



ARTICLE



<https://doi.org/10.1057/s41599-025-04437-3>

OPEN

The impact of long-term inquiry-based science education on students' motivation and knowledge acquisition: the role of gender, subject, and level of inquiry

Milada Teplá¹ & Petr Distler¹✉

The long-term effects of students' perceptions of inquiry-based science education when the introductory phase (comprising confirmation and structured inquiry) is omitted and guided laboratory tasks are directly implemented have not yet been thoroughly investigated. This study aims to assess the long-term effects of inquiry-based science education (IBSE) on student motivation and knowledge acquisition, comparing it with traditional laboratory activities across selected STEM subjects. Initial findings reveal that introducing inquiry-based activities in Chemistry classes reduces student interest, increases pressure on female students, and reduces the overall effort students put into learning Chemistry. Conversely, in Biology, these activities appear to reduce pressure for all students. The data indicates that while student motivation is initially low, it improves with continued exposure to IBSE methods. Furthermore, the research demonstrates that IBSE enhances students' skills and knowledge acquisition over time compared to traditional laboratory instruction. The effectiveness of IBSE varies significantly depending on the STEM subject and is influenced by student gender. Overall, IBSE shows a positive impact on long-term knowledge acquisition and motivation. These findings support previous recommendations, particularly for Chemistry, suggesting that IBSE should be introduced gradually, beginning at lower levels of inquiry.

¹Department of Chemistry Education, Faculty of Science, Charles University, Albertov 6, 128 00 Praha 2, Prague, Czech Republic. ✉email: petr.distler@natur.cuni.cz

Introduction

Curricula across the Czech Republic and other European countries have undergone reforms, driven in part by the Treaty on the Functioning of the European Union and recommendations from the Key Competences for Lifelong Learning framework (European Commission, 2019). These reforms aim to increase the participation of young people in science, technology, engineering, and mathematics (STEM) fields, recognizing that students with above-average interest in STEM are more likely to pursue further studies in these areas (Klahr and Nigam, 2004). This could be achieved by implementing Inquiry-based science education (IBSE) or the integration of inquiry-based approaches in STEM teaching as progressive learning methods (Palmer, 2009; Tuan et al. 2005). STEM and IBSE are intrinsically linked, with IBSE serving as an approach to teaching STEM subjects. IBSE is a student-centered methodology that emphasizes the development of scientific inquiry skills, including questioning, investigation, and experimentation. This approach focuses on hands-on and inquiry-based learning, encouraging students to ask questions, make observations, and formulate explanations based on evidence (Bell et al. 2010). The primary objectives of IBSE are to enhance students’ critical thinking and problem-solving abilities while simultaneously stimulating their interest in STEM subjects (Hanley et al. 2020). Therefore, IBSE can be viewed as a key element of STEM education, particularly within the area of science education.

Inquiry-based science education. Inquiry-based science education focuses primarily on developing students’ thinking and problem-solving abilities, as well as on their practical skills and dexterity (Hattie, 2008). Hattie (2008) also adds that inquiry-based teaching involves generating situations where students are required to observe and question phenomena, suggest explanations for the observations they have made, design and carry out experiments that provide evidence to support or contradict hypotheses, and analyze and draw conclusions from data.

Many authors (e.g. Banchi and Bell, 2008; Bell et al. 2005; Eastwell, 2009) have identified multiple levels of inquiry based on the amount of information provided to students and the extent of teacher guidance, e.g. through questions, comments, or instructions. In practice, the four-level research model is the most commonly used, which categorizes IBSE levels as follows: 1. confirmation inquiry, 2. structured inquiry, 3. guided inquiry, and 4. open inquiry (Banchi and Bell, 2008). It should be noted that the first two levels of inquiry cannot be considered typical inquiry-based science teaching, and are therefore referred to as “traditional laboratory activities” in the paper. The levels of IBSE are shown in Table 1. In traditional laboratory activities, students already know both the research question and procedure, and the conclusions are made by a teacher. This is similar to a structured inquiry, but in addition students also formulate the conclusions. They do not have to establish a hypothesis or search for a solution procedure. Nevertheless, it is one of the most frequently used forms of inquiry (Eastwell, 2009). Students who develop their skills through an IBSE approach to learning, go through the following stages of the research process: Engage, Explore, Explain,

Elaborate, and Evaluate. The phases of the Primary Connections 5E teaching and learning model are based on the 5Es instructional model (Koyunlu Ünlü and Dökme, 2022).

Confirmation inquiry: Students are provided not only with the instructions but also the research question and methodology used to verify previously known results. This level is, therefore, critical in gaining experience with practical and research activities, especially when collecting and recording data. *Structured inquiry:* In structured inquiry, the teacher sets a research question for the students to answer based on their collected data, which is arrived at by following a precisely given research procedure (instructions). *Guided inquiry:* At this level, students are close to conducting research activities independently. The teacher generates the topic and research question, but students are responsible for determining procedures and methods for data collection and interpretation. Students learn to independently plan an experiment from the hypothesis to the interpretation of results. *Open inquiry:* Once students have gained significant experience at the previous levels, they can conduct completely independent research driven by their own motivation. They formulate their own research questions, and answer them based on their own interpretation of the results, having followed their own procedures.

Using the IBSE approach, students should gradually change their role from passive listeners/learners to active researchers. The easiest start for students is level 1 (confirmation inquiry), where the teacher maintains responsibility. Students should gradually progress to higher levels of the IBSE approach, learning different procedures for solving teaching tasks and interpreting and generalizing any information obtained (Bell, 2004; Carpineti et al. 2015). Depending on the level of inquiry, students take more responsibility for their learning and the teacher moves into the role of supporter and observer (Banchi and Bell, 2008; Eastwell, 2009).

It is argued by some authors (Kirschner et al. 2006; Klahr and Nigam, 2004; Mayer, 2004) that students need very specific professional guidance during the entire learning process, and therefore, teaching with IBSE elements is considered less effective. They contend that novice learners should be provided with precise instructions and work procedures rather than being expected to discover these independently. Kirschner et al. (2006) highlight the possible development of misconceptions and the acquisition of knowledge without the full context when using inquiry-based approaches. They argue that if students with minimal knowledge and limited instruction begin searching independently for methodologies on a given topic, they could get lost in the inquiry process. This confusion could lead to students having a negative attitude, not only towards the given topic but towards the subject as a whole (Kalyuga et al. 2010).

The effect of an IBSE approach on motivation and study results. Motivation plays an important role in students’ learning processes and results, and in creating and maintaining a long-term interest in learning. It allows students to improve their skills by increasing the amount of effort they put into their studies. In

Table 1 Level of IBSE-a difference between traditional laboratory courses and IBSE.					
Laboratory course	Level of IBSE	Forming research question	Designing procedure	Testing the procedure	Explanations, conclusions
Traditional	Confirmation inquiry	Teacher	Teacher	Student	Teacher
	Structured inquiry	Teacher	Teacher	Student	Student
	Guided inquiry	Teacher	Student	Student	Student
	Open inquiry	Student	Student	Student	Student

addition to this, highly motivated students tend to perceive study materials as more significant and relevant. When students understand the importance and relevance of a subject in relation to their future goals, it can increase their motivation to study. However, if students feel pressured when learning, for example, due to high expectations or an excessive workload, it can lead to a decrease in motivation (Castle and Buckler, 2021). Motivation and related factors play a pivotal role in determining the effectiveness of learning (Kwarikunda et al. 2022). Evaluating motivation using appropriate, well-established tools from both psychology and sociology (Deci et al. 1999; Lei et al. 2024; Pintrich, Smith et al. 1991), can help increase the efficiency of learning processes.

The effect of an IBSE approach on motivation and attitude towards STEM subjects. Motivation and motivational attitudes toward learning STEM subjects are crucial in STEM education. Tuan et al. (2005) demonstrated that implementing IBSE positively influences junior high school students' motivational attitudes toward STEM subjects. Kane (2013) further supports this, noting that IBSE enhances students' motivation, fostering interest and maintaining positive attitudes toward natural sciences. The benefits of IBSE extend beyond subject-specific motivation. Madden (2011) and Patrick et al. (2009) found that an IBSE approach positively affects students' overall self-confidence in STEM education. Moreover, Palmer (2009) and Tuan et al. (2005) observed that IBSE positively impacts students' motivation for education in general. However, it is important to note potential drawbacks. Kalyuga et al. (2010) caution that students may develop negative attitudes not only towards a specific laboratory course but also towards the subject as a whole when tasked with solving problems about which they have minimal knowledge and instructions.

The effect of an IBSE approach on acquired knowledge. Researchers are still evaluating the effectiveness of an IBSE approach. Hattie (2008) published a monograph that synthesized four meta-analyses, including a total of 205 studies focusing on the effectiveness of an IBSE approach in relation to outcomes. It was found that an IBSE approach has the greatest positive impact on first-grade students in primary school. Furtak et al. (2012) conducted another meta-analysis of 37 papers, through which they found the positive impact of an IBSE approach in relation to acquired knowledge and skills. Minner et al. (2010) analyzed 138 papers, showing the positive influence of IBSE, which emphasizes student active thinking and concluding data.

Despite the positive findings in many studies, some research has not identified a significant effect of the IBSE approach on students' learning outcomes. Cobern et al. (2010) conducted a 2-week study on high-school students, measuring their acquired level of knowledge and skills during laboratory courses. The results of this experiment revealed no significant difference between the groups taught using traditional methods and those using the IBSE approach; similar conclusions were reached by Sadeh and Zion (2009). A negative correlation between the level of students' acquired knowledge and the IBSE approach was found in the PISA 2006 research (Gee and Wong, 2012; Lavonen and Laaksonen, 2009) and subsequently corroborated by McConney et al. (2014) and by Cairns and Areepattamannil (2017).

The effect of the level of an IBSE approach on acquired knowledge. Lederman et al. (2008) conducted a significant study comparing different levels of inquiry within laboratory activities. They assigned teachers a unified topic and instructed them to teach using either traditional methods, confirmatory inquiry, structured

inquiry, or a combination thereof. Their findings, along with those of other researchers, suggest a positive correlation between a higher level of inquiry, and increased effectiveness of teaching in terms of skill acquisition and understanding of scientific content (Blanchard et al. 2010; Chang and Mao, 2010; Sadeh and Zion, 2009). A study by Blanchard et al. (2010) compared the level of knowledge as a result of structured and guided inquiry across 1700 American high school students. Students working with guided inquiry teaching tasks achieved better results than those working with tasks with confirmation inquiry. Another combination is the impact of structured and guided inquiry, which Bunterm et al. (2014) observed: In this study, students engaged in practical laboratory tasks for 14–15 h analysis of the didactic test results showed that the group working with tasks at the level of guided inquiry demonstrated significantly better understanding of the content and consequently performed better in the practical tasks associated with scientific work than the group engaged in structured inquiry.

During the same period, Chatterjee et al. (2009) conducted research on American university students during a semester-long General Chemistry course. However, their results showed the opposite to those of Blanchard et al. (2010), with students preferring structured inquiry to guided inquiry. Cobern et al. (2010) conducted a 2-week test with American high school students. They monitored students' levels of acquired knowledge and skills in laboratory activities, comparing an experimental group (solving guided-inquiry tasks) with a control group (solving confirmation and structured guided tasks). Their evaluation showed no significant difference between the groups in terms of learning outcomes. However, students expressed a preference for lower levels of inquiry. These results can be interpreted similarly to the results of Chatterjee's (2009) study, as students in both groups worked for a limited time period (15 weeks/2 weeks) and did not work with tasks assigned sequentially, i.e. from confirmation inquiry to guided/open inquiry. Rather, they immediately started with challenging guided inquiry (Bell, 2004; Cairns and Areepattamannil, 2017; Carpineti et al. 2015).

Building on the outcomes of Banchi and Bell (2008), Bell (2004), Carpineti et al. (2015), and Eastwell (2009), students who have not previously encountered an IBSE approach may face difficulties in solving practical laboratory tasks assigned in the guided inquiry. Based on the results of various studies (e.g. Blanchard et al. 2010; Bunterm et al. 2014; Chang and Mao, 2010; Lederman et al. 2008; Sadeh and Zion, 2009) showing that the higher the IBSE level (higher level of inquiry) the better the outcomes students achieve, many educators and researchers try to assign guided inquiry-style tasks to students possessing little or no experience with an IBSE approach (e.g., Chatterjee et al. 2009; Cobern et al. 2010). Assigning practical activities in this way and the subsequent analysis of an IBSE approach's impact on student knowledge acquisition and attitudes, are contrary to the aforementioned widely accepted principles governing the implementation of IBSE methodologies in educational contexts. It is not yet clear what the long-term effects are regarding how students perceive an IBSE approach if the introductory phase (confirmation and structured inquiry) is skipped and guided-style laboratory tasks are assigned straight away. Similarly, the effects of how this frequently practiced approach affects their knowledge remains uninvestigated. Given the relatively limited experience with implementing IBSE approaches in the Czech educational context compared to international counterparts, there exists a notable paucity of longitudinal and methodologically consistent studies comparing a classical (tasks assigned at the level of confirmation and structured inquiry) and an IBSE approach (tasks at the guided inquiry level). The findings from

such comparative analyses could potentially inform and guide future educational policy directives in the Czech Republic.

This informed the research design, which aims to assess the impacts of an IBSE approach on both the motivational and knowledge levels of high school students in Chemistry and Biology classes. The study was conducted as a longitudinal experiment spanning an entire academic year.

The role of gender and subject in an IBSE approach. Many pedagogical experiments on an IBSE approach are performed for STEM subjects in general. However, it was shown that the results are very often subject-dependent (e.g. Vlckova et al. 2019; Potvin and Hasni, 2014) and, therefore, appropriate attention should be dedicated to certain subjects. Researchers (e.g. Gebhard et al. 2017; Wegner and Schmiedebach, 2020) have identified biology as the preferred STEM subject, mainly due to its connection to everyday life. Jansen (2015) identified that students have a more positive attitude towards biology, as it was seen as an easier subject compared to chemistry/physics.

Gender plays an important role in attitudes towards STEM subjects. It was shown that physics is preferable among males (Direito et al. 2017), whereas females prefer biology (Uitto et al. 2006). In chemistry, the gender situation is much more mixed. Direito et al. (2017) and Salta and Tzougraki (2004), found no evidence of gender differences, whereas less favorable attitudes were presented by Vilia and Candeias (2020). Cheung (2009) reports more positive attitudes toward females (Cheung, 2009).

Aim of the research, research questions, and hypotheses

The primary aim of this study is to assess and compare the effects of an IBSE approach (guided inquiry) versus traditional laboratory activities (confirmation and structured inquiry) on the intrinsic motivation and knowledge acquisition of high school students aged 13–14.

Derived from the research aim, the following questions guide this study:

RQ1. How does an IBSE approach to students with no previous experience, impact their learning motivation? Are the impacts dependent on the subject taught and students' gender?

RQ2. How does the students' learning motivation change over the intervention time with regular use of an IBSE approach (comparison between first and second time periods)?

RQ3. How does an IBSE approach affect students' knowledge acquisition? Are these impacts on knowledge dependent on the subject taught and students' gender?

Based on the research questions, the following hypotheses are proposed:

H1. (Linked to RQ1) An IBSE approach significantly enhances students' learning motivation compared to traditional laboratory activities.

H2. (Linked to RQ2) Continued exposure to an IBSE approach increases students' learning motivation over the school year.

H3. (Linked to RQ3) An IBSE approach positively affects learning outcomes compared to traditional laboratory activities.

Methods/Experimental

Participants. This case study was conducted during the school year of 2019/2020 in two parallel high school 8th classes (8A and 8B) during STEM subjects. One class was randomly assigned as the control group (CG, class 8A, 30 students), and the remaining was the experimental group (EG, class 8B, 32 students). In total, 62 students were included in this research, specifically 32 females and 30 males. A student was occasionally absent from an activity and did not complete the activity. The laboratory activities in EG used the IBSE approach at the guided inquiry instruction level.

The laboratory activities in CG were taught in a traditional way, with students following the instructions of their teacher. Both classes were taught by the same teacher who, at the time of the research, had four years of experience with an IBSE approach. The teacher was introduced to the concept of IBSE during the Teaching Inquiry with Mysteries Incorporated (TEMI) project (McOwan, 2016) in 2015. Part of the project includes the teacher receiving 2 years of training. Once the TEMI project ended, they completed several additional courses focussing on the IBSE approach. Moreover, the teacher was familiar with student-centered classroom principles (Keiler, 2018). By contrast, students had no or only very limited experience with this type of instruction before the research began. At the beginning of the research, the legal guardians of all students were informed how and why the pedagogical research would be conducted, and how the anonymity of their children would be assured.

Learning environments. Both groups were taught in a traditional way (structure inquiry) in their first laboratory activity of the year, and in an experimental way in their second laboratory activity (guided inquiry). The teaching in the first two laboratory sessions was designed differently to determine if both groups (EG and CG) were comparable. After these two lessons, the EG was taught using the IBSE approach, whereas the CG continued to be taught in the traditional way, as described above. Both groups were always taught the same topic. After analyzing several sources, reviewed laboratory activities created in the project entitled "Věda není žádná věda" [Science is no big deal] (Bludská et al. 2013; Vadasová et al. 2013; Volmutová et al. 2013) were selected as the source of the materials for the experimental group. These laboratory activities mimic real-life situations, and students in the EG used the 5Es instructional model: Engage, Explore, Explain, Elaborate, and Evaluate. These activities were modified for use in the CG by the authors of this paper so that they correspond to a traditionally conducted laboratory exercise, during which students work entirely according to the instructions (clear, step-by-step instructions) of the teacher. The teaching was designed in both the CG and EG as a 2 h (90 min) laboratory activity. The tasks with their brief characteristics are as follows:

- *September 2019: The surface of mammal bodies:* The task focused on observing the body coverings of different species of mammals, describing the structure of the skin, clarifying its meaning, stating the function, and calculating the surface area to volume ratios and its connection to breathing and heart rate. This topic was taught using only the control method of teaching.
- *October 2019: The difference between physical and chemical processes:* The goal was to distinguish between physical and chemical processes. Students proved with simple experiments whether it is a physical or chemical process and drew general conclusions based on the experiments. This topic was taught using only the experimental method of teaching.
- *November 2019: Separation methods:* The goal was to familiarize students with separation methods with an emphasis on everyday life. Students had to discover not only the principles, but also the advantages and disadvantages of individual separation methods, and apply them correctly.
- *December 2019: Winter sleep:* The task focussed on the energy requirements of different organisms during the winter period and their survival strategy. Students studied the properties of natural materials suitable for an animal's burrow, assessed their advantages, disadvantages, and heat maintenance.

- *January 2020: Preparation of solutions of a given concentration:* The goal of the task was to familiarize students with the properties of solutions with different salt concentrations. Students prepared their “micro-cocktail” in a straw and explored the properties of the “layers” of their cocktails.
- *February 2020: Construction of a fire extinguisher:* The goal was for students to be able to apply their skills to the construction of a fire extinguisher. Students assembled, tested, and evaluated their fire extinguishers, based on the principles of simple chemical reactions.
- *March 2020: Tooth cavity prevention:* In this task, students analyzed and interpreted the obtained data, described the structure of the tooth, and explained how pH in the mouth changes during the day or assessed the effect of different drinks on the development of dental caries.

Measures and questionnaires. Several research tools were used: (i) a standardized questionnaire: the Intrinsic Motivation Inventory, (ii) knowledge pretests, posttest, and retention tests, and (iii) interviews with students. To measure motivation, students were asked the following key questions:

- Are laboratory activities fun for students?
- Are students willing to put effort in during a laboratory activity in order to succeed?
- Do students feel under pressure while doing a laboratory activity?
- Is doing laboratory activities beneficial for students?

To obtain the desired results, the proposed didactic research focused on four categories: interest, effort, perceived pressure, and perceived usefulness. A standardized questionnaire was chosen as an appropriate tool—the Intrinsic Motivation Inventory (IMI) (McAuley et al. 1989; Ryan, 1982). The tool originally consisted of 45 items (statements) divided into 7 scales, namely: interest/enjoyment, effort/importance, pressure/tension, value/usefulness, perceived competence, perceived choice and relatedness. Due to the focus of the research, only items that related to the investigated scales were selected (interest/enjoyment, effort/importance, pressure/tension, and value/usefulness). The resulting research tool had only 16 items (which students found acceptable as it shortened the time spent filling out the questionnaire). The same items have also been used to monitor motivation in relation to STEM education in other research (e.g. Šmejkal et al. 2018). The first scale (interest/enjoyment) indicates the students’ perception of their interest in laboratory activities, with higher scores indicating higher interest. The second scale (effort/importance) illustrates how much energy and effort students are willing to put into the activity, with higher scores showing higher effort. The third scale (pressure/tension) reflects the students’ feelings toward the laboratory activity, with higher scores indicating that students felt less pressure or were more relaxed. The last scale (value/usefulness) represents the students’ perceived value of the laboratory activity, with higher scores indicating a higher level of perceived usefulness. All scales consist of four statements. The wording of all statements can be obtained online at <https://selfdeterminationtheory.org/intrinsic-motivation-inventory/> (account creation/login required). The respondents used the Likert scale to express how much they agreed or disagreed with each statement by scoring the responses from 1 to 7, with 1 meaning “Totally Disagree” and 7 “Totally Agree” (Pintrich et al. 1991; Ryan, 1982). The questionnaire was given to both the EC and CG students at the end of each laboratory session. One of the advantages of this

research tool is that it is designed to be flexible and modular according to research needs. Therefore, it is not necessary to use it in its complete version (Markland and Hardy, 1997; Pintrich et al. 1991; Rotgans and Schmidt, 2010), resulting in the other three scales not being used in the research.

Knowledge tests were designed for each laboratory session—4 tests related to Chemistry and 3 tests related to Biology. These tests aimed to measure the influence of IBSE on the students’ level of acquired knowledge and skills. Each test consisted of three types of tasks—open, closed-ended, and multiple-choice tasks. Some tasks were not only knowledge-based but required complex skills—e.g. formulating a hypothesis, planning laboratory work, and interpreting charts. Each test was evaluated by a panel of experts who approved the content validity of the tool. The panel consisted of 5 experts: 3 Chemistry/Biology teachers and 2 education researchers specializing in IBSE approaches in the Czech Republic. The education researchers are recognized experts in the field. All of the teachers have many years of experience in teaching IBSE in high schools and have also participated in the TEMI project. The experts’ thorough review and approval of the questionnaire and each test ensured its appropriateness and relevance for the study. The students took each test three times. The first (pretest) was completed one week before each laboratory session, the second (posttest) was performed within one week of each laboratory session, and the third (retention test) was undertaken 3–4 months later.

All data were statistically processed in IBM SPSS Statistics 25. Significance was set at $\alpha = 0.05$. The effect size was interpreted using Cohen’s *d*. Reversed items were recoded by flipping the response scale allowing the scoring to be consistent with the other items. The “pressure” scale items were recoded so that the statement with a higher score corresponds to the feelings that the students feel when relaxed. Statements with a lower score then meant that the students felt under pressure. Due to the small sample size, the data were not normally distributed. For this reason, non-parametric analysis methods were used, namely the two-sample Mann–Whitney *U*-test and the paired Wilcoxon’s test. The time-scale of the use of the research tools listed above (knowledge tests, questionnaires, and interviews) is shown in Fig. 1.

The reliability of the questionnaire was evaluated based on Cronbach’s alpha (α) for each scale of the Intrinsic Motivation Inventory (interest/enjoyment 0.91, effort/importance 0.75, pressure/tension 0.75 and value/usefulness 0.96) and for each knowledge test (Cronbach’s alpha ranged from 0.71 to 0.82). Based on the results, Cronbach’s alpha is considered acceptable because it surpassed 0.70 in all scales (Nunnally, 1978) and in each knowledge test (George and Mallery, 2003). The data were internally consistent and thus reliable.

Lastly, the students were interviewed in groups by their teacher as soon as possible after each laboratory session. All audio recordings of interviews lasted between 10 and 15 min and were subsequently analyzed by the authors of this paper. Students’ responses were transcribed and sorted by the frequency of their occurrence; they were later used to support and interpret the results from the quantitative part of the pedagogical research. Students expressed their feelings about and impressions of the laboratory activity. The purpose of the semi-structured interviews was to find out the students’ perceptions concerning how they worked during the laboratory activities, what specifically interested them or not about the activity, what caused them problems, why they felt under pressure, and why they considered some activities to be more important, etc. However, a detailed analysis of the interviews was not the aim of the research. The aim of the interviews was simply to further explain the students’ observed attitudes.

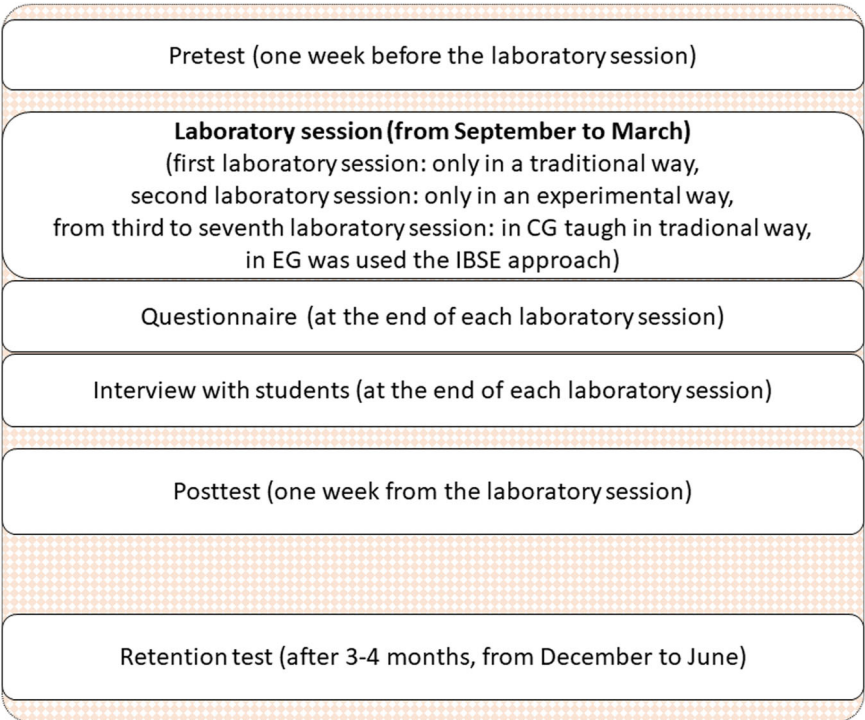


Fig. 1 Design of pedagogical research. The time-scale of the knowledge tests, questionnaires, and interviews.

Table 2 Overall evaluation of the laboratory activities by Intrinsic Motivation Inventory questionnaires.						
Group	Variables	N	Mean	Standard mean error	Standard deviation	Median
CG	interest/enjoyment	185	6.18	0.08	1.06	6.50
	effort/importance	185	5.68	0.08	1.14	5.75
	pressure/tension	185	5.83	0.09	1.28	6.25
	value/usefulness	185	5.94	0.10	1.41	6.50
EG	interest/enjoyment	199	6.12	0.08	1.09	6.50
	effort/importance	199	5.76	0.07	0.94	6.00
	pressure/tension	199	5.76	0.09	1.29	6.25
	value/usefulness	199	5.95	0.09	1.31	6.50

CG control group, EG experimental group.

Results

Overall evaluation of the laboratory activities. Based on the students’ feedback, all 7 laboratory activities were evaluated highly positively because the averages of all variables ranged from 5.68 to 6.18 on a scale from 1 to 7 (see Table 2). Therefore, the laboratory activities sparked the students’ interest, the students were willing to put some effort into succeeding in tasks, they did not feel pressure during the activities, and they found the tasks useful.

The influence of an IBSE approach on students’ motivation
Differences in students’ motivation between CG and EG (RQ1). The first research question (RQ1) was answered using a two-step approach). Step 1 (A) examined the differences between the CG and EG during the first and second sessions to determine whether both groups were balanced before the implementation of the research. In the second step (B), differences between the CG and EG during the 3rd–7th sessions were compared to assess the effect of the IBSE approach. Differences were then also examined in relation to the subject taught and the student’s gender.

Comparability of CG and EG before the start of the research: Firstly, results of the CG and EG ($N_{CG} = 30$, $N_{EG} = 32$) were compared regarding the students’ attitudes towards (i) the traditional instruction, based on questionnaires from the first traditional laboratory session and (ii) the experimental instruction, based on questionnaires from the sessions based on a guided inquiry activity. The Mann–Whitney *U*-test did not show any significant difference between the CG and EG in the first session (the *p*-value of all *U*-tests was higher than 0.05: interest: $p = 0.83$, $d = 0.06$, $Md_{CG} = 6.00$, $Md_{EG} = 6.00$, $U = 406.5$; effort: $p = 0.31$, $d = 0.27$, $Md_{CG} = 5.50$, $Md_{EG} = 5.88$, $U = 354.5$; pressure: $p = 0.56$, $d = 0.15$, $Md_{CG} = 6.25$, $Md_{EG} = 5.63$, $U = 383.0$; value: $p = 0.80$, $d = 0.06$, $Md_{CG} = 6.25$, $Md_{EG} = 6.50$, $U = 404.0$, as shown in Table 3).

Similar results were achieved when comparing groups in the second laboratory session which was conducted as an inquiry laboratory activity. The Mann–Whitney *U*-test did not reveal any significant difference between the CG and EG in the second session either (the *p*-value of all tests was higher than 0.05).

The students’ answers were not significantly different on any scale of motivation in either of the laboratory sessions. All students felt a similar level of interest in the subject matter, and

they were willing to put a similar amount of effort into the activities. Furthermore, during the laboratory session, they felt a similar level of pressure and usefulness. No differences were found in any of the initial laboratory sessions, nor when assessing the data with respect to gender (the p -value of all tests was higher than 0.05).

Comparison of the control and experimental instruction: A comparison was made between the results of the CG and EG questionnaires from other laboratory sessions (3rd–7th). The findings were merged, resulting in a higher sample size for both the CG and EG ($N_{CG} = 127$, $N_{EG} = 130$). However, the results of the Mann-Whitney U -test did not show any significant difference between the groups (the p -value of all 4 tests was higher than 0.05, interest: $p = 0.43$, $d = 0.10$, $Md_{CG} = 6.75$, $Md_{EG} = 6.50$, $U = 8474$; effort: $p = 0.82$, $d = 0.03$, $Md_{CG} = 6.00$, $Md_{EG} = 6.00$, $U = 8810$; pressure: $p = 0.33$, $d = 0.12$, $Md_{CG} = 6.50$, $Md_{EG} = 6.25$, $U = 8352$; value: $p = 0.99$, $d = 0.00$, $Md_{CG} = 6.50$, $Md_{EG} = 6.50$, $U = 8943.5$). Additionally, no significant difference was found in any scale when analyzing the questionnaires' results. The results, therefore, show that the students felt a similar level of interest and pressure, put an equal amount of effort into the activities, and perceived a similar value of the activities, regardless of instruction.

The next step examined whether the impact was dependent on the subject taught or the gender of the student. The students of the EG and CG were divided according to the subject and

assessed in isolated subgroups; the results for Chemistry and Biology were analyzed separately. Data were processed using the Mann-Whitney U -test. Subsequently, in these isolated groups, the impact of gender was assessed (see Table 4).

Two significant differences were found between Chemistry (c.) students in the EG and CG ($N_{CG, c.} = 72$, $N_{EG, c.} = 87$): (1) The CG students expressed more interest than the EG students (small effect; $p = 0.03$; $d = 0.36$, $Md_{CG} = 6.75$, $Md_{EG} = 6.50$, $U = 2520$), regardless of gender. (2) The CG students felt less pressure than the EG students (medium effect; $p = 0.01$, $d = 0.42$, $Md_{CG} = 6.75$, $Md_{EG} = 5.75$, $U = 2387.5$). Females (f.) compared in the CG and EG, felt under higher pressure in the EG ($p = 0.05$, $d = 0.44$, $N_{CG, c. f.} = 72$, $N_{EG, c. f.} = 87$, $Md_{CG} = 6.75$, $Md_{EG} = 5.62$, $U = 674.5$). However, there were no differences between males. Dividing Chemistry students into female and male subgroups showed that the CG females were willing to put more effort into the educational process than the EG female students (medium effect, $p = 0.03$; $d = 0.47$, $N_{CG, c. f.} = 72$, $N_{EG, c. f.} = 87$, $Md_{CG} = 6.50$, $Md_{EG} = 6.00$, $U = 656.5$). This was not the same for the male students.

A significant difference was also found in Biology (b.): The CG students felt more pressure than their peers in the EG ($p = 0.04$, $d = 0.39$, $N_{CG, b.} = 55$, $N_{EG, biology} = 54$, $Md_{CG} = 5.50$, $Md_{EG} = 6.00$, $U = 1160.5$). This was in contrast to the findings from the Chemistry laboratory activity. The second distinction was that there were no differences between male and female subgroups.

The influence of the intervention time of regular usage of the IBSE approach on students' motivation (RQ2). The second research question was answered by dividing the obtained regardless of school subject data into 2 equal periods—the first period (f. p.) from October to December 2019 (containing the first 3 laboratory sessions) and the second period (s. p.) from January to March 2020 (the last 3 laboratory sessions). It was observed that students' attitudes towards the experimental instruction changed over time. A Mann-Whitney U -test was used to determine differences in results between the first and second periods ($N_{EG, f. p.} = 113$, $N_{EG, s. p.} = 86$). The results showed a significant increase in two scales, specifically: interest/enjoyment (small effect; $p = 0.01$; $d = 0.38$, $Md_{CG} = 6.25$, $Md_{EG} = 7.00$, $U = 3788.5$) and pressure/tension (medium effect; $p = 0.001$; $d = 0.46$, $Md_{CG} = 5.75$, $Md_{EG} = 6.50$, $U = 3590$; (a reminder: a higher score corresponded to the feeling of being more relaxed). During the second period, students were more interested in the subject matter and they felt less pressure.

Table 3 Traditional teaching style comparison between CG and EG in terms of intrinsic motivation (after the first laboratory session).

	Interest/ enjoyment	Effort/ importance	Pressure	Value/ usefulness
Md _{CG} (SD)	6.00 (1.42)	5.50 (1.10)	6.25 (1.61)	6.25 (1.71)
Md _{EG} (SD)	6.00 (1.34)	5.88 (0.88)	5.63 (1.38)	6.50 (1.30)
Mann-Whitney U	406.5	354.5	383.0	404.0
Wilcoxon W	812.5	819.5	789.0	869.0
Z	−0.21	−1.02	−0.58	−0.26
p -value (2-tailed)	0.83	0.31	0.56	0.80
Cohen's d	0.06	0.27	0.15	0.07

CG control group, EG experimental group, SD standard deviation, Md median.
Number of students in CG = 30 and in EG = 28.

Table 4 Comparison between experimental and traditional teaching styles in terms of intrinsic motivation by school subject (third to seventh laboratory sessions).

	Chemistry				Biology			
	Interest/ enjoyment	Effort/ importance	Pressure	Value/ usefulness	Interest/ enjoyment	Effort/ importance	pressure	value / usefulness
Md _{CG} (SD)	6.75 (0.74)	6.13 (1.19)	6.75 (1.25)	6.75 (1.30)	6.50 (1.10)	5.50 (1.17)	6.25 (1.20)	6.50 (1.57)
Md _{EG} (SD)	6.50 (1.07)	6.00 (0.91)	5.75 (1.23)	6.50 (1.33)	6.50 (1.05)	6.00 (1.07)	6.75 (1.34)	6.63 (1.44)
N_{CG}	72	72	72	72	55	55	55	55
N_{EG}	87	87	87	87	54	54	54	54
Mann-Whitney U	2520.0	2836.0	2387.5	2846.5	1261.0	1241.5	1160.0	1308.0
Wilcoxon W	6348.0	6664.0	6215.5	6674.5	2801.0	2781.5	2700.0	2848.0
Z	−2.21	−1.03	−2.61	−1.01	−1.39	−1.48	−2.02	−1.10
p -value (2-tailed)	0.03 ^a	0.30	0.01 ^a	0.31	0.17	0.14	0.04 ^a	0.27
Cohen's d	0.36	0.17	0.42	0.16	0.27	0.29	0.39	0.21

CG control group, EG experimental group, SD standard deviation, Md median.

^aSignificant at the level of 0.05.

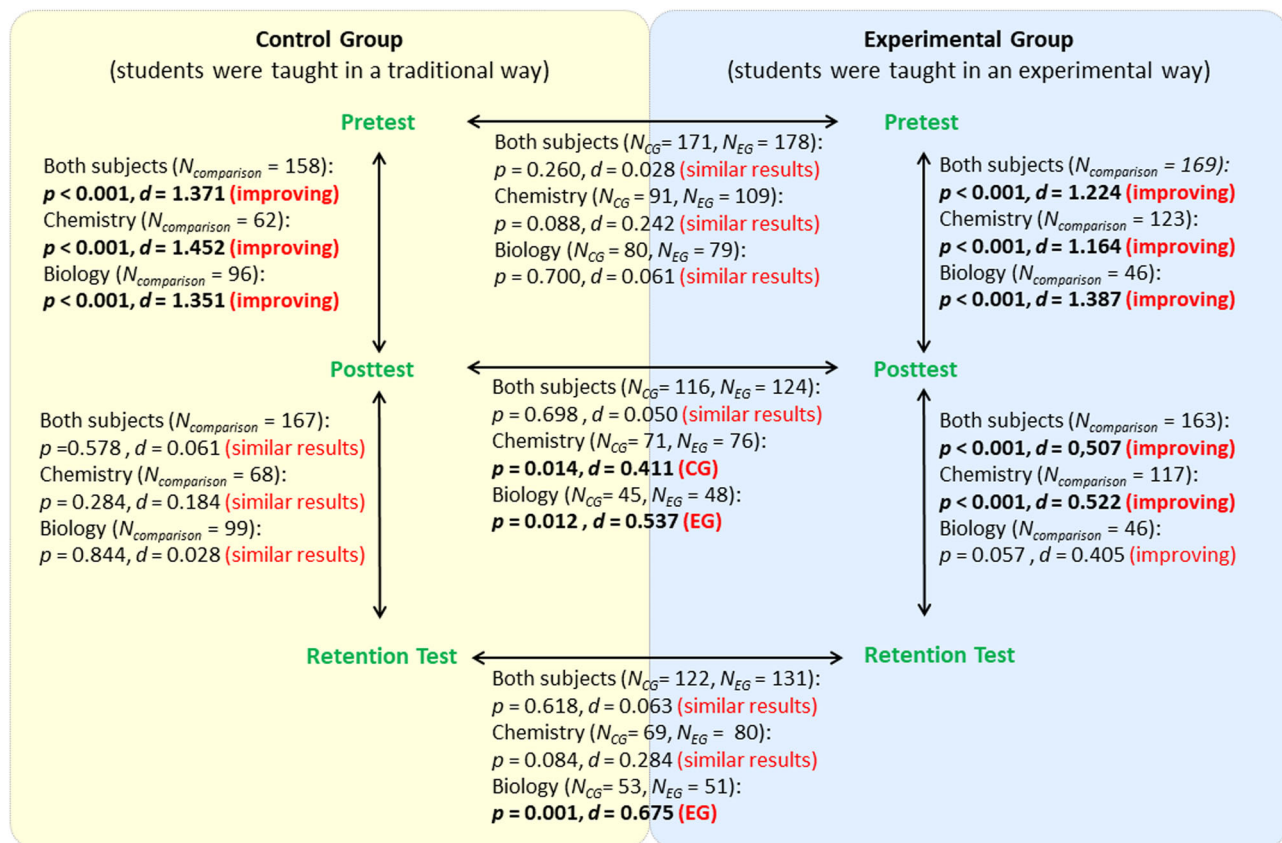


Fig. 2 Comparison of the control (CG) and experimental groups (EG). Differences in knowledge test scores (pretest, posttest, and retention test) between CG and EG (significant at the level of 0.05, the brackets indicate the group with a higher score). N number of students, p p -value of the Mann-Whitney U -test, d Cohen's d .

The influence of an IBSE approach on the level of acquired knowledge (RQ3). To answer the third research question (RQ3), it was necessary to compare the data at two different levels (Fig. 2). In the first level (A-horizontal comparison Fig. 2), differences in pretest, posttest, and retention test results between CG and EG students were examined. In the second level (B-vertical comparison in Fig. 2), differences between pretest and posttest results and between posttest and retention test results of the same students were observed to determine whether there was an improvement in their performance.

Differences in pretests, posttests, and retention tests between CG and EG students

The comparison of pretests: The first step was to determine whether there were statistically significant differences (whether the groups were balanced before the start of each laboratory activity) in the results of the pretests between the CG and EG students. The results of knowledge tests for the 1st–7th sessions were merged, resulting in a higher sample size for both the CG and EG. The Mann-Whitney U -test was chosen as a suitable statistical method and no significant differences in the results of the pretests were found when comparing CG and EG students. CG and EG students achieved similar results ($p = 0.26$, $d = 0.028$, $N_{CG} = 171$, $N_{EG} = 188$, $Md_{CG} = 8.00$, $Md_{EG} = 7.00$, $U = 14969$), thus acquiring similar levels of knowledge on both subjects. No differences were found even when assessing each subject separately (Fig. 2-first horizontal comparison).

The comparison of posttest and retention tests: Differences between posttests and retention tests were calculated for all laboratory sessions using both teaching methods, i.e., for the

3rd–7th laboratory sessions. The findings from the knowledge tests for these sessions were merged, resulting in a higher sample size for both the CG and EG. The Mann-Whitney U -test was chosen as a suitable statistical method and no significant differences in the results of the posttests and retention tests were found when comparing CG and EG students (posttest: $p = 0.70$, $d = 0.05$, $N_{CG, posttest} = 116$, $N_{EG, posttest} = 124$, $Md_{CG} = 10.00$, $Md_{EG} = 10.25$, $U = 6984$; retention test: $p = 0.62$, $d = 0.06$, $N_{CG, retention_test} = 122$, $N_{EG, retention_test} = 131$, $Md_{CG} = 9.50$, $Md_{EG} = 10.0$, $U = 7702$).

However, significant differences were found when analyzing the results of the posttests and retention tests of CG and EG students, for each subject separately. In Chemistry, CG students scored significantly higher marks in the posttest than their peers in the EG ($p = 0.01$, $d = 0.41$, $N_{CG} = 71$, $N_{EG} = 76$, $Md_{CG} = 9.00$, $Md_{EG} = 8.50$, $U = 2068.5$) regardless of gender, but not in the retention test ($p = 0.08$, $d = 0.28$, $N_{CG} = 69$, $N_{EG} = 80$, $Md_{CG} = 9.00$, $Md_{EG} = 9.00$, $U = 2309$).

In Biology, EG students scored significantly higher in the posttest ($p = 0.01$, $d = 0.54$, $N_{CG} = 45$, $N_{EG} = 48$, $Md_{CG} = 10.00$, $Md_{EG} = 11.50$, $U = 754.5$) and in the retention test ($p = 0.001$, $d = 0.68$, $N_{CG} = 53$, $N_{EG} = 51$, $Md_{CG} = 10.00$, $Md_{EG} = 12.50$, $U = 850$) than CG students, regardless of gender. Significant differences were found between females in the retention test ($p = 0.02$, $d = 0.70$, $N_{CG} = 27$, $N_{EG} = 25$, $U = 207.5$) and between males in both the posttest ($p = 0.05$, $d = 0.61$, $N_{CG} = 22$, $N_{EG} = 24$, $U = 174$) and the retention test ($p = 0.02$, $d = 0.67$, $N_{CG} = 26$, $N_{EG} = 26$, $U = 212.5$). These results will be discussed further.

Differences in the level of knowledge over time. The differences in knowledge test scores throughout the study were examined on an

individual student basis and per subject. Areas assessed included the difference between (i) pretest and posttest results and (ii) posttest and retention test results (Fig. 2-vertical comparisons). The Wilcoxon paired test was chosen as a suitable statistical method. The results of knowledge tests from the first laboratory sessions were included in the CG (students were taught in a traditional way [t.w.]), whereas the results from the second laboratory sessions were included in the EG (students were taught in an experimental way [e.w.]). The results from these laboratory sessions were merged, resulting in a higher sample size for both the CG and EG (number of comparisons: $N_{CG, \text{ t.w., pretest-w- posttest}} = 158$, $N_{EG, \text{ e.w., pretest-posttest}} = 169$; $N_{CG, \text{ t.w., posttest-retention_test}} = 167$, $N_{EG, \text{ e.w., posttest-retention_test}} = 163$).

The comparison of pretest and posttests: Comparing pretests with posttests showed that both CG and EG students scored significantly higher in the posttest than in the pretest, with a large size effect ($p < 0.001$, $d > 1.1$). This trend in the data was expected because students developed their knowledge and skills during the laboratory activities-regardless of group (CG and EG), subject, or gender.

The comparison of posttest and retention test: Scores revealed a significant improvement when students were taught in an experimental way ($p < 0.001$, $d = 0.51$, $N_{EG, \text{ posttest-retention_test}} = 163$, $Md_{\text{posttest}} = 13.00$, $Md_{\text{retention_test}} = 14.0$, $Z = -4.43$) as opposed to a traditional way ($p = 0.58$, $d = 0.06$, $N_{CG, \text{ posttest-retention_test}} = 167$, $Md_{\text{posttest}} = 11.00$, $Md_{\text{retention_test}} = 11.00$, $Z = -0.56$). A more detailed analysis of the posttest versus the retention test by subject showed a significant improvement in knowledge of Chemistry only when students were taught in an experimental way ($p < 0.001$, $d = 0.52$, $N_{EG, \text{ posttest-retention_test}} = 117$, $Md_{\text{posttest}} = 16.00$, $Md_{\text{retention_test}} = 18.00$, $Z = -3.86$), among females ($p = 0.01$, $d = 0.47$, $N_{EG, \text{ posttest-retention_test}} = 63$, $Md_{\text{posttest}} = 17.00$, $Md_{\text{retention_test}} = 17.00$, $Z = -2.56$) and males ($p = 0.003$, $d = 0.58$, $N_{EG, \text{ posttest-retention_test}} = 54$, $Md_{\text{posttest}} = 16.00$, $Md_{\text{retention_test}} = 18.50$, $Z = -2.923$). A nonsignificant decrease was identified in test scores when students were taught in a traditional way ($p = 0.284$, $d = 0.184$, $N_{CG, \text{ posttest-retention_test}} = 68$, $Md_{\text{posttest}} = 9.00$, $Md_{\text{retention_test}} = 9.00$, $Z = -1.071$).

Although an improvement in knowledge of Biology was also observed when students were taught in an experimental way, the difference was nonsignificant (regarding both genders, $p = 0.057$, $d = 0.405$, $N_{EG, \text{ posttest-retention_test}} = 46$, $Md_{\text{posttest}} = 11.50$, $Md_{\text{retention_test}} = 12.00$, $Z = -1.902$). However, closer analysis revealed that females improved significantly ($p = 0.01$, $d = 0.836$, $N_{EG, \text{ posttest-retention_test}} = 23$, $Md_{\text{posttest}} = 11.00$, $Md_{\text{retention_test}} = 14.00$, $Z = -2.62$). The Biology test yielded similar results (though a small nonsignificant improvement was achieved) when students were taught in a traditional way ($p = 0.84$, $d = 0.03$, $N_{CG, \text{ posttest-retention_test}} = 99$, $Md_{\text{posttest}} = 12.00$, $Md_{\text{retention_test}} = 12.00$, $Z = -0.20$).

The results of interviews. The interview results corresponded with those obtained from the questionnaires (as described in the previous section). However, it was possible to interpret the data obtained from the quantitative research more effectively. Student responses and their frequency of responses are mentioned to support quantitative results in the “Discussion” section.

Discussion

The influence of an IBSE approach on students’ motivation. The first hypothesis (H1), “An IBSE approach significantly enhances students’ learning motivation compared to traditional laboratory activities”, was not confirmed overall using combined

data (regardless of subject or gender) because the results were comparable in both groups. Cobern et al. (2010), Chatterjee et al. (2009), and Kalyuga et al. (2010) reached similar conclusions also unable to confirm the positive impact of an IBSE approach on students learning natural sciences. However, further analysis revealed significant differences between subjects, genders, and the first and second periods of the study.

In Chemistry, the IBSE approach decreases the students’ interest in laboratory activities while increasing the pressure on female students and consequently lowering their effort in completing the activities. This causal relation was supported by the students’ interviews, mainly by females, as reported also by Vilja and Candeias (2020) or by different gender patterns (Freedman et al. 2023; Stolk et al. 2021). These trends may also be caused by a lack of experience with the IBSE approach, as confirmed by Hattie (2008), who highlights the need for incorporating IBSE more frequently into lower education stages-even during primary education. Similar conclusions were reached through the analysis of students’ interviews. Students were not used to working independently, having to propose and verify hypotheses without the step by step guidance from the teacher they were accustomed to. In the second laboratory activities especially, approximately half of the students reported a lack of workflow they were used to, which limited their interest in participating in the activity and eventually led them to decrease their efforts.

Critics of the IBSE approach often cite a decline in student interest due to its lack of explicit instructions (Klahr and Nigam, 2004; Mayer, 2004). Moreover, IBSE can push students out of their comfort zones, leading to feelings of disorientation during laboratory activities (Kalyuga et al. 2010). This discomfort can heighten perceived pressure, particularly among female students in chemistry laboratory activities, potentially fostering negative attitudes and decreased effort (Kalyuga et al. 2010; Gebhard et al. 2017) as confirmed in the research among females based on their interviews (approximately 1/3 of females).

Despite following the provided instructions, approximately one-third of female students reported higher stress levels due to IBSE’s time-consuming nature (Sadeh and Zion, 2009). This was made worse by the absence of previous experience of time-planning activities, leading to fear of mistakes and incomplete work. To address these challenges, Fussy et al. (2023) suggest linking IBSE activities to more practical, real-world scenarios. Additionally, as demonstrated in Biology studies (Jansen, 2015), providing encouragement and reassurance can significantly benefit female students. Conversely, many male students appreciated the opportunity to explore different approaches, highlighting the potential influence of gender-specific behavioral patterns on IBSE outcomes (Cooper et al. 2015; Secker, 2010).

Unlike in Chemistry, Biology classes showed different trends, with EG students feeling more relaxed than CG students regardless of gender. Interviews confirmed that students perceived Chemistry activities (regardless of the group) as being more difficult than Biology activities. Their responses indicated that Chemistry activities are more difficult due to frequent abstract concepts as well as Chemistry being a new subject. Conversely, Biology activities were described by more than 50% of students as having a context closer to everyday life and thus easier to solve, confirming Jansen’s conclusions (2015). A positive effect of the IBSE approach on students’ attitudes toward Biology was also proved by Sandika and Fitrihidajati (2018).

As the students lacked any previous experience with an IBSE approach, the negative effects of IBSE could be eliminated by regularly using this approach and starting from the confirmation level. The second hypothesis (H2), “Continued exposure to an IBSE approach increases students’ learning motivation over the school year”, was validated. EG students described a higher

interest and felt less pressure in the second period than in the first period. This conclusion corroborates previous studies (e.g. Kanter and Konstantopoulos, 2010).

The influence of an IBSE approach on knowledge and skills.

The third hypothesis (H3), “An IBSE approach positively affects learning outcomes compared to traditional laboratory activities” was evaluated. The comparison of the posttest results between CG and EG students showed that EG students scored (i) significantly lower than CG students in Chemistry and (ii) significantly higher than CG students in Biology. These results show that IBSE does not affect the level of knowledge acquisition in Chemistry, as previously concluded by Lechová (2014). Conversely, it was found that an IBSE approach positively affects the level of acquired knowledge in Biology, as conceded in the findings of Radvanová (2017). To examine long-term changes, the scores between posttests and retention tests were compared, revealing that: (i) in Chemistry, only students who were taught in an experimental way, regardless of gender, scored significantly higher in retention tests compared with posttest; (ii) in Biology, students who were taught in an experimental way, regardless of gender also scored higher in retention tests nonsignificantly but the retention was significantly higher for females. The gender differences found are in contrast to Ješková et al. (2016), who evaluated the efficacy of the IBSE approach in Mathematics, Physics, and Informatics, or in PISA 2012 results (OECD, 2014), which assessed students performance in science in general. The results were also supported by the measured differences between the CG and EG in the retention test, where EG students scored significantly higher than CG students in Biology. The results in Chemistry in the retention tests between CG and EG were comparable as EG students performed worse in the posttests. Despite significant improvements in long-term knowledge, EG students achieved statistically the same results as the CG students, who showed no improvement.

The data shows that an IBSE approach has a positive, long-term effect on the knowledge of high school students, resulting in increased autonomy in problem-solving. This trend was also reported by Freeman et al. (2014). The IBSE approach encourages students to search for answers on their own, promoting the development of long-term memory. Schmid and Bogner (2015) also noted an increase in knowledge acquisition. The results are consistent with published research comparing the IBSE approach with traditional approaches to teaching and learning (Blanchard et al. 2010; Chang and Mao, 2010; Furtak and Alonzo, 2009; Lechová, 2014; Minner et al. 2010). An IBSE approach appears suitable for both male and female students, as suggested by the results of Akcay and Yager (2016) the IBSE approach did produce a stronger positive effect on females in Biology.

Study limitations. The results of the research have several limitations. Only two classes from a single high school were included in the study, therefore the results cannot be generalized for the whole population. Nevertheless, the data were collected between September 2019 and March 2020, generating a sufficient number of questionnaires and knowledge tests to perform statistical analyses. A further limitation was the difference in the popularity/attractiveness of the studied topic, but this was compensated for by teaching the same topics in the CG (traditionally) and EG (applying the IBSE approach). There were also fewer students in Biology (3 laboratory activities) than in Chemistry (4 laboratory activities) due to the closure of Czech schools during the COVID-19 pandemic. A potential limitation of the research is also the risk of data amplification due to repeated measurements from the same participants over several months.

Conclusions

The long-term effects of an IBSE approach on students were evaluated. No statistically significant difference was found in motivation levels between the CG (using confirmation and structured inquiry) and EG (using guided inquiry) during the research in tested STEM subjects. Based on the student's feedback, all 7 laboratory activities were evaluated highly positively, which confirms the significance of implementing laboratory activities into teaching in general. Students reported similar levels of interest and pressure, put an equal amount of effort into the activities, and perceived a similar value in the activities regardless of instruction. Nevertheless, differences in gender, subject, and the length of the pedagogical experiment were found.

When focusing on Chemistry during the first period (months 2–4), a lack of experience with the IBSE approach decreases the students' interest (both genders), and for female students, it also increases the pressure during laboratory activities and decreases the effort put into the educational process (The above-mentioned unfavorable effects of an IBSE approach may be minimized by starting with confirmation inquiry, and by regularly incorporating this approach into the educational process as soon as Chemistry teaching commences). Without implementing these recommendations, it is unlikely that overall interest in STEM would rise, and the existing gender disparity, with only 30% of STEM students being female, would likely persist. Previous findings are supported by the results indicating that long-term (months 4–7) experience with IBSE increases motivation and interest in the subject matter, reduces perceived levels of pressure, and increases the activity's perceived value-data showed that students needed more time to familiarize themselves with guided inquiry compared to confirmation/structured inquiry. With Biology, however, the IBSE approach decreased the students' perceived pressure, and students successfully followed the guided inquiry from the beginning of the laboratory activities.

Regarding knowledge acquisition in Chemistry, CG students scored higher immediately after the laboratory session, but EG students scored significantly higher in tests after a 3-month period (i.e. retention tests). In Biology, EG students scored higher in posttests and retention tests. The research proved that an IBSE approach has a positive long-term effect on the acquired knowledge of high school students in Biology and Chemistry.

To conclude, the guided inquiry was more difficult for students in the more abstract subject of Chemistry than Biology, and female students experienced more difficulties than males. Providing more support and motivation to female students at the beginning of laboratory activities is recommended; this should increase gender equity in STEM education and future careers (Ortiz-Martínez et al. 2023). Despite these differences, an IBSE approach improved long-term knowledge acquisition in both subjects.

A longitudinal study should be undertaken to assess the effects of IBSE on students' motivation over a few years. These suggestions aim to highlight the importance of continued research aimed at optimizing and understanding the long-term impact of IBSE approaches in STEM education. IBSE should be gradually incorporated into STEM lessons and teachers of these subjects should be encouraged to use an IBSE approach frequently. Teachers acquainted with active learning strategies (Nguyen et al. 2021), will enable students to work more independently, and creating a safe environment where students are open to learning from their own mistakes.

Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Received: 18 February 2024; Accepted: 15 January 2025;

Published online: 22 February 2025

References

- Akçay H, Yager RE (2016) Students learning to use the skills used by practicing scientists. *Eurasia J Math Sci Technol Educ* 12(3):513–525. <https://doi.org/10.12973/EURASIA.2015.1395A>
- Banchi H, Bell R (2008) The many levels of inquiry. *Sci Child* 46(2):26–29
- Bell RL (2004) Perusing Pandora's box: exploring the what, when, and how of nature of science instruction. In: Flick LB, NG Lederman NG (eds) *Science inquiry and nature of science. Implications for teaching, learning, and teacher education*. Kluwer Academic Publisher, pp. 427–446
- Bell RL, Smetana L, Binns I (2005) Simplifying inquiry instruction. *J Sci Teacher Educ* 72(7):30–34
- Bell T, Urhahne D, Schanz S, Ploetzner R (2010) Collaborative inquiry learning: models, tools, and challenges. *Int J Sci Educ* 32(3):349–377. <https://doi.org/10.1080/09500690802582241>
- Blanchard MR, Southerland SA, Osborne JW, Sampson VD, Annetta LA, Granger EM (2010) Is inquiry possible in light of accountability? A quantitative comparison of the relative effectiveness of guided inquiry and verification laboratory instruction. *Int J Sci Educ* 94(4):577–616. <https://doi.org/10.1002/SCE.20390>
- Bludská M, Balounová R, Bojkovský M, Fendrychová A, Hájková Z, Kozlovská L, Mokrejšová A, Prazienka M, Vozka J, Žabová I (2013) *Chemie-Výukové materiály pro střední školy*, 1. vydání [Chemistry-teaching tasks for high schools, 1st edn]. Conatex-didactic
- Bunterm T, Lee K, Kong JNL, Srikoorn S, Vangpoomyai P, Rattavongsa J, Rachahoon G (2014) Do different levels of inquiry lead to different learning outcomes? A comparison between guided and structured inquiry. *Int J Sci Educ* 36(12):1937–1959. <https://doi.org/10.1080/09500693.2014.886347>
- Cairns D, Areepattamannil S (2017) Exploring the relations of inquiry-based teaching to science achievement and dispositions in 54 countries. *Res Sci Educ* 49(1):1–23. <https://doi.org/10.1007/S11165-017-9639-X>
- Carpinetti M, Childs PE, Dittmar J, Eiliks I, Fortus D, Giliberti M, Hofstein A, Jordan J, Katzevich D, Mamlok-Naaman R, Peleg R, Sherborne T, Yayon M (2015) Teaching the TEMI way: how using mysteries supports science learning. Queen Mary University
- Castle P, Buckler S (2021) *Teach psychology*. Sage Publications
- Chang C-Y, Mao S-L (2010) Comparison of Taiwan science students' outcomes with inquiry-group versus traditional instruction. *J Educ Res* 92(6):340–346. <https://doi.org/10.1080/00220679909597617>
- Chatterjee S, Williamson VM, McCann K, Peck ML (2009) Surveying students' attitudes and perceptions toward guided-inquiry and open-inquiry laboratories. *J Chem Educ* 86(12):1427–1432. <https://doi.org/10.1021/ed086p1427>
- Cheung D (2009) Students' attitudes toward chemistry lessons: the interaction effect between grade level and gender. *Res Sci Educ* 39(1):75–91. <https://doi.org/10.1007/s11165-007-9075-4>
- Cobern WW, Schuster D, Adams B, Applegate B, Skjold B, Undreiu A, Loving CC, Gobert JD (2010) Experimental comparison of inquiry and direct instruction in science. *Res Sci Educ* 28(1):81–96. <https://doi.org/10.1080/02635140903513599>
- Cooper TE, Bailey B, Briggs KS (2015) Gender differences in achievement in an inquiry-based learning precalculus course. *IEJME* 10(2):97–110. <https://doi.org/10.29333/iejme/294>
- Direito I, Connolly J, Simon S, Trevethan J (2017) Gender differences in science interest, self-concept and career aspirations. In: Finlayson O, McLoughlin E, Erduran S, Childs P (eds) *Proceedings of the ESERA 2017 Conference*. European Science Education Research Association, Dublin, Ireland
- Deci EL, Koestner R, Ryan RM (1999) A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychol Bull* 125(6):627–668. <https://doi.org/10.1037/0033-2909.125.6.627>
- Eastwell P (2009) Inquiry learning: elements of confusion and frustration. *Am Biol Teach* 71(5):263–264. <https://doi.org/10.2307/27669426>
- European Commission (2019) Key competences for lifelong learning. Publications Office. <https://data.europa.eu/doi/10.2766/569540>
- Freedman G, Green MC, Kussman M, Drusano M, Moore MM (2023) “Dear future woman of STEM”: letters of advice from women in STEM. *Int J STEM Educ* 10:20. <https://doi.org/10.1186/s40594-023-00411-0>
- Freeman S, Eddy SL, McDonough M, Smith MK, Okoroafor N, Jordt H, Wenderoth MP (2014) Active learning increases student performance in science, engineering, and mathematics. *Proc Natl Acad Sci USA* 111(23):8410–8415. <https://doi.org/10.1073/PNAS.1319030111/-DCSUPPLEMENTAL>
- Furtak EM, Alonzo AC (2009) The role of content in inquiry-based elementary science lessons: an analysis of teacher beliefs and enactment. *Res Sci Educ* 40(3):425–449. <https://doi.org/10.1007/S11165-009-9128-Y>
- Furtak EM, Seidel T, Iverson H, Briggs DC (2012) Experimental and quasi-experimental studies of inquiry-based science teaching: a meta-analysis. *Rev Educ Res* 82(3):300–329. <https://doi.org/10.3102/0034654312457206>
- Fussy DS, Iddy H, Amani J, Mkimibili ST (2023) Girls' participation in science education: structural limitations and sustainable alternatives. *Int J Sci Educ* 45(14):1141–1161. <https://doi.org/10.1080/09500693.2023.2188571>
- Gebhard U, Höttecke D, Rehm M (2017) *Pädagogik der naturwissenschaften*. In: Ein studienbuch [Science pedagogy: a study handbook]. Springer Fachmedien
- Gee KA, Wong KK (2012) A cross national examination of inquiry and its relationship to student performance in science: evidence from the Program for International Student Assessment (PISA) 2006. *Int J Educ Res* 53:303–318. <https://doi.org/10.1016/J.IJER.2012.04.004>
- George D, Mallery P (2003) *SPSS for Windows step by step: a simple guide and reference*. Routledge
- Hanley P, Wilson H, Holligan B, Elliott L (2020) Thinking, doing, talking science: the effect on attainment and attitudes of a professional development programme to provide cognitively challenging primary science lessons. *Int J Sci Educ* 42(15):2554–2573. <https://doi.org/10.1080/09500693.2020.1821931>
- Hattie J (2008) *Visible learning: a synthesis of over 800 meta-analyses relating to achievement*. Routledge
- Jansen M (2015) *Akademische selbstkonzepte in den naturwissenschaftlichen fächern: ausdifferenzierung, Geschlechtsunterschiede und effekte dimensionaler fergleiche* [Academic self-concept in scientific subjects: differentiation between gender and effects of dimensional comparisons]. Doctoral dissertation Humboldt-Universität zu Berlin
- Ješková Z, Lukáč S, Hančová M, Šnajder L, Guniš J, Balogová B, Kireš M (2016) Efficacy of inquiry-based learning in mathematics, physics and informatics in relation to the development of students' inquiry skills. *J Balt Sci Educ* 15(5):559–559–574. <https://www.proquest.com/scholarly-journals/efficacy-inquiry-based-learning-mathematics/docview/2343744264/se-2>
- Kalyuga S, Ayres P, Chandler P, Sweller J (2010) The expertise reversal effect. *Educ Psychol* 38(1):23–31. https://doi.org/10.1207/S15326985EP3801_4
- Kane EM (2013) Urban student motivation through inquiry-based learning. *J Educ Stud* 3(1):155–168. <https://doi.org/10.5296/JSE.V3I1.3076>
- Kanter DE, Konstantopoulos S (2010) The impact of a project-based science curriculum on minority student achievement, attitudes, and careers: the effects of teacher content and pedagogical content knowledge and inquiry-based practices. *Sci Educ* 94(5):855–887. <https://doi.org/10.1002/SCE.20391>
- Keiler LS (2018) Teachers' roles and identities in student-centered classrooms. *Int J STEM Educ* 5:34. <https://doi.org/10.1186/s40594-018-0131-6>
- Kirschner PA, Sweller J, Clark RE (2006) Why minimal guidance during instruction does not work: an analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educ Psychol* 41(2):75–86. https://doi.org/10.1207/s15326985ep4102_1
- Klahr D, Nigam M (2004) The equivalence of learning paths in early science instruction: effects of direct instruction and discovery learning. *Psychol Sci* 15(10):661–667. <https://doi.org/10.1111/J.0956-7976.2004.00737.X>
- Koyunlu Ünlü Z, Dökme İ (2022) A systematic review of 5E model in science education: proposing a skill-based STEM instructional model within the 21-st century skills. *Int J Sci Educ* 44(13):2110–2130. <https://doi.org/10.1080/09500693.2022.2114031>
- Kwarikunda D, Schiefele U, Muwonge CM, Ssenyonga J (2022) Profiles of learners based on their cognitive and metacognitive learning strategy use: occurrence and relations with gender, intrinsic motivation, and perceived autonomy support. *Humanit Soc Sci Commun* 9:337. <https://doi.org/10.1057/s41599-022-01322-1>
- Lavonen J, Laaksonen S (2009) Context of teaching and learning school science in Finland: reflections on PISA 2006 results. *J Res Sci Teach* 46(8):922–944. <https://doi.org/10.1002/TEA.20339>
- Lechová P (2014) *Prírodné látky v projektovom vyučovaní* [Natural products in project-based education]. Doctoral dissertation, Charles University
- Lederman NG, Abell S, Akerson V (2008) Students' knowledge and skills with inquiry. In: Abrams E, Southerland SA, Silva P (eds) *Inquiry in the classroom: realities and opportunities*. Information Age Publishing, pp. 3–38
- Lei H, Chen C, Luo L (2024) The examination of the relationship between learning motivation and learning effectiveness: a mediation model of learning engagement. *Humanit Soc Sci Commun* 11:137. <https://doi.org/10.1057/s41599-024-02666-6>
- Madden KR (2011) The use of inquiry-based instruction to increase motivation and academic success in a high school biology classroom. Doctoral dissertation, Montana State University
- Markland D, Hardy L (1997) On the factorial and construct validity of the intrinsic motivation inventory: conceptual and operational concerns. *Res Q Exerc Sport* 68(1):20–32. <https://doi.org/10.1080/02701367.1997.10608863>
- Mayer RE (2004) Should there be a three-strikes rule against pure discovery learning? *Am Psychol* 59(1):14–19. <https://doi.org/10.1037/0003-066X.59.1.14>
- McAuley ED, Duncan T, Tammen VV (1989) Psychometric properties of the intrinsic motivation inventory in a competitive sport setting: a confirmatory factor analysis. *Res Q Exerc Sport* 60(1):48–58. <https://doi.org/10.1080/02701367.1989.10607413>
- McConney A, Oliver MC, Woods-McConney A, Schibeci R, Maor D (2014) Inquiry, engagement, and literacy in science: a retrospective, cross-national analysis using PISA 2006. *Sci Educ* 98(6):963–980. <https://doi.org/10.1002/SCE.21135>

- McOwan P (2016) Final report summary. Teaching Enquiry with Mysteries Incorporated (TEMI), London
- Minner DD, Levy AJ, Century J (2010) Inquiry-based science instruction-what is it and does it matter? Results from a research synthesis years 1984 to 2002. *J Res Sci Teach* 47(4):474–496. <https://doi.org/10.1002/TEA.20347>
- Nguyen KA, Borrego M, Finelli CJ, DeMonbrun M, Crockett C, Tharayil S, Shekhar P, Waters C, Rosenberg R (2021) Instructor strategies to aid implementation of active learning: a systematic literature review. *Int J STEM Educ* 8:9. <https://doi.org/10.1186/s40594-021-00270-7>
- Nunnally JC (1978) An overview of psychological measurement. In: Wolman BB (ed) Clinical diagnosis of mental disorders. Springer US, pp. 97–146
- OECD (2014) PISA 2012 results: what students know and can do-student performance in mathematics, reading and science. OECD Publishing
- Ortiz-Martínez G, Vázquez-Villegas P, Ruiz-Cantisani MI, Delgado-Fabián M, Conejo-Márquez DA, Membrillo-Henández J (2023) Analysis of the retention of women in higher education STEM programs. *Humanit Soc Sci Commun* 10:101. <https://doi.org/10.1057/s41599-023-01588-z>
- Palmer DH (2009) Student interest generated during an inquiry skills lesson. *J Res Sci Teach* 46(2):147–165. <https://doi.org/10.1002/TEA.20263>
- Patrick H, Mantzicopoulos P, Samarapungavan A (2009) Motivation for learning science in kindergarten: is there a gender gap and does integrated inquiry and literacy instruction make a difference. *J Res Sci Teach* 46(2):166–191. <https://doi.org/10.1002/TEA.20276>
- Pintrich PR, Smith DAF, García T, McKeachie WJ (1991) A manual for the use of the motivated strategies for learning questionnaire (MSLQ). National Center for Research to Improve Postsecondary Teaching and Learning
- Potvin P, Hasni A (2014) Analysis of the decline in interest towards school science and technology from grades 5 through 11. *J Sci Educ Technol* 23:784–802. <https://doi.org/10.1007/s10956-014-9512-x>
- Radvanová S (2017) Efektivita vybraných vzdělávacích postupů ve výuce biologie [Efficiency of selected educational practices in biology teaching]. Doctoral dissertation, Univerzita Karlova
- Rotgans JI, Schmidt HG (2010) The motivated strategies for learning questionnaire: a measure for students' general motivational beliefs and learning strategies?. *Asia-Pac Educ Res* 19(2):357–369. <https://ejournals.ph/article.php?id=4022>
- Ryan RM (1982) Control and information in the intrapersonal sphere: an extension of cognitive evaluation theory. *J Pers Soc Psychol* 43(3):450–461. <https://doi.org/10.1037/0022-3514.43.3.450>
- Sadeh I, Zion M (2009) The development of dynamic inquiry performances within an open inquiry setting: a comparison to guided inquiry setting. *J Res Sci Teach* 46(10):1137–1160. <https://doi.org/10.1002/TEA.20310>
- Salta K, Tzougraki C (2004) Attitudes toward chemistry among 11th-grade students in high schools in Greece. *Sci Educ* 88(4):535–547. <https://doi.org/10.1002/sce.10134>
- Sandika B, Fitrihidajati H (2018) Improving creative thinking skills and scientific attitude through inquiry-based learning in basic biology lecture toward students of biology education. *J Pendidik Biol Indones* 4(1):23–28. <https://doi.org/10.22219/jpbi.v4i1.5326>
- Schmid S, Bogner FX (2015) Does inquiry-learning support long-term retention of knowledge? *Int J Learn Teach Educ Res* 10(4):51–70. <http://www.ijlter.org/index.php/ijlter/article/view/289>
- Secker vonC (2010) Effects of inquiry-based teacher practices on science excellence and equity. *J Educ Res* 95(3):151–160. <https://doi.org/10.1080/00220670209596585>
- Šmejkal P, Skoršepa M, Stratilová Urválková E (2018) Evaluation and comparison of newly designed IBSE oriented MBL activities and of work with MBL systems by Slovak and Czech students. In: Ciesla P, Michniewska A (eds) Experiments in teaching and learning natural sciences. Pedagogical University of Krakow
- Stolk JD, Gross MD, Zastavker YV (2021) Motivation, pedagogy, and gender: examining the multifaceted and dynamic situational responses of women and men in college STEM courses. *Int J STEM Educ* 8:35. <https://doi.org/10.1186/s40594-021-00283-2>
- Tuan HL, Chin CC, Tsai CC, Cheng SF (2005) Investigating the effectiveness of inquiry instruction on the motivation of different learning styles students. *Int J Sci Math Educ* 3(4):541–566. <https://doi.org/10.1007/S10763-004-6827-8>
- Uitto A, Juuti K, Lavonen J, Meisalo V (2006) Students' interest in biology and their out-of-school experiences. *J Biol Educ* 40(3):124–129. <https://doi.org/10.1080/00219266.2006.9656029>
- Vadasová R, Černohlávková K, Hošková M, Pouchová M, Andresková J, Křestanová V (2013) Biologie-Výukové materiály pro 2. stupeň ZŠ (1. vydání) [Biology-teaching tasks for high schools, 1st edn]. Conatex-didactic
- Vília P, Candeias AA (2020) Attitude towards the discipline of physics-chemistry and school achievement: revisiting factor structure to assess gender differences in Portuguese high-school students. *Int J Sci Educ* 42(1):133–150. <https://doi.org/10.1080/09500693.2019.1706012>
- Vlckova J, Kubiak M, Usak M (2019) The perception of biology by Czech lower Secondary school students. *Eurasia J Math Sci Technol Educ* 15(5):em1714. <https://doi.org/10.29333/ejmste/105277>
- Volmutová I, Kuchařová E, Mokřejšová O, Müllerová L, Pražienka M, Strnadová H, Vondráčková K (2013) Chemie-Výukové materiály pro 2. stupeň ZŠ (1. vydání) [Chemistry-teaching tasks for high schools, 1st edn]. Conatex-didactic
- Wegner C, Schmiedebach M (2020) Interest in biology: grade-dependent differences and benefits of participating in out-of-school interventions. *Int J Res Method Educ* 6(3):427–434. <https://doi.org/10.46328/ijres.v6i3.1051>

Author contributions

Both authors contributed to the final version of the paper equally. The support of the project COOPERATIO in the field of Subject Specific Education Research awarded by Charles University is gratefully acknowledged.

Competing interests

The authors declare no competing interests.

Ethical approval

This study meets the Charles University guidelines for research on humans, which are in agreement with the ethical standards of the Czech Republic and the European Union. The pedagogical research was conducted during regular school lessons, incorporating new teaching methods, in full compliance with Czech educational laws, and the quality of the teaching process was not disrupted. The Institutional Review Board (IRB) of Charles University, Faculty of Science, approved the study “The impact of long-term inquiry-based science education on the students' motivation and knowledge acquisition in Biology and Chemistry” on 25 June 2019 under approval number 2019/31. The approval specifically covers pedagogical research involving human participants in regular educational settings, ensuring compliance with all relevant ethical standards and guidelines, including the Declaration of Helsinki. The research commenced in September 2019, and at that time, no additional ethical requirements or approvals were mandated by Czech legislation or institutional policies. The scope of the approval thus includes the implementation of innovative teaching methods and their assessment in regular classroom conditions.

Informed consent

At the beginning of this research, oral informed consent was obtained from the headmaster of the school on 6 September 2019. Subsequently, written informed consent was obtained from the legal representatives of all participating students during the month of September 2019 via the school's online platform. The consent was collected by the lead researcher in cooperation with classroom teachers and the headmaster. Legal representatives were provided with clear and detailed information, including (1) the purpose and objectives of the pedagogical research, (2) the procedures involved and the methods being introduced in the classroom, (3) how the anonymity of their children would be ensured through the use of unique, anonymized codes, (4) how the collected data would be securely stored and processed, and (5) that participation in the study carried no associated risks for the students. The consent specifically covered participation in the study, collection of anonymized data, and consent to publish the results in an aggregated and anonymized form. No incentives or payments were provided for participation in this study. Additionally, to maintain confidentiality, the name of the school is not mentioned in the paper.

Additional information

Correspondence and requests for materials should be addressed to Petr Distler.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons

Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025