 1999). For example, in the 5Es learning circle, the inquiry-based activities are divided into five phases: *Engage; Explore; Explain; Elaborate or Extend; Evaluate* (Bybee, 2002; 2009). According to the scheme, first of all, the interest in the problem should be developed, followed by a reference to the

https://doi.org/10.33225/jbse/19.18.184

already possessed knowledge, and then, there is time for carrying out the experiments and research, after which the knowledge should be generalized and applied in different situations. The final stage consists of the student's self-assessment related to the acquired information. Alternatively, the inquiry can be divided into stages of implementation of a research process. In Figure 1, the six-step cycle of inquiry-based process is presented (National Research Council, 1996).

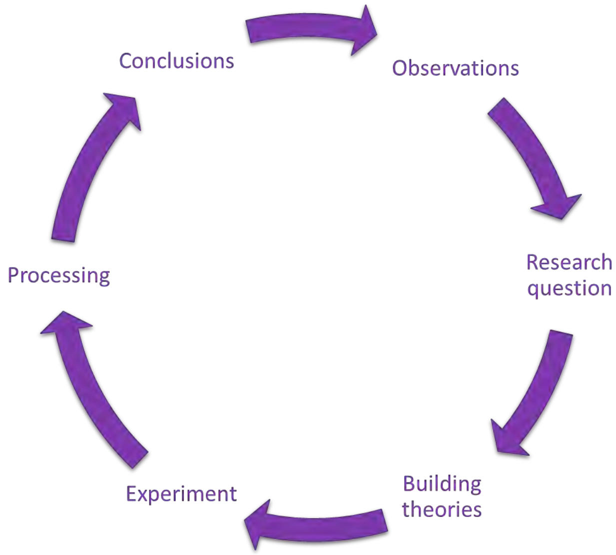

Figure 1. A 6-stage cycle for inquiry investigations and modelling.

This cycle describes the inquiry process well but does not precise students' activities, actions, and skills used and developed during the inquiry. The division based on the groups of skills developed by students through IBSE is shown in the form of a cycle in Figure 2.

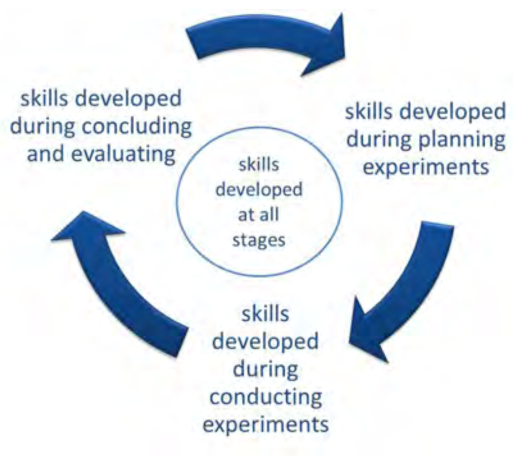

Figure 2. A simplified version of the experimental cycle based on the groups of skills being developed.

There are two widely recognized taxonomies of inquiry skills: one described by Fradd, Lee, Sutman and Saxton (2001), and the other by Wenning (2007). Table 1 shows the classification of the IBSE skills, presenting merged approaches mentioned above by Orwat, Bernard and Dudek (2014; 2016).

https://doi.org/10.33225/jbse/19.18.184

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

Table 1. List of skills developed through IBSE.

Stage of the inquiry	Developed skills
1. Planning	• Formulation of a research question • Formulation of a research hypothesis (with justification) • Defining variables • Development of a method to control variables (description of the mathematical dependencies) • Development of a method to collect raw data (description of the experimental procedure)
2. Conducting an experiment	• Collection of raw data • Processing data and determination of measurement error • Calculation/estimation of experimental errors • Presentation of results
3. Conclusions and evaluation	• Drawing conclusions • Evaluation of the plan and application of the experiment • Proposals of modifications, and improvement of the experiment
4. Skills developed at all stages	• Searching for information, critical analysis of information • Teamwork • Working safely at the laboratory • Using ICT

Assessment of Students Learning Through Inquiry

The listed skills can/should be acquired by inquiry-oriented learners, and the level of its mastering can be assessed. Of course, for the assessment, a proper selection of evaluation criteria is the key element, but at the same time, it is very problematic. In the literature, many examples of such criteria can be found (e.g. Ebenezer, Kaya, & Ebenezer (2011)), but creating a uniform, ready-to-use criteria is almost impossible. These have to be adjusted to students' age, prior experience, subject knowledge and the type of inquiry used. Moreover, the teacher has to consider possible methods of evidence collection for assessment the level of skills mastering, the number of students in the class, and even the equipment and facilities available in the school classroom/laboratory. Therefore, the assessment tools must be carefully planned and dedicated to particular inquiry-based classes. The intended role of the assessment is another problem. Wynne Harlen (2007) points out that a summative assessment has an undeniably strong impact on teaching methods and content, but she claims also that a formative assessment plays a special role in the evaluation of the inquiry skills (Harlen, 2000). Formative assessment, also called 'assessment for learning', is '*the process used by teachers and students to recognize and respond to student learning in order to enhance that learning, during the learning*' (Bell & Cowie, 2001). Thanks to it, the students receive feedback on the skills they have already fully mastered and those, which should be developed further (Black & Harrison, 2004). It enables an ongoing evaluation of almost all skills trained by inquiry. Such an evaluation enables the planning of changes in the teaching and learning process in order to improve students' achievements (Harrison, 2014). The formative assessment can be used at any stage of learning, contrary to the summative assessment, which is usually applied at the end of this process (Olson & Loucks-Horsley, 2000) and provides information about students' progress in relation to the selected population (Harlen, 2000).

According to Llewellyn (2007), the tools sufficient for the evaluation of the students' inquiry skills are: multiple-choice and constructive response questions; rubric; monitoring charts; concept maps; journal entries, lab reports; oral presentations; performance and self-assessment. While the multiple-choice questions are usually applied for the assessment of the content knowledge, rubrics and oral presentations can help in the evaluation of the students' skills to perform the inquiry process (Llewellyn, 2007). It should be also emphasized that rubrics are considered particularly useful for the evaluation of the inquiry skills. Barron and Darling-Hammond (2008) consider rubrics (analytical and holistic ones) to be one of the most effective tools for students' evaluation. Learning through IBSE enables development of basic, but also higher order reasoning skills (Csapó et al., 2012;

https://doi.org/10.33225/jbse/19.18.184

Lawson, 2005). The assessment of students' scientific reasoning competences is a complex task, and requires more advanced tools, like Lawson's tests (Lawson, 1978; 2000a; 2000b) or another reasoning test; e.g. those proposed by Gobert, Sao Pedro, Raziuddin and Baker (2013), Surim and Humburg (2011) or Donovan, Hutton, Lennon, O'Connor and Morrissey (2008). The above-mentioned tools can be used both in the summative and formative assessment, as the character of the evaluation should mostly depend on its objective, and not on the tools applied (Black & Harrison, 2004). Moreover, Zhang and Misiak (2015) claim that the most effective is a combination of standard-based grading with written feedback.

There are many known limiting factors affecting the implementation of inquiry-based instruction in school practice. The most common are teachers' beliefs and attitudes toward the method (Arce, Bodner, & Hutchinson, 2014; Luft, 2001; Windschitl, 2004), understanding of inquiry and the nature of science (Roehrig & Luft, 2004; Morrison, 2014; Zhang, Krajcik, Sutherland, Wang, Wu, & Qian, 2005), and lack of content or pedagogical content knowledge (Luft, 2001; Jeanpierre, Oberhauser, & Freeman, 2005; Taitelbaum, Mamlok-Naaman, Carmeli, & Hofstein, 2008).

The aim of the research was to examine how teachers having a positive attitude toward IBSE, knowing inquiry-based methodology, having a broad content and pedagogical content knowledge, are able to integrate inquiry-based instruction with the formative assessment; what difficulties they face, what problems they are able to solve on their own, and what motivates them to further actions?

Research Methodology

Background

Answering the research questions required very specific participants – teachers who were well educated, experienced in school teaching and had a positive attitude toward inquiry. It was agreed that the participants should have an MSc and PhD degree in the taught subject, and at least 10 years of professional experience. Also, they should have broad theoretical knowledge on the inquiry methodology (preferably confirmed by a training certificate), some experience in teaching using this method and a positive attitude toward it. Participants were recruited among participants of teacher training realized within the SAILS (Strategies for Assessment of Inquiry Learning in Science) FP7 project.

Participants

The chosen criteria were met by two of 75 teachers participating in the project in 2014. Both were chemistry teachers who completed earlier training on teaching with the use of the IBSE during the ESTABLISH FP7 project (Bernard, Maciejowska, Krzeczkowska, & Odrowąż, 2015). Moreover, both teachers had an academic background in in-service and pre-service teachers' tutoring. Even though the participants of the study met all recruitment criteria, there were still some differences between them. They were teaching at different educational levels: teacher 1 (pseudonym Adriana) at the lower secondary school level, K7-9, age 15-17; and teacher 2 (pseudonym Barbara) at the upper secondary school level, K10-12, age 17-19. The study was carried out after a basic (theoretical) training in the assessment of students learning through IBSE. The training was delivered in the form of a 2-day-long workshop, including 12 hours of classes (Jönsson, Lundström, Finlayson, McLoughlin, & McCabe, 2014). The research was conducted during the teachers' first attempt to introduce IBSE with the integrated assessment in practice.

Participation in the research was voluntary. The teachers and the students were aware of the data to be collected, the goal of the collection, and the mode of processing, according to the Jagiellonian University ethical standards. The participants might renounce their participation in the study at any stage.

Instrument and Procedures

The teachers' task was to run inquiry-based classes and evaluate selected inquiry skills. The teachers could use the teaching materials gathered during the training, covering a set of chemical inquiry-based activities for various levels of education (final version of materials: Finlayson, McLaughlin, Coyle, McCabe, Lovatt, & Van Kampen, 2015a; 2015b). In the materials, they could find a theoretical background, a proposed strategy of educational approach, and suggested IBSE skills that could be evaluated during teaching. The teachers were also provided with the proposals of methods for collecting information for assessment and the types of tools that could be

https://doi.org/10.33225/jbse/19.18.184

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

applied in the evaluation process but without the provision of exact tools and evaluation criteria. The teachers were also informed that the assessment should have a formative character. The teachers could choose freely the class they are going to work with, develop their own educational strategy, select the skills for evaluation, select the data collecting tools and develop the criteria. The summary of teachers' and groups' characteristics is shown in Table 2.

Table 2. Comparison of teachers and students participating in the study

	Adriana	Barbara
Teachers' background	MSc in chemistry (specialization: chemistry education) PhD in chemistry (specialization: chemistry education) 20 years of experience in teaching chemistry at the lower secondary school level	MSc in chemistry (specialization: chemistry education) PhD in chemistry (specialization: inorganic chemistry) 10 years of experience in teaching chemistry at the upper secondary school level
Student group	Students' age: 16 years Type of school: 3rd grade of public lower secondary school. Number of students: 11 Gender: 7 female students, 4 male students Students experienced IBSE for the first time.	Students' age: 17-18 years Type of school: Public upper secondary school, a class with an extended programme of biology and chemistry. Number of students: 5 Gender: 4 female students, 1 male student Students experienced IBSE for the first time.
Chosen activity	"Properties of Plastic" (Ganajová, 2015)	"Which is the Best Fuel?" (Cakmakci, 2015)
Inquiry skills selected for evaluation	• Discussion with their peers • Planning investigations • Data collection	• Developing hypotheses • Defining variables • Taking notes (collecting raw data)

The research had a qualitative character, and a mixed-methodology approach during the data collection was used. The teachers compiled a description of the strategy used with a description of the designed assessment tools, and the data were collected during classes in the form of a short report, the classes were observed by the researchers, and finally, the teachers participated in a semi-structured interview.

Data Analysis

The results of the class observations were compared with teachers' reports. That allowed for the determination of contradictory elements, and identification of gaps in teacher's reports. Identified issues were discussed and clarified while follow-up interview. Finally, two extensive case reports were prepared. The reports underwent comparative content analysis in the following areas: design of the classes, assessment tools and strategy, evaluation of the educational process.

Results of Research

Design of the Classes

Case I – Adriana

The classes began with a presentation of the investigation objective and its general structure, aimed at the students. Then, the teacher presented the engaging questions suggested in the 'unit':

- Is plastic useful?
- What properties make the widespread use of plastic possible?
- Do all kinds of plastic have the same properties?

https://doi.org/10.33225/jbse/19.18.184

- Do plastics change with time?
- Which plastic material properties would you like to explore in more detail?
- Do the materials made of plastic have negative properties?

The questions were displayed using an overhead projector. Before proceeding to a discussion of the above-mentioned issues, the students were supposed to write down their answers on sheets of paper. When they were ready, there was a discussion on those. All the students were sitting at one large table so that they could see each other. The teacher was moderating the discussion but did not interfere with what the students were saying. Next, the signed sheets with students' answers were collected as the basis for the students' evaluation.

When the discussion was complete, the students were divided into five groups: four pairs and one group consisting of three students. The students formed the groups by themselves – three groups were single-gender and the other two – mixed. In the groups, the students planned and performed laboratory experiments and completed their worksheets. Each group had access to common laboratory reagents and equipment, and the students were free to choose the reagents and equipment to perform the experiments determining: 1) density; 2) flammability; 3) thermal stability; 4) thermal conductivity, and 5) electrical conductivity of the chosen plastics.

Case II – Barbara

It was the teacher's first attempt to use this kind of investigation. The classes lasted 135 minutes – 3 lesson periods, even though 4 school periods were suggested in the delivered materials. Before the classes, the students were informed about the topic to be discussed. The teacher sent them self-designed worksheets, which they were asked to complete individually prior to the lesson. The teacher pointed out that the aim of the worksheet was to introduce the topic, as it was not possible to achieve it during the three-lesson-period long classes. The teacher began the class with a discussion about concepts of fuel, various aspects of its' practical relevance, the heat capacity and enthalpy of combustion. After the discussion, the students formed two groups; group A was composed of one boy and two girls, and group B consisted of two girls. The students formed the groups by themselves. In the groups, they were asked to perform the laboratory tasks. Each group had all reagents and equipment needed for the experiments prepared at their workstations. The students were free to choose the reagents and equipment for the experiment they planned to carry out.

Assessment

Case I – Adriana

The teacher selected three skills for the formative assessment, and developed a strategy of evidence collection for each of the skills:

- *discussion with peers* – the teacher decided to evaluate this skill, although it was not detailed in the unit provided. The teacher used two sources of information for the assessment of this skill:
    1. observation card – completed by the teacher during the discussion,
    2. student artefacts – collection and evaluation of the sheets of paper on which the students wrote down their ideas for discussion.
- *planning investigations* – the assessment of the skill was based on Activity 1: 'Determination of density of plastics. The materials used for evaluation were the groups' worksheets and the student self-assessment questionnaires.
- *data collection* – the assessment of this skill was performed during Activity 3: 'Thermal stability of plastics. The teacher evaluated the worksheets filled in by the groups and the student self-assessment questionnaires. The methods for the assessment of this skill were based on the tools provided in the unit.

In the unit, 3-level analytical rubrics were presented for the assessment of the skills of planning investigations and data collection. The teachers decided to introduce the fourth level to the rubrics and for the skill of the discussion with peers, which was not suggested for the assessment in this unit, the teacher developed her own rubrics. The final rubrics used for the assessment are presented in Table 3.

https://doi.org/10.33225/jbse/19.18.184

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

Table 3. 4-level analytical rubrics developed and used by Adriana.

Skill	1 point	2 points	3 points	4 points
Discussion with peers	The student rarely takes part in the discussion. The student does not listen to the other members of the group. The student is not interested in the discussion (e.g. s/he does something else).	The student takes part in the discussion but only at the request of the person moderating the discussion. The student's statements are not always factually correct. The student listens to other students' statements.	The student occasionally takes part in the discussion. The student's suggestions are correct. The student respects the opinions of other people, but s/he is not always able to notice incorrect (illogical) statements.	The student often takes part in the discussion without the teacher's encouragement. The student provides suggestions that may be used by the group The student provides correct substantive justifications. The student can notice erroneous statements made by other discussion participants and s/he is able to correct them.
Planning investigations	The student understands the task, but s/he does not know what density is. S/he cannot plan the experiment independently.	The student knows what density is; s/he can suggest how to assess the density of plastics in relation to water, but s/he cannot justify why a given scenario has been applied in the experiment.	The student is able to define density; s/he can suggest a procedure for determining the density of plastics in comparison to water; s/he is able to substantiate the course of the experiment.	The student is able to define density; s/he can suggest a procedure for determining the density of plastics in comparison to water; s/he is able to substantiate the course of the experiment. S/he is able to list the limitations of the method.
Data collection	Missing or incorrect data in the table.	The data for four substances completed correctly in the table.	The data for five substances completed correctly in the table. An attempt to describe the structure of substance after taking it out of the water.	All data in the table completed correctly, described extensively, and all data that can be observed is presented.

Case II – Barbara

Barbara decided to base the assessment on students' notes taken during the classes. The students received a clear instruction to write down their research question, hypothesis, collected raw data, to calculate the results and present the conclusions. From the whole range of skills, three were evaluated: developing hypotheses, defining variables and collecting experimental data. The students' notes were collected by the teachers after the laboratory classes. The teacher decided not to use the worksheet suggested in the unit, allowing the students to arrange notes on their own. The criteria of the assessment were arranged in a 3-level analytical rubric designed by the teacher and presented in Table 4. The rubrics were introduced to the students before the classes.

Table 4. Rubrics developed by Barbara.

Skill	Fail	Satisfactory	Very good
Developing hypotheses	The student does not formulate a hypothesis appropriate to the research problem raised.	With the teacher's assistance, the student formulates a hypothesis for the research problem raised.	The student independently formulates a correct hypothesis, adequate for further experiments and referring to a correctly raised research problem.
Defining variables	The student does not define variables associated with the planned experiment.	The student defines some variables, and with the teacher's assistance s/he is able to identify other relevant variables.	The student independently defines appropriate dependent and independent variables.

https://doi.org/10.33225/jbse/19.18.184

Collection of experimental data	The student prepares incomplete, unreadable notes containing information unusable in terms of finding an answer to a research question raised.	With the teacher's help, the student is able to write down some information – but not enough to present and interpret the results, e.g. obtained data without symbols and units.	The student independently prepares appropriate notes, taking into account relevant units, quantities, and symbols, relationships between quantities reflected in formulas, presenting a logical cause and effect sequence that contains all the necessary information, which, in the end, allows for the calculation of the combustion efficiency of the tested fuels.

Teachers' Reflections

The teachers were asked to prepare reports describing the course of the classes. The report template had a semi-open structure, presented only the aspects that should be described by teachers. In the case of unclear or incomplete statements, the reports were supplemented basing on the interviews with the teachers. Below, fragments of reports describing the teachers' reflections on the strategies and assessment tools are presented.

Case I – Adriana

…The discussion, in my opinion, should guarantee the students an opportunity to present their thoughts/ideas, even if their statements are not always correct. However, when they know that this element is subjected to assessment, they may be afraid of expressing their own opinions. My students behaved differently than normally, ever since they knew that they were being assessed. Usually, they participate in the discussion more willingly, but I did want to ask questions and force them to give answers…

…The ability to take part in the discussion turned out to be very difficult for me to assess. While listening to what the students were saying, I found it difficult to take notes simultaneously. It would have been easier to assess this ability if the students had conducted the discussion in small groups, but in such case, I would be able to assess only one, selected group – I wouldn't know what the others are doing at that time…

…In the self-assessment sheets, only one female student indicated that she rarely took part in the discussion. The rest claimed that they spoke often or very often. Clearly, there was a great discrepancy between what the students said and my own observations, most probably due to the fact that they completed the self-assessment sheets after the class, i.e. after the discussion in small groups. I was not able to assess the discussion in pairs, since, for safety reasons, I had to walk around the classroom, and I could not observe the work of one pair only. The workplace of each group should be arranged so that the groups do not interfere with each other...

…I think that, in general, the worksheets are helpful when assessing the data collection skill. The tool for the assessment of this skill and the part of the experiment when the assessment should be done was suggested in the unit. However, the degree to which the teacher should help the students was a reason for doubts. I tried not to interfere with the students' work. As a result, the students did not know that they were required to provide a broader description of substances in the worksheet, so they put only plus or minus signs (+/-), or they wrote YES/NO. The skill should be assessed on the basis of the experiment, which provides a lot of data, or the worksheet needs to be very precise. However, in my opinion, a very detailed worksheet may reduce the creativity of students…

Case II – Barbara

…Initially, the students' activity was very cautious and characterized by uncertainty and shyness. Over time, their level of activity increased, first in the field of discussion, and then in laboratory work…

…In my opinion, it makes no sense to prepare rubrics including three levels only (as I did) – I should have prepared a complete scale of all school marks at once. Rubrics preparation is very difficult. I didn't know whether the descriptions should be quite detailed, or more general. The students were given the rubrics before the beginning of the class, i.e. they were

https://doi.org/10.33225/jbse/19.18.184

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

informed of what would be assessed. I do not know if the lack of variables resulted from the pressure of time or their poor skill of defining the variables…

…The topic of the unit was very interesting for both girls and boys. It presents a practical application of some chemical concepts, such as heat capacity, or enthalpy of combustion. The possibility to work independently in a laboratory revealed the involvement of the students that was previously unknown to the teacher, especially in the case of students who had low achievements in chemistry. The investigation can be run as open inquiry, but it requires a large workload and more time than suggested 4 lessons…

…I definitely wanted to see what the students understand by the concept of notes, and how they will cope with noting down certain facts, and what facts they will consider as worth noting…

Discussion

Both teachers successfully prepared and implemented the inquiry-based classes, together with the integrated assessment. However, in both cases, one can observe that the complexity of the challenge exceeded the teachers' expectations. For example, *Adriana* prepared appropriate tools for the assessment of the skill of 'discussion with peers' and adopted adequate assessment strategies and yet she had significant difficulties while implementing her plan. Filling in the observation cards and moderating the discussion in a large group turned out to be impossible. It was found also that the self-assessment tool was applied too late, and consequently, the students did not know which part of the classes it is referred to. As a result, the teacher's observations differed significantly from the self-assessment results. The teacher points out that this skill should be assessed on the basis of discussions in small groups, but she does not see the possibility to complete such a task. She does not take into account the possibility of assessing only a selected group of students, as, at all costs, she wants to evaluate all the students at the same time. The teacher informed the students which skills would be assessed, but she did not present them the prepared rubrics with the evaluation criteria. Black and William (1998) claim that the awareness of the guidelines/assessment criteria increases the student's learning efficiency. This fact was communicated to the teachers during the training.

Defining variables is an important step in the inquiry process (Chin & Osborne, 2008; Rodríguez-Arteche & Martínez-Aznar, 2016; Windschitl, 2004). Unfortunately, the skill is almost entirely overlooked in the Polish core curriculum (Bernard, Maciejowska, Odrowąż, Dudek, & Geoghegan, 2012). *Barbara* introduced the concept of variables to the students two years earlier. Before the laboratory part, the teacher discussed the elements of a proper experimental plan with the students. The teacher intentionally didn't discuss 'defining variables' as a separate issue during classes and did not address it in the worksheet, since she wanted to check what they remember from previous lessons when she drew the students' particular attention to the necessity of defining variables when planning an experiment. As a result, the students did not define any variables in their notes and it is unclear if they forgot to do it, had no time to do it, or they did not know how to do it. Such a situation is contrary to the formative assessment principles (Black, Harrison, Lee, Marshall, & Wiliam, 2002; Harlen, 2003), and it does not support students in the skills development nor create a friendly environment for learning.

The fear of the teachers of interfering in the students' activity is easily noticeable, both in the open and the guided parts of the activities. *Adriana* did not want to interfere in the students' discussions; she was only acting as a moderator and then took care of the safety during the experimental part. She made no immediate comments to the students on the fact that their notes were too brief. *Barbara* acted in a similar way. She took care of the safety of the laboratory work but did not draw the students' attention to the lack of defined variables. Moreover, teachers used only the engaging questions, but not the guiding questions suggested in the units. This indicates that both teachers taking part in the study have difficulties with the interpretation of the form and the role of formative assessment and scaffolding the inquiry process through questioning. Research in many classrooms showed that supporting teacher-student dialogue is rarely achieved (Black, Harrison, Lee, Marshall, & Wiliam, 2002; Blanchard, Southerland, & Granger, 2009; Harrison, 2016; Hughes, Green, & Greene, 2014; van Zee, 2000). The role of questions during inquiry enactment was underlined during the training. Still, the conflict between attributing the leading role to the students and the implementation of formative assessment is clearly noticeable.

Previous studies showed that teachers engaged in using formative assessment appreciate the value of self-designed rubrics (Britner & Finson, 2005). Similarly, in this research, both teachers felt that the rubrics are an appropriate tool supporting the assessment during the chosen activities and the used evaluation strategies. They

https://doi.org/10.33225/jbse/19.18.184

demonstrated high confidence when adapting and creating rubrics. Both teachers prepared a description of the assessment criteria in a more analytical than holistic way. The rubrics prepared show that the teachers understand the process of inquiry and the features of the skills developed through IBSE, but neither of them included the students in the process of creating rubrics. The developed rubrics describe the assessment criteria in detail, which may constitute an important element of the formative assessment. Both teachers worked with students who had no experience in learning through inquiry. In both cases, one may observe the problem of estimating the level of the students' skills and the preparation of rubrics in such a way as to reflect the actual skills of different groups of students in the class. In this respect, once again, the teachers' commitment to summative assessment is easily observable. Both teachers tried to directly translate the descriptive assessment into a specific mark written down in a class register. Noticed difficulties with linking formative and summative assessment are typical and appear not only at an early stage of formative assessment implementation (Black, Harrison, Hodgen, Marshall, & Serret, 2010; 2011; Harlen, 2013).

Conclusions

The results clearly demonstrate that the integration of the assessment process and inquiry is a difficult task. The teachers with extensive professional experience, broad content and the IBSE methodology knowledge, and experience in conducting research (both have a PhD) have been selected for the study intentionally. Such a selection was to ensure that the main problem they would have to face would be the integration of the assessment in the process of inquiry rather than other conceptual or methodological issues. The participants had not tried to integrate the assessment with the process of inquiry before; however, they received training in this field. The task was accomplished successfully; in both cases, the process of inquiry was carried out satisfactorily. Unfortunately, due to the significant divergence of the age of the students, it was not possible to implement the same unit in both cases, which makes the comparative analysis of the results more difficult. Nevertheless, some common trends for both cases are observed. The biggest challenge consisted in the implementation of the assessment in a formative way. It is probably due to the fact that such a type of assessment is not often used in Polish lower and upper secondary schools. The teachers are committed to summative assessment and marks put in the register. Descriptive assessment is mainly associated with the primary school. The feedback provided to students after summative tests of acquired knowledge is a commonly used component of the formative assessment.

As mentioned earlier, the teachers participating in the study completed the training in the field of the assessment of students learning through inquiry. Their strong conviction of the high value of the rubrics is clearly visible. It is also noticeable that they are aware of a wide range of criteria, which may be taken into account, even during the assessment of basic skills. The problem consists in the awareness of the actual level of students' skills and adaptation of the criteria in the rubrics to it. The tendency to create 'exaggerated' rubrics can be observed, where the highest skill level reflects the ideal. On one hand, this discrepancy between the actual and the ideal state of affairs shows the direction of development to the students, but on the other hand, higher levels of skills may seem unattainable, which might have a demotivating effect.

The diagnosed main difficulties, together with a number of other, minor ones, are largely determined by the long-standing education system and the lack of the teachers' experience in the use of formative assessment. However, the fact that the participation in the training on the integration of IBSE and assessment has resulted in the teachers' extensive theoretical knowledge and developed strong self-esteem in design assessment tools, allows us to look positively to the future. Those values, combined with the gained experience, are likely to transform into actual skills and refine the teaching process.

The main limitation of the presented research lies in the number of cases. The strict rules of participation in the study ensured the desired model situation but at the same time limited the number of possible participants. Moreover, both participating teachers specialized in teaching chemistry, what made comparison easier but didn't deliver information if the observed difficulties are common for teachers of all sciences' subjects. Finally, it should be remembered that the teachers participating in the research were extraordinary cases. They were very well educated, experienced and motivated, and still, they faced major problems while completing the task. 'Ordinary' teachers have to deal not only with a challenging combination of inquiry and assessment but also face problems with the application of the inquiry process and the introduction of content knowledge. A description of such cases would be an interesting step forward.

https://doi.org/10.33225/jbse/19.18.184

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

Acknowledgement

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

Anderson, R. (2002). Reforming science teaching: What research says about inquiry. *Journal of Science Teacher Education, 13*(1), 1-12.
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.
Atkin, J. M., & Karplus, R. (1962). Discovery or invention? *The Science Teacher, 29*(5), 45-51.
Barron, B., & Darling-Hammond, L. (2008). How can we teach for meaningful learning? In: L. Darling-Hammond, B. Barron, P. D. Pearson, A. H. Schoenfeld, E. K. Stage, T. D. Zimmerman, J. L. Tilson (Eds.) *Powerful learning: What we know about teaching for understanding.* San Francisco: Jossey-Bass.
Bell, B., & Cowie, B. (2001). The characteristics of formative assessment in science education. *Science Education, 85*(5), 536-553.
Bernard, P., Maciejowska, I., Krzeczkowska, M., & Odrowąż, E. (2015). Influence of in-service teacher training on their opinions about IBSE. *Procedia – Social and Behavioral Sciences, 177,* 88-99. 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. doi.org/10.2478/cdem-2013-0004
Black, P., & Harrison, C. (2004). *Science inside the black box: Assessment for learning in the classroom.* London: GL Assessment.
Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education, 5*(1), 7-71.
Black, P., Harrison, C., Hodgen, J., Marshall, B., & Serret, N. (2010). Validity in teachers' summative assessments. *Assessment in Education: Principles, Policy & Practice, 17*(2), 215-232. doi.org/10.1080/09695941003696016
Black, P., Harrison, C., Hodgen, J., Marshall, B., & Serret, N. (2011). Can teachers' summative assessments produce dependable results and also enhance classroom learning? *Assessment in Education: Principles, Policy & Practice 18*(4), 451-469. doi.org/10.1080/0969594X.2011.557020.
Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2002). *Working inside the black box: Assessment for learning in the classroom.* London: King's College.
Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning–putting it into practice.* Buckingam: Open University Press.
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. doi.org/10.1002/sce.20298.
Britner, S. L., & Finson, K. D. (2005). Preservice teachers' reflections on their growth in an inquiry-oriented science pedagogy course. *Journal of Elementary Science Education, 17*(1), 39-54.
Bybee, R. (2002). *Learning science and the science of learning.* Arglington: NSTA Press.
Bybee, R. (2009). A commissioned paper prepared for a workshop on exploring the intersection of science education and the development of 21st century skills. *The BSCS 5E instructional model and 21st century skills.* USA: National Science Teachers Association.
Cakmakci, G. (2015). Which is the best fuel? In: O. Finlayson, E. McLaughlin, E. Coyle, D. McCabe, J. Lovatt, & P. Van Kampen (Eds.). *SAILS Inquiry and Assessment Units, Vol. 2* (pp. 119-129). Dublin: DCU. Retrieved from http://sails-project.eu/sites/default/files/outcomes/SAILS_units_volume-2.pdf.
Chin, C., & Osborne, J. (2008) Students' questions: a potential resource for teaching and learning science. *Studies in Science Education, 44*(1), 1-39. doi.org/10.1080/03057260701828101.
Csapó, B., Csíkos, C., Korom E., Németh, M. B., Black, P., Harrison, C., van Kampen, P., & Finlayson, O. (2012). *Report on the strategy for the assessment of skills and competencies suitable for IBSE.* Strategies for the Assessment of Inquiry Learning in Science (SAILS FP7 project). Retrieved from http://sails-project.eu/sites/default/files/outcomes/d2-1.pdf.
Donovan, J., Hutton, P., Lennon, M., O'Connor, G., & Morrissey, N. (2008). *National Assessment Program – science literacy year 6 school release materials, 2006*. Carlton, South Victoria: Ministerial Council on Education, Employment, Training and Youth Affairs.
Ebenezer, J., Kaya, O., & Ebenezer, D. (2011). Engaging students in environmental research projects: Perceptions of fluency with innovative technologies and levels of scientific inquiry abilities. *Journal of Research in Science Teaching, 48*(1), 94-116.
Finlayson, O., McLaughlin, E., Coyle, E., McCabe, D., Lovatt, J., & Van Kampen, P. (Eds.). (2015a). *SAILS Inquiry and Assessment Units, Vol. 1*. Dublin: DCU. Retrieved from 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, Vol. 2*. Dublin: DCU. Retrieved from http://sails-project.eu/sites/default/files/outcomes/SAILS_units_volume-2.pdf.
Fradd, S., Lee, O., Sutman, F., & Saxton , M. (2001). Promoting science literacy with English language learners through instructional

https://doi.org/10.33225/jbse/19.18.184

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

materials development: A case study. *Bilingual Research Journal, 25*(4), 417-439.

Ganajová, M. (2015). Polymers. In: O. Finlayson, E. McLaughlin, E. Coyle, D. McCabe, J. Lovatt, & P. Van Kampen (Eds.). *SAILS Inquiry and Assessment Units, Vol. 1* (pp. 77-92). Dublin: DCU. Retrieved from http://sails-project.eu/sites/default/files/outcomes/SAILS_units_volume-1.pdf.

Gobert, J.D., Sao Pedro, M., Raziuddin, J., & Baker, R.S. (2013). From log files to assessment metrics: Measuring students' science inquiry skills using educational data mining. *Journal of the Learning Sciences, 22*(4), 521-563.

Harlen, W. (2000). Assessment in the inquiry classroom. In: *Foundations: A monograph for professionals in science, mathematics, and technology education, Vol. 2. Inquiry. Thoughts, Views, and Strategies for the K-5 Classroom* (pp. 87-98). USA: National Science Foundation. Retrieved from https://www.nsf.gov/pubs/2000/nsf99148/htmstart.htm

Harlen, W. (2003). *Enhancing inquiry through formative assessment.* San Francisco: Exploratorium.

Harlen, W. (2007). *Assessment of learning.* London: SAGE Publishing.

Harlen, W. (2013). *Assessment & inquiry-based science education: issues in policy and practice.* Trieste: Global Network of Science Academies.

Harrison, C. (2013). Collaborative action research as a tool for generating formative feedback on teachers' classroom assessment practice: the KREST project. *Teachers and Teaching, 19*(2), 202-213. doi.org/10.1080/13540602.2013.741839.

Harrison, C. (2014). Assessment in the inquiry classroom. *Science Education International, 25*(1), 112-122.

Hughes, S., Green, C., & Greene, V. (2014). *Report on current state of the art in formative and summative assessment in IBE in STM - Part II.* (ASSIST-ME Report Series, Number 2). Denmark: University of Copenhagen.

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. doi.org/10.1002/tea.20069.

Jönsson, A., Lundström, M., Finlayson, O., McLaughlin, E., & McCabe, D. (2014). *Report on IBSE teacher education and assessment programme – stage 1*. Strategies for the assessment of inquiry learning in science (SAILS FP7 project). Retrieved from http://sails-project.eu/sites/default/files/outcomes/d4-2.pdf.

Lawson, E. A. (1978). The development and validation of a classroom test of formal reasoning. *Journal of Research in Science Teaching, 15*(1), 11-24.

Lawson, E. A. (2000a). *Classroom test of scientific reasoning*. Test revision, Arizona State University, Arizona. Retrieved from http://www.public.asu.edu/~anton1/LawsonAssessments.htm.

Lawson, E. A. (2000b). The generality of hypothetico – deductive reasoning: making scientific thinking explicit. The American Biology Teacher, *62*(7), 482-495.

Lawson, E. A. (2005). What is the role of induction and deduction in reasoning and scientific inquiry? *Journal of Research in Science Teaching, 42*(6), 716-740.

Linn, M. C., Davis, E. A., & Bell, P. (2004). *Internet environments for science education.* London: Lawrence Erlbaum Associates.

Llewellyn, D. J. (2007). *Inquire within: implementing inquiry-based science standards in grades 3-8*. Thousand Oaks, CA: 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.

Martin, R., Sexton, C., & Gerlovich, J. (1999). *Science for all children lessons for constructing understanding.* Needham Heights: Allan & Bacon.

Morrison, J. A. (2014). Scientists' participation in teacher professional development: the impact on fourth to eighth grade teachers' understanding and implementation of inquiry science. *International Journal of Science and Mathematics Education*, *12*(4), 793-816. doi.org/10.1007/s10763-013-9439-3.

National Research Council. (1996). *National science education standards*. Washington, D.C: National Academy Press.

Olson, S., & Loucks-Horsley, S. (2000). *Inquiry and the national science education standards: A guide for teaching and learning.* Washington, D.C.: National Academy Press.

Orwat, K., Bernard, P., & Dudek, K. (2016). Starego drzewa nie zegniesz, czyli jak szkolić nauczycieli z zakresu stosowania metod samodzielnego dociekania wiedzy przez uczniów (IBSE) [Old trees do not bend, how to train teachers in the application of inquiry-based methods (IBSE)]. *Aktualne problemy dydaktyki przedmiotów przyrodniczych* (pp. 165-179). Kraków: Wydział Chemii UJ.

Orwat, K., Bernard, P., & Dudek, K. (2014). Inquiry based science education – bringing theory to practice. In: *Science and Technology Education for 21st Century. 9th IOSTE Symposium for Central and Eastern Europe* (pp. 225-238). Hradec Kralove: Gaudeamus.

Rocard, M., Csermely, P., Jorde, D., Lenzen, D., Walberg-Henriksson, H., & Hemmo, V. (2007). *Science education now: A renewed pedagogy for the future of europe.* Brussels: European Communities.

Rodríguez-Arteche, I., & Martínez-Aznar, M. M. (2016). Introducing Inquiry-Based Methodologies during Initial Secondary Education Teacher Training Using an Open-Ended Problem about Chemical Change. *Journal of Chemical Education, 93*(9), 1528-1535. doi.org/10.1021/acs.jchemed.5b01037.

Roehrig, G. H., & Luft, J. A. (2004). Constraints experienced by beginning secondary science teachers in implementing scientific inquiry lessons. *International Journal of Science Education, 26*(1), 3-24.

Sirum, K., & Humburg, J. (2011). The experimental design abillity test (EDAT). *Bioscene*, 37, 8-16.

Škoda, J., Doulík, P., Bílek, M., & Šimonová, I. (2015). The effectiveness of inquiry-based science education in relation to the learners´ motivation type. *Journal of Baltic Science Education*, *14*(6), 791-803.

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. doi.org/10.1080/09500690701854840.

https://doi.org/10.33225/jbse/19.18.184

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

van Zee, E. H. (2000). Analysis of a student-generated inquiry discussion, *International Journal of Science Education, 22*(2), 115-142. doi.org/10.1080/095006900289912.
Wenning, C. (2007). Assessing inquiry skills as a component of scientific literacy. *Physics Teacher Education Coordinator, 4*(2), 21-24.
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. doi.org/10.1002/tea.20010
Zhang, B., & Misiak, J. (2015). Evaluating Three Grading Methods in Middle School Science Classrooms. *Journal of Baltic Science Education*, *14*(2), 207-215.
Zhang, B., Krajcik, J. S., Sutherland, L. M., Wang, L., Wu, J., & Qian, Y. (2005). Opportunities and challenges of China's inquiry-based education reform in middle and high schools: Perspectives of science teachers and teacher educators. *International Journal of Science and Mathematics Education*, *1*(4), 477-503.

Received: *November 21, 2018*   Accepted: *March 23, 2019*

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: 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 Website: http://www.zdch.uj.edu.pl/karol-dudek ORCID: https://orcid.org/0000-0003-3518-2089
Kinga Orwat	PhD, Assistant, Jagiellonian University, Department of Chemical Education, 2 Gronostajowa Str. 30-387 Kraków, Poland. E-mail: kinga.orwat@uj.edu.pl

https://doi.org/10.33225/jbse/19.18.184