How to conduct inquiry: Planning skills of lower secondary school students

Adam Nejedlý[1,*], Karel Vojíř[1]

[1] Faculty of Education, Charles University, Magdalény Rettigové 4, 116 39 Praha 1, Czech Republic; adam.nejedly@pedf.cuni.cz

The inquiry-based approach is the fundamental way of acquiring new knowledge in science. It is therefore essential to incorporate it into science education. A key principle is bridging the gap between asking questions and final answers, which requires an appropriate inquiry procedure. This study examined the planning skills of lower secondary school students when solving biology tasks based on an inquiry-based approach. Sets of solutions to six tasks from 51 Czech students were analysed. Closed and open coding based on research design aspects derived from the EDAT framework was used. The results show that most students recognized the need for an empirical approach and frequently identified observed variables and instruments. However, aspects of the inquiry process, such as controlling variables, replication, data analysis, and reducing uncertainty, were rarely included or were omitted. Differences were also found between tasks with descriptive and causal problems and across task contexts. The results suggest that inquiry-based learning needs more explicit support for the development of planning skills, with a focus on the complexity of research procedures. It also appears necessary to address these skills across various thematic units.

Key words: inquiry-based learning, planning skills, biology education.

Received 3/2026
Revised 5/2026
Accepted 6/2026

1 Introduction

Climate change, biodiversity loss, the spread of misinformation, and rapid technological advancement are just some of the complex global challenges facing society today (OECD, 2025; World Economic Forum, 2026). Solving these problems requires the active participation of citizens who are able to understand science, technology, and mathematics, evaluate them critically, and apply them in both individual and collective decision-making (Mocanu et al., 2025). In this context, science literacy is considered a key outcome of science education (Elhai, 2023; OECD, 2023b).

Empirical data suggest that achieving these goals remains challenging (Kranz et al., 2023). The PISA 2022 results show that a significant proportion of students in European Union countries do not meet the basic level of science literacy; in the Czech Republic, more than one-fifth of 15-year-old students fall below this level (OECD, 2023a). These findings indicate that some students have limited ability to apply science knowledge to solve real-world problems and to critically evaluate scientific information and distinguish warranted trust in science from either uncritical acceptance or sceptical rejection of expert knowledge (Baltikian et al., 2024; Osborne & Allchin, 2024; Teig, 2024).

Research has repeatedly highlighted the difficulties students face when solving tasks that require a higher level of inquiry-related skills, particularly in identifying and operationalizing variables (Čipková et al., 2025), formulating research questions (Nejedlý & Vojíř, 2022) and hypotheses (Ješková et al., 2022) or supporting conclusions with relevant evidence (Nejedlý & Vojíř, 2025).

Planning an investigation is commonly regarded as one of the core components of scientific inquiry and scientific literacy. In assessment frameworks such as PISA, planning, interpreting evidence, and evaluating scientific procedures are conceptualised as related dimensions of students' scientific competence. Nevertheless, the relationship between these dimensions should not be treated as self-evident. It requires empirical examination, particularly in the context of inquiry-based tasks implemented in real educational settings. For this reason, a detailed analysis of students' planning skills can provide important evidence about how lower secondary school students approach the methodological process of answering a research question. Such analysis may also help clarify the extent to which students' planning of an investigation supports, or is associated with, their later interpretation and evaluation of evidence. The aim of the research was to determine the skills of lower secondary school students in planning the process of inquiry for solving biology tasks. This aim was specified using the following research questions:

1. What aspects do students include when planning a methodological approach to answering a research question while working on inquiry-based tasks in biology?
2. Do students include different aspects when planning a methodological approach to answering descriptive and causal research questions while working on inquiry-based tasks in biology?
3. What are the typical student solutions when planning a methodological procedure for answering research questions in model inquiry-based tasks in biology?

2 Theoretical background

2.1 Science literacy as a key outcome of science education

According to the OECD (2023b) conceptual framework, science literacy encompasses the ability to (a) explain natural phenomena scientifically, (b) design and evaluate research procedures and critically interpret data and evidence, and (c) locate, assess, and use scientific information in decision-making and action (OECD, 2023b). Science literacy thus goes beyond the acquisition of factual knowledge and includes an understanding of the scientific process, the nature of science, the role of evidence and uncertainty, distinguishing between correlation and causation, and applying knowledge in new contexts (NASEM, 2016; Osborne & Allchin, 2024). Recent work further emphasises that science education should prepare students not only to evaluate evidence directly, but also to recognise the limits of their own expertise and to make informed judgements about the credibility of scientific sources (Osborne & Allchin, 2024). Osborne and Allchin (2024) argue that citizens are inevitably epistemically dependent on scientific experts and therefore need to develop informed epistemic trust, epistemic humility, and the competencies of a "competent outsider": the ability to assess who is legitimately speaking for science, whether a relevant expert consensus exists, and how scientific claims have been produced, reviewed, communicated, and potentially distorted in public contexts.

2.2 Skills in the inquiry process

The emphasis is on the limitations in students' skills connected with inquiry, which constitute a key component of science literacy and encompass a set of cognitive and metacognitive processes associated with conducting and interpreting scientific investigations (National Research Council, 2012; Osborne & Dillon, 2008). In the literature, skills connected with inquiry are often conceptualised as a set of interrelated processes that encompass the various stages of the inquiry cycle (Pedaste et al., 2015). These skills can be further divided into planning skills, analytical skills, and interpretive skills (Fugarasti et al., 2019; Pedaste et al., 2021). These components are interrelated, and their development depends on both a sufficient knowledge base and opportunities for active participation in research activities (Seeratan et al., 2020). In this context, planning skills deserve special attention, as they represent a cognitively demanding yet crucial stage of the inquiry process (Pedaste et al., 2015). Research planning requires the integration of subject-matter expertise with an understanding of the principles of scientific methodology and the ability to anticipate the potential outcomes and limitations of the proposed approach (Crujeiras-Pérez & Jiménez-Aleixandre, 2017; Strat et al., 2024). Inadequate planning skills can lead to methodological errors that negatively affect both the validity of the data collected and its subsequent interpretation (Bewersdorff et al., 2023; Kranz et al., 2023). In the study by Pedaste et al. (2021), planning skills are understood in a narrower sense as skills related to the design of a research procedure—that is, the transition from the formulation of a research problem to the empirical implementation of the investigation. This approach corresponds to the operationalization of planning skills in a concept that distinguishes several key categories of planning, such as the formulation of a research question or hypothesis, the identification and control of variables, the design of data collection methods and procedures, and the anticipation of possible experimental outcomes (Pedaste et al., 2021). Focusing on this phase allows for a more detailed analysis of how students design the structure of their research and what typical difficulties arise in this part of the inquiry process (Pedaste et al., 2021). Within inquiry cycle models (e.g., orientation – conceptualization – investigation – conclusion – discussion), planning serves as a key transition between problem formulation and the empirical conduct of research, and thus significantly determines the quality of the entire process (Pedaste et al., 2015). This model of sequential phases is a certain didactic simplification (Windschitl et al., 2008); in reality, the development of scientific knowledge is more iterative and pluralistic, involving ongoing evaluation and interpretation of the results obtained (see Harlen, 2021). However, it can provide a foundation for thinking about scientific inquiry as a response to a specific problem, which is a key element for further, more complex reflections on scientific epistemology.

2.3 Planning skills development

Despite the acknowledged importance of developing inquiry-related skills, many educational contexts continue to be dominated by transmissive teaching approaches focused on the reproduction of factual knowledge (Karlsen et al., 2025). Inquiry activities are often carried out on a limited basis or without sufficient methodological support (Strat et al., 2024). Furthermore, secondary analyses of international surveys suggest that the frequency of Inquiry-based learning activities alone does not necessarily lead to improved student performance in science literacy (cf. Jerrim et al., 2022, Oliver et al., 2021, Sjøberg,

2018). This discrepancy highlights the importance of the quality of implementation and the need for targeted scaffolding for the development of specific components of inquiry-related skills, rather than simply increasing the number of inquiry activities (Kang, 2022; Van Uum et al., 2017). As Riga et al. (2017) note, simply using the inquiry cycle may not be sufficient for the effective implementation of inquiry-based science education. One promising approach is the systematic use of inquiry-based learning tasks explicitly designed to develop planning skills (Schwichow et al., 2022). Tasks designed in this way should be based on a meaningful, context-based problem; they should require the formulation of a research question or hypothesis, the identification and operationalization of variables, and the design of a methodological approach (Nejedlý & Vojíř, 2024; Volkmann & Abell, 2003). It appears necessary to create a framework that will enable students to gradually take responsibility for the individual steps of the planning process and reflect on the limitations of the proposed research (Lombardi et al., 2018). Research also suggests that the impact of inquiry-based activities on planning skills represents a significant, but yet under-explored, aspect of inquiry-based education (Chen & Chen, 2025). The effectiveness of this approach depends not only on designing tasks considering the various stages of the inquiry process, but also on a deeper analysis of how students solve these tasks (Lazonder & Harmsen, 2016). It is becoming clear that, in order to understand the reasons behind students' success or failure in solving inquiry-based tasks, it is essential to focus on evaluating student performance at each stage of the inquiry process and to systematically identify the typical mistakes students make (Pedaste et al., 2021). This approach makes it possible to assess the level of students' initial planning skills, thereby providing a basis for targeted interventions that support the development of their scientific thinking.

3 Methodology

To answer the research questions, a mixed-methods study was conducted in 2024 at seven Czech elementary schools. The study combined quantitative analysis of the frequency with which students included selected aspects of inquiry design in their written plans with qualitative analysis of typical patterns in students' proposed methodological procedures.

3.1 Inquiry-based biology tasks

The study involved six inquiry-based tasks designed to encourage students to develop their own methodological approach to conducting inquiry. The biology tasks were piloted as part of the study by Nejedlý & Vojíř (2024).

Each inquiry task was designed according to a standard format. Each task always included:

1. Thematic motivational text;
2. instructions for students and the assigned research question;
3. an open space for students to propose an inquiry methodology;
4. opportunity for students to suggest the tools and materials to be used.

A uniform task structure was chosen to minimize the cognitive load associated with understanding the assignment, so that students could focus their attention on the planning process itself. Students worked with a given research question, and their task was to propose a methodological approach to investigating it, including research materials and tools. The assignment intentionally did not include any further specifics regarding the design of the methodological procedure.

The tasks represented two basic types of scientific research based on the research question: a) causal and b) descriptive. Each of these types was represented by three tasks.

Tasks containing a descriptive type of research problem:

- Observation of a slug – What are the length, width, and surface texture of a Spanish Slug?
- Observation of blooms – What are the colour, size, and number of petals Snowdrops and Spring Snowflakes blooms?
- Observation of fingers – What length are the fingers on the left and right hands of 14-years-old students?

These research questions were intended to guide students in developing a systematic approach to observing biological specimens and in designing an appropriate method for collecting and recording data. Students were asked to propose a comprehensive observation procedure that would allow them to answer the given question empirically.

Tasks involving a causal type of research problem:

- Experiment with yeast – How does temperature of the environment affect yeast activity?
- Experiment with a runner – How does running 800 meters affect a runner's body temperature?

- Experiment with tomatoes – How does enriching the soil with horse manure affect the weight of tomato fruit?

These research questions were intended to guide students in considering how to conduct systematic experiments in relation to the variables contained in the research question and in designing an appropriate method for collecting and recording data. Students were asked to propose a comprehensive experimental procedure that would enable them to answer the given question.

3.2 Assignment procedure and research sample

As part of the study, 15-year-old students gradually completed all six inquiry-based tasks during their regular biology classes. A total of 207 students initially participated in the study. Of these, 156 students were excluded from the analysis because they did not complete all six tasks. The final analytical sample therefore consisted of 51 students, each of whom completed all six inquiry-based tasks. Because every student solved each of the six tasks, the data have a repeated-measures structure: the same students contributed responses across all task contexts and both types of research problems. Before the study began, all teachers were trained in how to implement the inquiry-based tasks and received detailed methodological guidelines to ensure consistency in the study's implementation and adherence to the principle of not providing students with feedback on the tasks. To preserve mental capacity, students worked on only one task per class period. Biology classes were scheduled back-to-back to minimize external influences, and teachers did not provide direct feedback between tasks. Students worked on the tasks individually. The tasks were presented to them in a fixed order: Observation of a slug – Experiment with yeast – Observation of blooms – Experiment with a runner – Observation of fingers – Experiment with tomatoes. This order was chosen deliberately so that tasks involving descriptive and causal problems would alternate.

The study was conducted at seven public elementary schools. All schools followed the national curriculum, and science education was delivered in the format typical in the Czech Republic, namely as separate subjects of biology, chemistry, and physics. It was confirmed that the students at these schools had no prior educational experience with solving problems of this thematic nature and teachers at these schools do not use inquiry-based education in their lessons. Biology classes at all participating schools were taught by fully qualified teachers. The schools were selected with regard to sociodemographic diversity and the potential influences of the educational environment (Gay et al., 2019). The breakdown included the regions of Bohemia (5) and Moravia (2), specifically the Ústí nad Labem region (1), the Pardubice region (1), the Prague region (3), the South Moravia region (1), and the Central Bohemia region (1), as well as school sizes: up to 350 students (2), up to 700 students (3), and up to 1,000 students (2). The analysis included solutions from students who completed all 6 tasks. In total, solution sets from 51 students were further analysed.

3.3 Analysis procedure

The analysis of the students' solutions focused on evaluating the aspects of the inquiry procedure plans that the students proposed to answer the assigned research questions. Each solution was analysed using a tool based on categories formulated according to the EDAT framework (see Sirum & Humburg, 2011). This framework has already been tested in studies focused on assessing experimental design skills (e.g. Shanks et al., 2017). The aspects under consideration are also consistent with other definitions of research plan aspects in inquiry-based learning (see Fugarasti et al., 2019; McComas, 2014; National Research Council, 2013). The aspects evaluated in the students' solutions are listed in Table 1. Since it appears that students often find it difficult to identify the nature of the research problem when dealing with research tasks (Nejedlý & Vojíř, 2025), all aspects were evaluated for every task. Even in the case of tasks with a descriptive type of research problem, it was monitored whether students included irrelevant aspects, such as manipulated variables, in their solutions.

In the first phase of the analysis, closed coding was used. For each aspect, it was recorded whether it was **included** or **not included** in the student solution under review. Based on this coding, it was possible to determine the proportion of solutions in which students had taken the given aspect into account. To increase reliability, the coding was performed by two researchers. In the second phase, the analysis was expanded to include open coding focused on specific ways in which students addressed the assessed aspects of the methodological plan leading to the answer to the research question in individual tasks. To identify typical solution patterns, open codes were assigned to individual student solutions for each assessed aspect. Coding proceeded iteratively in a cyclical sequence, with solutions repeatedly revised and refined. The coding process continued until theoretical saturation was reached, at which point no new categories of meaning were identified. This process was carried out by two researchers, who continuously compared their coding and reached consensus in cases of disagreement.

Table 1: Aspects of inquiry design and characteristics of the evaluation

Aspects of Inquiry Design	*Characteristics of the evaluation*
inquiry approach	Whether the student proposed a procedure based on the systematic collection of empirical data through observation or experimentation. E.g. the student mentioned measuring a runner's temperature or observing differences in spring snowflake blooms.
manipulated variables	Whether the student identified a variable whose value they will intentionally change during the experiment. E.g., changing the ambient temperature when studying yeast activity, or adding manure to the soil when growing tomatoes.
observed variables	Whether the student has identified the variables or characteristics whose values will be recorded during the study. E.g., the length and width of a slug's body, the number of petals on blooms, or a runner's body temperature.
method of variable observation	Whether the student specified a specific method for measuring or observing the variables under study. E.g., measuring finger length with a ruler in centimetres, or measuring temperature with a thermometer before, during, and after a run.
constant conditions	Whether the student specified conditions that should be kept constant during the experiment. E.g., the same amount of yeast in the test tubes, the same amount of water, or the same measurement intervals.
sample characteristics	Whether the student described the basic characteristics of the objects or subjects on which the research will be conducted. E.g., the Spanish slug, snowdrop and snowflake blooms, 14-year-old students.
sample selection	Whether the student specified the method used to select the objects or subjects for the study. E.g., undamaged blooms, volunteers among 14-year-old students, or several specimens of slugs.
Repetition	Whether the student suggested repeating the measurement, observation, or experiment. E.g., The measurement was performed on five slugs, or the temperature was measured three times after the run.
data recording method	Whether the student described a method for systematically recording the collected data. E.g., entering values into a table or a graph.
tools and materials	Whether the student listed the instruments or supplies needed to conduct the research. E.g., ruler, magnifying glass, thermometer, stopwatch, test tubes, scale.
data analysis	Whether the student described how the collected data will be analysed or compared. For example, this could involve comparing values, calculating averages, identifying trends, or drawing conclusions based on the data.
uncertainty and error reduction	Whether the student considered potential sources of error and proposed methods to improve the reliability of the results. E.g., repeated measurements, a control sample, or verification of the results on multiple specimens.

3.4 Data processing and analysis

The data were descriptively analysed, standardized, and prepared for further analysis in MS Excel. In the analysis of data obtained through closed coding for the selected aspects, the percentage frequency of the inclusion of individual aspects in the entire set of solutions was first calculated based on the closed dichotomous coding. To further analyse the inclusion of procedural aspects according to the type of research problem, the occurrence of each aspect in tasks requiring experimental and observational approaches was calculated as a simple sum.

IBM SPSS Statistics software was used for the statistical analysis. As the Shapiro-Wilk test indicated significant deviations from normality for all items ($p < 0.01$), nonparametric tests were used in the analysis. Differences in the inclusion of individual aspects between students' solutions to tasks involving descriptive and causal problems were tested using the Wilcoxon signed-rank test. The use of the Wilcoxon signed-rank test and Cochran's Q test reflected the repeated-measures design of the study, as the same students provided responses to all six tasks. The Rosenthal correlation coefficient r was used to assess effect size. Differences in the inclusion of individual aspects in students' solutions across tasks were tested using Cochran's Q test, which allows for the comparison of multiple related binary values. In this study, the test examined whether a given aspect was included in the solutions to several tasks. Effect size was expressed using η^2. Effect size coefficients were interpreted according to Cohen (1988). Statistical significance was assessed at a p-value of < 0.05.

For data obtained through open coding, the most frequent combinations of student solutions were quantified. This identified typical patterns of methodological approaches used to answer the research questions in individual tasks. These solutions were described qualitatively, and representative examples were selected for them.

4 Results

4.1 Differences between task types in specific aspects

The vast majority of students chose an inquiry-based approach in their inquiry plans to answer the assigned research questions. However, when considering the individual aspects examined, significant differences were found in the inquiry plans; see Table 2. In the vast majority of cases, students included the tools and/or materials they plan to use for their inquiry and considered the variables whose values they will monitor.

However, students specified the specific method for tracking these variables or provided a detailed description of the research sample on which they would conduct their study in only about half of their solutions. Other aspects were rarely considered. Students included methods for evaluating the collected data only very sporadically, and after describing the setup of the observation or experiment, they typically provided general statements such as: "I will write everything down and answer the research question." In their plans, students practically did not address solutions to potential limitations of the research or procedures for reducing errors. This aspect was addressed only in a handful of cases where students chose to use a control sample, repeated control measurements, verification of the procedure's functionality on a different sample, measurements using different devices and comparison of findings, or a review by another researcher.

Table 2: Consideration of individual aspects in student inquiry plans and statistical significance of differences between task solutions. Statistically significant p-values are marked in bold. Rosenthal r and eta squared values indicating medium effect sizes are marked in italics, and values indicating large effect sizes are marked in bold.

	proportions in solutions			difference according research problem type			difference among particular tasks		
	All tasks	Descriptive problems	Causal problems	Z	p	r	Q	p	η^2
inquiry approach	99%	98%	99%	−1.41	0.16	0.20	12.22	**0.03**	0.05
manipulated variables	44%	1%	88%	−6.51	< **0.01**	**0.91**	212.49	< **0.01**	**0.83**
observed variables	79%	84%	75%	−2.45	**0.01**	*0.34*	41.51	< **0.01**	**0.16**
method of variable observation	53%	58%	47%	−2.44	**0.02**	*0.34*	23.58	< **0.01**	*0.09*
constant conditions	14%	8%	20%	−3.48	< **0.01**	*0.49*	32.57	< **0.01**	*0.13*
sample characteristics	53%	53%	54%	−0.17	0.87	0.02	35.55	< **0.01**	*0.14*
sample selection	14%	27%	2%	−4.37	< **0.01**	**0.61**	62.86	< **0.01**	**0.25**
Repetition	26%	40%	12%	−4.23	< **0.01**	**0.59**	55.00	< **0.01**	**0.22**
data recording method	24%	29%	19%	−2.58	**0.01**	*0.36*	9.52	0.09	0.04
tools and materials	98%	97%	98%	−0.58	0.56	0.08	12.04	**0.03**	0.05
data analysis	4%	1%	7%	−2.71	**0.01**	*0.38*	1.95	0.06	0.04
uncertainty and error reduction	3%	2%	4%	−1.00	0.32	0.14	6.95	0.22	0.03

Several statistically significant differences were found in how students considered individual aspects in tasks requiring an experimental approach and an observational approach without manipulating variables. Statistically significant differences were absent only for aspects that students considered either almost universally or almost never; see Table 2. Furthermore, no statistically significant difference was found in the case of the characteristics of the research sample, which slightly exceeded half of the solutions in both types of approaches. The most pronounced difference was found in the inclusion of manipulated variables. This difference stems from the nature of the research questions posed; in the case of descriptive research questions, students failed to correctly mention variable manipulation in the vast majority of cases. Conversely, in most plans for answering research questions involving a causal type of research problem, students correctly reported variable manipulation. Other significant differences of large effect size were found regarding research sample selection and replication of the procedure. In both cases, students more frequently mentioned these aspects in procedures related to descriptive problems. Here, students more frequently cited the need to select a representative sample from a larger population that, for example, would not be biased. They mentioned selecting a larger number of subjects for observation. In contrast, when planning an experiment to address a causal problem, sample selection was practically non-existent, and only in a small number of cases did students mention conducting experiments on a larger number of samples or repeating the procedure.

4.2 Differences between individual tasks in specific aspects

A more detailed analysis further revealed differences between individual tasks. In the case of tasks with a descriptive problem type, the frequency of mentioning the selection of the research sample varied considerably (see Fig. 1). This aspect was very rarely mentioned in the observation of the slug (8%), while it was mentioned more frequently in the observation of blooms (35%) and fingers (37%). The greatest differences in these tasks were found in constant conditions. These were most frequently considered in the observation of fingers (20%), while in other tasks focused on description, this aspect did not appear at all or only rarely (4%).

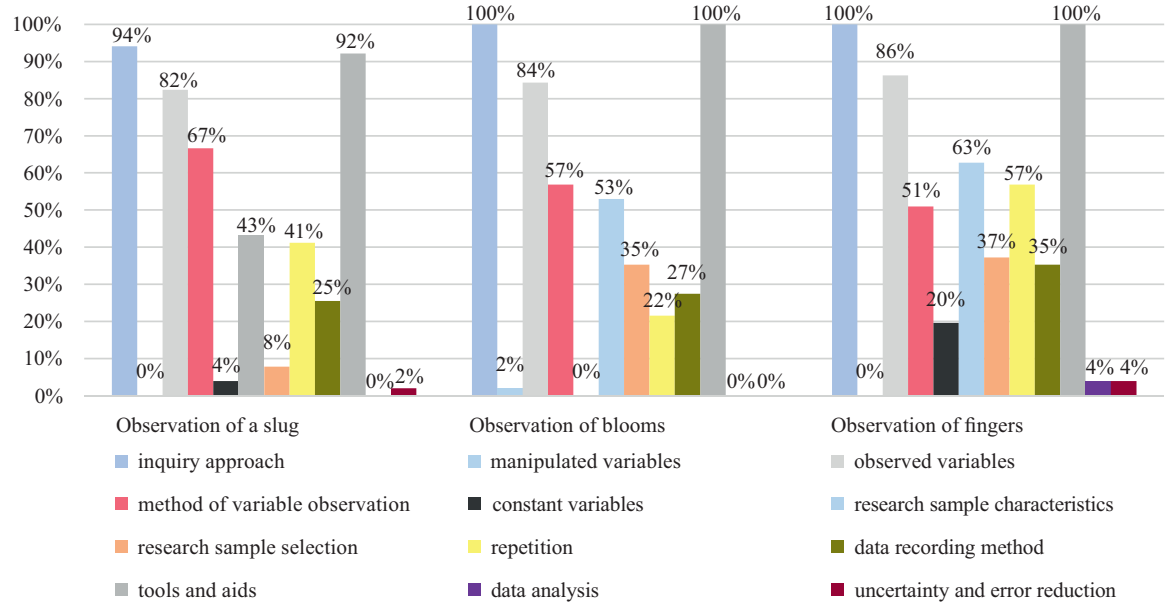

Fig. 1: The proportion of solutions containing the assessed aspects in tasks with a descriptive problem type

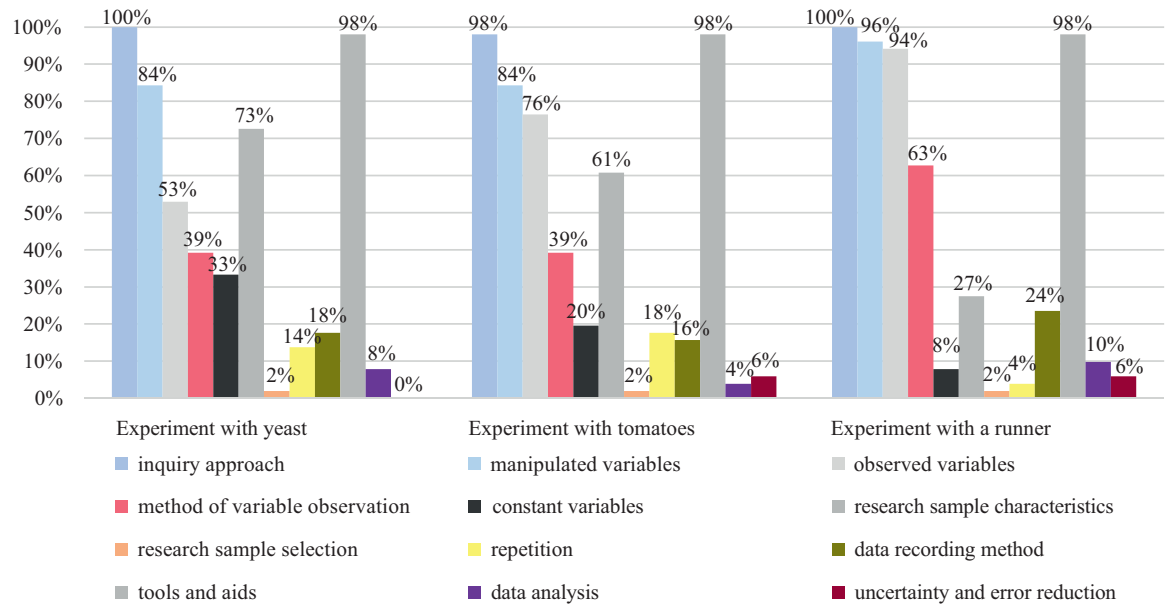

Fig. 2: The proportion of solutions containing the assessed aspects in tasks with a causal problem type

The inclusion of the constant conditions aspect was strongly influenced by the specific topic of the problem being addressed in tasks focused on causal relationships (see Fig. 2). While it was considered in one-third of the cases in the yeast experiment, it was considered in only 8% of the cases in the runner experiment. Furthermore, significant differences were found in these tasks regarding the reporting of observed variables. While the observed variable was reported in 94% of solutions for the runner experiment, it was 76% for the tomato experiment and only 53% for the yeast experiment. Conversely, in the runner task, students more frequently specified the method of monitoring the variable (63%) compared to the

yeast and tomato experiments (both 39%). A significant divergence in experimental design plans was also found in the characterization of the research sample. The highest representation of this aspect was recorded in the yeast experiment (73%), while it was lowest in the runner experiment (27%). A similar trend of low representation in the runner experiment was also found for the aspect of repetition, which was included in only 4% of the solutions.

4.3 Typical solutions based on the considered aspects

4.3.1 Observation of a slug

In the case of the task focused on observing the Spanish slug, 17 types of student solutions were identified based on the aspects of the inquiry design that were considered. The most common type of solution (23% of student solutions) was a simple proposal for an observation procedure focused on a single Spanish slug.

> *After the rain, we'll find a Spanish slug. We'll use a ruler to measure its length and width. We'll use a magnifying glass to examine the slug's structure. We'll record our findings. Then we'll release the slug back into the wild.*
>
> *Materials and tools: ruler, Spanish slug, magnifying glass.*

In this solution, the students adopted an inquiry-based approach and identified the key variables of interest, specifically the length, width, and surface structure of the body. They also listed the tools and aids they used, primarily a ruler for measuring dimensions and a magnifying glass for closer observation of the body's surface. The procedure also included recording the information gathered, which indicates at least a basic understanding of data documentation.

From a methodological standpoint, however, this approach remains rather limited. The students did not describe the characteristics of the research sample, as the observation involved only a single specimen. Furthermore, there was no repetition of measurements, which would have allowed for data comparison and increased the reliability of the results. The solution also did not include a detailed specification of how the variables would be monitored, such as how body length would be measured given the varying dimensions of different body parts or considering the slug's movement, nor how the observed traits would be systematically recorded or analysed. Overall, this is a simple, intuitive observation design that allows for the collection of basic descriptive data but does not provide a sufficient methodological framework for more systematic biological investigation.

The second type of common student solution, which appeared with the same frequency (23% of student solutions), was significantly more sophisticated in terms of inquiry design.

> *We'll collect Spanish slugs in a bowl. I'll prepare a mat, paper, or newspaper and some string. We'll transfer one slug onto the mat, paper, or newspaper. We'll observe its movements as it stretches and contracts. I'll take the string and measure the length and width of one slug. We record the measured values. We perform the measurements on two more specimens. We examine the structure of the slug with a magnifying glass. We record our observations in the log. I make sure not to mix up the millipedes. I place them in different bowls. We evaluate the data. We answer the research question.*
>
> *Materials and tools: ruler, spiral notebook, 3× Spanish slugs, paper, pencil, 3× bowls, mat, paper/newspaper, magnifying glass, string.*

In this case, the students defined a larger research sample and incorporated repeated measurements. They also described how individual specimens would be handled and separated and specified the use of string to measure body dimensions despite the slug's movement and changing body shape. Compared with the first solution, this design also includes more explicit data recording and subsequent evaluation. However, it still lacks a more systematic approach to data analysis.

4.3.2 Observation of blooms

In the task focused on observing bloom parts, 20 types of student solutions were identified based on the aspects of the inquiry design that were considered. The most common type of solution (18%) involved identifying the key variables under observation, namely the colour, size, and number of petals in the blooms of two plant species.

> *Count the petals on the blooms and record the number. Measure the size of the blooms with a ruler and record the measurement. Identify the colour. Pay close attention to all the results and answer the question.*
>
> *Materials and tools: blooms from the plants listed, a magnifying glass, and a ruler.*

The students proposed methods for monitoring variables, such as measuring bloom size with a ruler and counting the number of petals. However, this proposal remained very general, and the students did not specify, for example, which parts they would measure. The solution also listed the tools to be used—a ruler and a magnifying glass. The students stated that they would record the results but did not specify how they would keep records. This type of solution did not include repeating the observations on a larger number of specimens or a systematic description of data analysis, such as how to compare results among different plant species. From a methodological perspective, this is a relatively functional proposal for simple observation, but with limited depth in the inquiry design.

The second most common type of solution (16%) also considered sample selection, emphasizing that the blooms should be whole and undamaged. These solutions included observing a larger number of blooms and comparing them, which suggests an understanding of how appropriate object selection can affect the quality of results.

> *I will prepare both types of plants—avoiding damaged or incomplete blooms. I will examine the blooms and first assess their colour. I will pluck one petal and measure its size with a ruler. I'll record the colour and size in my notebook. I'll count the number of petals on each plant. I'll record the results. I'll evaluate and compare the results. I'll answer the research question.*
>
> *Materials and tools: two snowdrop blooms, two snowflake blooms, a pencil, a notebook, a ruler.*

4.3.3 Observation of fingers

In the task focused on observing fingers, 27 types of student solutions were identified based on the aspects of the inquiry design that were considered. The most common type of solution (16%) involved only the basic identification of the research sample—namely, 14-year-old students—and a proposal to measure the length of their fingers using a ruler.

> *I'll bring in some 14-year-old students and ask them to help with the research. I'll measure their fingers from all sides. Done.*
>
> *Materials and tools: ruler, students.*

The students did not specify the exact method of measurement or the point from which the finger length would be measured. Their solution did not include repeated measurements, systematic data recording, or subsequent data analysis.

The second most common type of solution (14%), on the other hand, was significantly more complex.

> *I will ask the volunteers to undergo a measurement. I will use a tape measure or ruler to measure the length of their finger from the centre of the joint to the tip. I will take measurements on both the right and left hands. I will record the data in centimetres. I will repeat this process for all 14-year-olds. I will compare and analyse the data. I will answer the research question.*
>
> *Materials and tools: ruler, volunteers, paper, pencil.*

Unlike the first solution, this response specified the measurement method, defining finger length as the distance from the joint to the fingertip. It also identified a sample of 14-year-old volunteers, included repeated measurements across multiple individuals, and involved both data recording and subsequent comparison, representing a basic form of data analysis.

4.3.4 Experiment with yeast

In the task focused on the yeast experiment, 22 types of student solutions were identified based on the aspects of inquiry design that were considered. The most common type of solution (16%) was a relatively complex proposal for an experimental procedure.

> *I'll prepare a mixture of yeast, water, and sugar in two test tubes. We want the same amount of mixture in both. I'll fill the beakers with cold and warm water and grab a stopwatch. I'll label the beakers 'T' for warm water and 'S' for cold water. I place one test tube in each beaker. I record what happens at each minute. I record the readings after two minutes and after ten minutes.*
>
> *Materials and tools: sugar, 2 test tubes, water, a stopwatch, paper, a pencil, warm water.*

The students clearly identified the manipulated variable (ambient temperature). The procedure also included the method of monitoring the variable: observing changes over time and recording them at regular intervals. They also considered constant conditions, such as using the same amount of yeast, water, and sugar in each test tube, and included the necessary materials and a plan for data recording.

The second most common type of solution (14%) was significantly simpler. The students did not specify the dependent variable or how it would be measured. The solution also lacked data recording and subsequent analysis.

> *Label test tubes one through four. Place yeast and water in each test tube. Place the first and second test tubes in a cool environment and the third and fourth test tubes in a warm environment. The experiment will last ten minutes.*
>
> *Materials and tools: four test tubes.*

4.3.5 Experiment with tomatoes

In the task focused on the tomato experiment, 21 types of student solutions were identified based on the aspects of the inquiry design that were considered. However, only a single type of solution significantly outnumbered the others in this case. This most common solution (14%) included the identification of the manipulated variable—adding manure to the soil—and the observed variable, which was the weight of the tomato fruits.

> *We will plant the tomato seedlings. We will label the seedlings A through D. We will apply fertilizer to three of them. After the tomatoes have grown, we will weigh the fruits from each plant. I will record the weights, analyse the data, and answer the research question.*
>
> *Materials and tools: manure, 4 tomato seedlings, scale, pen, paper, water.*

The students also described how to measure using a scale and planned to record and analyse the data. The experimental design included a control group, since only some of the plants were fertilized. However, it did not explicitly specify constant conditions, such as equal amounts of water or light, or include replication with a larger number of plants.

4.3.6 Experiment with a runner

In the task focused on the runner experiment, 18 types of student solutions were identified based on the aspects of inquiry design that were considered. The most common type of solution (25%) was a simple experiment in which students identified the manipulated variable, namely an 800 meter run, and the observed variable, namely the runner's body temperature.

> *We will measure the runner's temperature. We will have the runner run 800 meters. Once they finish, we will measure their temperature with a thermometer. We will compare it to their original temperature before running.*
>
> *Materials and tools: thermometer, measuring tape.*

The procedure involved measuring body temperature before, during, and after the run and then comparing the results. However, the study did not include repeated measurements, the selection of a research sample, or the control of other factors that could influence body temperature.

The second most common type of solution (18%) expanded the design by including repeated temperature measurements at several intervals after the run.

> *First, we'll take the runner's temperature twice while he's at rest, then he'll run 800 meters. When he returns, we'll take his temperature right away. We'll take his temperature again after one minute, and again after two minutes. We'll compare the temperatures we've measured.*
>
> *Materials and tools: a thermometer, a place where you can run 800 meters.*

4.3.7 Summary students' inquiry proposal design

An analysis of typical solutions across individual tasks reveals a relatively consistent pattern in how students approached the design of a research procedure. In most cases, students were able to identify the basic research approach and correctly identified the variables under study and the corresponding instruments or tools needed to measure or observe them. These elements appeared in both descriptive tasks (e.g., observing a slug, blooms, or fingers) and experimental tasks (e.g., yeast, tomatoes, or a runner).

Students also frequently mentioned basic data recording, indicating a fundamental understanding of the need to document the results obtained.

In contrast, more complex aspects of inquiry design appeared significantly less frequently in their solutions. Only a portion of the students included a description of the research sample or the repetition of the procedure, which are important for increasing the reliability of the data obtained. Even less frequently did students explicitly mention constant experimental conditions or methods for controlling them. Another significant shortcoming was the absence of a more systematic data analysis plan that would allow for a structured comparison of the results obtained and the formulation of conclusions. Only in very rare cases did the solutions take measurement uncertainty into account or propose procedures for reducing measurement errors.

Overall, the results suggest that students are relatively proficient at identifying variables and the practical steps involved in data collection, but they consider methodological principles that ensure the reliability and interpretability of results significantly less often. This pattern emerged in both descriptive and causal inquiry-based tasks, suggesting that students' difficulties are related more to the more complex aspects of research planning than to the type of research question itself.

5 Discussion

The results show that, in most cases, 15-year-old Czech students correctly recognized that answering the research question requires an empirical inquiry approach; this is also confirmed by a study conducted by Teig (2024). The nearly universal choice of observation or experimentation may indicate that, within the context of these written tasks, students were able to associate the research questions with an empirical approach; this connection is also confirmed by e.g. Pedaste et al. (2015) and van Uum et al. (2016). This result is particularly evident in the distinction between descriptive and causal tasks: while students generally did not identify the manipulated variables in descriptive problems, they often correctly identified them in causal tasks. These results are consistent with the findings of a qualitative study by Arnold et al. (2014) regarding issues in students' procedural understanding. A possible cause, as noted by Schwichow et al. (2022), is insufficient metacognitive knowledge of the relationship between manipulative and observed variables. Tairab (2015) notes and confirms that students can, to some extent, distinguish between when it is appropriate to use observation and when to use experimentation; however, he also points out that guided inquiry – focused primarily on verifying and illustrating phenomena – predominates in the classroom, leading to its dominance in the learning context and limiting opportunities for authentic scientific inquiry. This finding corresponds with the fact that understanding remains largely at the level of basic procedure selection and is less reflected in its methodological sophistication (see Kranz et al., 2023).

The analyses showed that the students were relatively proficient, particularly in those aspects of the design that are directly related to specific activities, namely identifying the variables of interest and listing the instruments and tools (cf. Arnold et al., 2014). Typical solutions were often based on a brief description of measurements or observations and their recording; more complex proposals involving multiple measurements, sampling, or basic data analysis were less common. This pattern can be linked not only to the cognitive demands of individual aspects of the inquiry design (cf. Schwichow et al., 2022), but also with the nature of science education in schools, where students are often guided more toward reproducing predetermined "experiments" and demonstrations than toward independently thinking through the structure of inquiry, as confirmed by, e.g., Abrahams and Millar (2008) and Janštová and Pavlasová (2019). In other words, students have a fairly good idea of what they will do, but they rarely explicitly consider why the procedure should be designed this way and what conditions it must meet in order to lead to reliable conclusions (see Zimmerman, 2007). The acquisition of these patterns is apparently also influenced by the nature of the investigative tasks presented in the teaching materials. For example, in chemistry textbooks, laboratory activities are typically described only procedurally, as a list of mechanical steps accompanied by a corresponding list of equipment. However, aspects of methodological planning—such as linking the procedure to the research question, selecting and specifying samples, evaluating results, or eliminating errors and uncertainties—are not presented in textbooks (see Vojíř, 2021).

Even in the students' plans for methodological approaches, constant conditions, repetition of procedures, data analysis, and—in particular—dealing with uncertainty were taken into account significantly less often; this pattern is consistent with earlier research, which identifies precisely these components as problematic in the acquisition of inquiry-related skills (cf. Arnold et al., 2014; Schwichow et al., 2022). It is precisely here that shortcomings in epistemic knowledge and its application in inquiry design are likely to become apparent (see Sandoval, 2005). While identifying a variable or a tool is relatively straightforward, taking into account variability, sources of error, or how to interpret results requires a deeper understanding of the nature of science of scientific inquiry (see Zetterqvist & Bach, 2023; Zimmerman, 2007). The

very low prevalence of procedures for reducing uncertainty therefore suggests that students often view measurement as a means of obtaining a single "correct answer", rather than as an error-prone process that requires verification and repetition, which is consistent with the findings reported Kok (2022).

The differences between individual tasks further demonstrate that, in addition to the type of research problem, the context of the specific task also plays a significant role, as noted Brown (2019). For example, in the task involving the runner, students more frequently specified the method of monitoring the variable than in other experimental tasks, likely because body temperature is a familiar and well-established concept in school settings. In contrast, in the yeast task, the monitored variable was specified less frequently, even though it involved a methodologically rich experimental situation. This suggests that students find it easier to work with variables that are commonly used in school instruction and have established terminology, while they face greater difficulties when they need to independently specify a variable in a less standard context, a finding confirmed by other studies (e.g. Shaffer et al., 2016; Schäfer et al., 2024). It is therefore not merely a matter of identifying variables as such, but also of the ability to translate a specific biological phenomenon into explicit research language (see Hamnell-Pamment, 2024).

It is also interesting to note that sample selection and repetition of the procedure occurred more frequently in descriptive tasks than in causal tasks. This result suggests that students associated the need for multiple examples more with comparing objects than with verifying the reliability of experimental results, which is consistent with the findings reported Klahr and Chen (2011). When it came to blooms and fingers, they more often considered multiple instances, whereas for experimental tasks, they often settled for the idea of a single trial. Thus, in their written plans, students appeared to apply repeatability and representativeness more as task-specific considerations than as general principles of inquiry design (see Arnold et al., 2014).

5.1 Implications

Based on the research conducted, it appears that instruction should not be limited solely to incorporating inquiry tasks (cf. Lazonder & Harmsen, 2016), but should more explicitly support the individual components of inquiry planning. Although students in most cases correctly recognized that answering a research question requires an empirical approach, their proposals often remained at the level of immediate procedural steps. This suggests that the experience of an inquiry task alone may not automatically lead to the development of more methodologically sophisticated thinking (see Heindl, 2019; Jerrim et al., 2022). It therefore seems particularly important to focus instruction on those aspects of planning that students spontaneously considered only to a limited extent, particularly the control of conditions, repetition of procedures, data analysis, and dealing with uncertainty. These aspects likely require targeted scaffolding (see van Uum et al., 2017), which will help students move beyond merely describing "what they will do" toward thinking through why a given procedure is methodologically sound and under what conditions it can lead to reliable conclusions.

This need is further underscored by the importance of explicitly mapping out a inquiry plan. The results suggest that students are relatively proficient at identifying variables and tools, but they less frequently consider the relationships between them, the methods for analysing the data, or the limitations of the proposed procedure. According to Renkl (2014), it can be pedagogically beneficial to make the structure of an inquiry design accessible through model solutions, joint analysis of examples, or a comparative examination of methodologically stronger and weaker designs. Such model solutions can serve not only as a support for students but also as a support for teachers in focusing attention on the key features of different types of tasks. In this sense, they can function as scaffolding tools that structure both student learning and pedagogical decision-making regarding which aspects of planning need to be emphasized in a specific type of task (see van Uum et al., 2017). It seems that it might be useful to break down inquiry planning into explicit steps and provide targeted questions or supporting tools for each of them. Such an approach could help ensure that the methodologically and epistemologically more demanding components of inquiry design become part of students' everyday thinking gradually, as noted by, for example, Sandoval (2005) and Zimmerman (2007), rather than as implicit expectations.

The results show that the quality of students' solutions is not determined solely by their general skill level but is significantly influenced by the specific thematic context of the task. Differences between individual tasks suggest that students find it easier to work with variables and procedures that are commonly used in school instruction and linguistically established, whereas in less familiar contexts they have greater difficulty with their explicit formulation and operationalization (see Shaffer et al., 2016). It follows that the development of planning skills should not be tied to individual, isolated activities or specific subject areas, but should be systematically supported across various science topics. What is important here is not only the repetition of inquiry tasks, but also the variability of contexts, which allows students to transfer planning strategies between different types of problems and gradually abstract them

from specific content. The study's findings thus support the need to design inquiry tasks not only in terms of their thematic content but also in terms of the specific planning skills and epistemic understanding they are intended to develop in students.

5.2 Limitations of the study

When interpreting the results, several limitations of the study must be considered. The research was conducted using a set of six research-oriented biology tasks, which limits the generalizability of the findings to other areas of science and types of tasks. Differences between tasks may thus have been influenced not only by the type of research problem but also by the specific content and the students' familiarity with the given topic. Although the order was chosen deliberately to alternate descriptive and causal research problems, possible order effects cannot be ruled out. Students' responses to later tasks may have been influenced by their experience with previous tasks.

Furthermore, the study analysed written proposals for methodological procedures, not the actual conduct of the research. The results therefore reflect the students' ability to explicitly articulate a research plan, not necessarily all aspects they would consider in practical work. While the open-ended nature of the assignment, without detailed guidance, it may have led to the underestimation of certain aspects of the research design that students did not consider necessary to explicitly mention.

Another limitation is the size of the analysed dataset, which included only 51 complete sets of solutions; this may limit the generalizability of the results and suggests possible sample bias. These limitations do not diminish the significance of the findings, but they do constrain their interpretation and highlight the need for further research, particularly linking the analysis of proposals to actual research activities and verifying the results on larger and more diverse student samples.

6 Conclusion

This study contributes to our understanding of how 15-year-old lower-secondary school students design inquiry procedures when solving inquiry-based biology tasks. The results show that, in most cases, students correctly recognize that answering the given research question requires an empirical inquiry approach, and at a basic level, they can distinguish between descriptive and causal types of problems. However, it appears that their proposals remain largely at the level of immediate procedural steps. Most frequently, they included the identification of observed variables and the specification of instruments or tools, while components requiring deeper methodological and epistemic understanding—particularly the control of conditions, repetition of the procedure, data analysis, and dealing with uncertainty—appeared significantly less often. Furthermore, the differences observed between individual tasks suggest that students' performance is determined not only by the type of research question but also by the specific thematic context and the extent to which the given variables or procedures are familiar to students from school and part of their vocabulary.

These results suggest that the development of planning skills cannot be reduced to merely incorporating inquiry-based activities but rather requires targeted and explicit support focused on the structure of the research design. From a didactic perspective, it therefore seems important to pay greater attention to those aspects of planning that students do not spontaneously include, and to support their development across different types of tasks and subject areas. Further research should therefore examine how planning skills change over time, how they develop gradually through inquiry-based tasks, how they are influenced by various forms of scaffolding support, and to what extent they transfer across different science contexts.

Declaration of interest statement

No conflict of interest to declare.

Ethics statement

We confirm that the study was carried out in accordance with the ethical principles outlined in the Belmont Report and complied with all relevant ethical, legal, and human subject protection standards applicable to this type of research. Informed consent was obtained from all participating schools prior to the inclusion of students in the study. All personal data were fully anonymised, and researchers had access only to coded identifiers that did not allow for the identification of individual participants or schools. During manuscript preparation, the authors used ChatGPT solely for AI-assisted copy editing to improve readability, style, grammar, spelling, punctuation, and tone. It was not used for content creation, research ideas, data analysis, interpretation, or any substantive scholarly contribution.

Funding

This publication was supported by Institutional Support for Longterm Development of Research Organizations – Cooperatio SOC/Subject Specific Education Research – Charles University, Faculty of Education.

References

Abrahams, I., & Millar, R. (2008). Does practical work really work? A study of the effectiveness of practical work as a teaching and learning method in school science. *International Journal of Science Education*, *30*(14), 1945–1969. https://doi.org/10.1080/09500690701749305

Arnold, J. C., Kremer, K., & Mayer, J. (2014). Understanding students' experiments—what kind of support do they need in inquiry tasks? *International Journal of Science Education*, *36*(16), 2719–2749. https://doi.org/10.1080/09500693.2014.930209

Baltikian, M., Kärkkäinen, S., & Kukkonen, J. (2024). Assessment of scientific literacy levels among secondary school students in Lebanon: exploring gender-based differences. *Eurasia Journal of Mathematics, Science and Technology Education*, *20*(3), em2407. https://doi.org/10.29333/ejmste/14279

Bewersdorff, A., Seßler, K., Baur, A., Kasneci, E., & Nerdel, C. (2023). Assessing student errors in experimentation using artificial intelligence and large language models: a comparative study with human raters. *Computers and Education: Artificial Intelligence*, *5*, 100177. https://doi.org/10.1016/j.caeai.2023.100177

Brown, J. P. (2019). Real-World Task Context: Meanings and Roles. In G. A. Stillman & J. P. Brown (Eds.), *Lines of inquiry in mathematical modelling research in education* (pp. 53–81). Springer International Publishing. https://doi.org/10.1007/978-3-030-14931-4_4

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Routledge. https://doi.org/10.4324/9780203771587

Crujeiras-Pérez, B., & Jiménez-Aleixandre, M. P. (2017). High school students' engagement in planning investigations: findings from a longitudinal study in Spain [10.1039/C6RP00185H]. *Chemistry Education Research and Practice*, *18*(1), 99–112. https://doi.org/10.1039/C6RP00185H

Čipková, E., Šmida, D., & Pecníková, K. (2025). Evaluation of the level of selected inquiry skills among grammar school students. *Science & Education*, *34*(5), 3727–3749. https://doi.org/10.1007/s11191-024-00602-3

Elhai, J. (2023). Science literacy: a more fundamental meaning. *Journal of Microbiology & Biology Education*, *24*(1), e00212-00222. https://doi.org/10.1128/jmbe.00212-22

Fugarasti, H., Ramli, M., & Muzzazinah. (2019). Undergraduate students' science process skills: a systematic review. In N. Y. Indryanti, M. Ramli, & F. Nurhasanah (Eds.), *AIP Conference proceedings: Vol. 2194 (1). 2nd International conference on Science, Mathematics, Environment, and Education.* AIP Publishing. https://doi.org/10.1063/1.5139762

Gay, L. R., Mills, G. E., & Airasian, P. W. (2019). *Educational research: competencies for analysis and applications.* Pearson.

Hamnell-Pamment, Y. (2024). The role of scientific language use and achievement level in student sensemaking. *International Journal of Science and Mathematics Education*, *22*(4), 737–763. https://doi.org/10.1007/s10763-023-10405-7

Harlen, W. (2021). *The case for inquiry-based science education (IBSE).* AIP. https://www.interacademies.org/publication/case-inquiry-based-science-education-ibse

Heindl, M. (2019). Inquiry-based learning and the pre-requisite for its use in science at school: a meta-analysis. *Journal of Pedagogical Research*, *3(2)*, 52–61. https://doi.org/10.33902/JPR.2019254160

Chen, F., & Chen, G. (2025). Learning analytics in inquiry-based learning: a systematic review. Educational technology research and development, *73*(4), 2131–2161. https://doi.org/10.1007/s11423-025-10507-9

Janštová, V., & Pavlasová, L. (2019). Inquiry vs. cookbooks in practical teaching biology viewed by teachers. In M. Rusek & K. Vojíř (Eds.), *Project-based education and other activating strategies in science education XVI* (pp. 30–36). https://page.pedf.cuni.cz/pbe/files/2019/07/sbornikPBE2018_wos.pdf

Jerrim, J., Oliver, M., & Sims, S. (2022). The relationship between inquiry-based teaching and students' achievement. New evidence from a longitudinal PISA study in England. *Learning and Instruction*, *80*, 101310. https://doi.org/10.1016/j.learninstruc.2020.101310

Ješková, Z., Lukáč, S., Šnajder, Ľ., Guniš, J., Klein, D., & Kireš, M. (2022). Active learning in STEM education with regard to the development of inquiry skills. *Education Sciences*, *12*(10), 686. https://doi.org/10.3390/educsci12100686

Kang, J. (2022). Interrelationship between inquiry-based learning and instructional quality in predicting science literacy. *Research in Science Education*, *52*(1), 339–355. https://doi.org/10.1007/s11165-020-09946-6

Karlsen, S., Kersting, M., Ødegaard, M., Olufsen, M., Lill Suhr, M., & Kjærnsli, M. (2025). A comparative study of teachers' conceptualisations and enactment of inquiry-based science education. *International Journal of Science Education*, 1–31. https://doi.org/10.1080/09500693.2025.2547411

Klahr, D., & Chen, Z. (2011). Finding one's place in transfer space. *Child Development Perspectives*, *5*(3), 196–204. https://doi.org/10.1111/j.1750-8606.2011.00171.x

Kok, K. (2022). *Certain about uncertainty* [Dissertation, Humboldt-Universität zu Berlin]. https://doi.org/10.18452/24782

Kranz, J., Baur, A., & Möller, A. (2023). Learners' challenges in understanding and performing experiments: a systematic review of the literature. *Studies in Science Education*, *59*(2), 321–367. https://doi.org/10.1080/03057267.2022.2138151

Lazonder, A. W., & Harmsen, R. (2016). Meta-analysis of inquiry-based learning: effects of guidance. *Review of Educational Research*, *86*(3), 681–718. https://doi.org/10.3102/0034654315627366

Lombardi, D., Bailey, J. M., Bickel, E. S., & Burrell, S. (2018). Scaffolding scientific thinking: students' evaluations and judgments during Earth science knowledge construction. *Contemporary Educational Psychology*, *54*, 184–198. https://doi.org/10.1016/j.cedpsych.2018.06.008

McComas, W. F. (2014). Science process skills. In W. F. McComas (Ed.), *The language of science education: An expanded glossary of key terms and concepts in science teaching and learning* (pp. 89–89). SensePublishers. https://doi.org/10.1007/978-94-6209-497-0_79

Mocanu, M., Bibiri, A.-D., Rusu, V. D., Moroşanu, A., & Bejan, I. G. (2025). Enhancing civic engagement with science: a comparative approach across European regions. *Scientometrics*, *130*(1), 447–468. https://doi.org/10.1007/s11192-024-05198-7

NASEM. (2016). *Science Literacy: concepts, contexts, and consequences.* The National Academies Press. https://doi.org/10.17226/23595

National Research Council. (2012). *A Framework for K-12 science education: practices, crosscutting concepts, and core ideas.* The National Academies Press. https://doi.org/10.17226/13165

National Research Council. (2013). *Next generation science standards: for states, by states.* The National Academies Press. https://doi.org/10.17226/18290

Nejedlý, A., & Vojíř, K. (2022). How do students formulate a research question and conclusions in science research? In M. Rusek & M. Tóthová (Eds.), *Project-based education and other student-activation strategies and issues in science education XIX.* (pp. 29–38). Charles University, Faculty of Education. https://pages.pedf.cuni.cz/pbe/files/2022/10/ProceedingsPBE2021_final.pdf

Nejedlý, A., & Vojíř, K. (2024). How to construct inquiry tasks: a methodological framework and task piloting. In D. Koperová & M. Rusek (Eds.), *Project-based education and other activating strategies in science education XXI.* (pp. 125–135). Charles University, Faculty of Education. https://pages.pedf.cuni.cz/pbe/files/2024/09/PBE2023_final.pdf

Nejedlý, A., & Vojíř, K. (2025). The role of inquiry-based learning: development of analytical skills in orientation and conceptualization by science inquiry at lower secondary school. In G. L. Chova, M. G. Ch., & L. J. (Eds.), *17th International conference on education and new learning technologies* (Vol. 17, pp. 7494–7501). IATED Academy. https://doi.org/10.21125/edulearn.2025.1845

OECD. (2023a). *PISA 2022 results (volume I): The state of learning and equity in education.* OECD Publishing. https://doi.org/10.1787/53f23881-en

OECD. (2023b). *PISA 2025 Science framework draft.* https://pisa-framework.oecd.org/science-2025/assets/docs/PISA_2025_Science_Framework.pdf

OECD. (2025). *OECD science, technology and innovation outlook 2025: driving change in a shifting landscape.* https://doi.org/10.1787/5fe57b90-en

Oliver, M., McConney, A., & Woods-McConney, A. (2021). The efficacy of inquiry-based instruction in science: a comparative analysis of six countries using PISA 2015. *Research in Science Educcation, 51* (Suppl 2), 595–616. https://doi.org/10.1007/s11165-019-09901-0

Osborne, J., & Allchin, D. (2024). Science literacy in the twenty-first century: informed trust and the competent outsider. *International Journal of Science Education*, *47*(15–16), 1–22. https://doi.org/10.1080/09500693.2024.2331980

Osborne, J., & Dillon, J. (2008). *Science education in Europe: Critical reflections.* King's College London. https://www.nuffieldfoundation.org/wp-content/uploads/2019/12/Sci_Ed_in_Europe_Report_Final1.pdf

Pedaste, M., Baucal, A., & Reisenbuk, E. (2021). Towards a science inquiry test in primary education: development of items and scales. *International Journal of Stem Education*, *8*(1), 19. https://doi.org/10.1186/s40594-021-00278-z

Pedaste, M., Maeots, M., Siiman, L. A., de Jong, T., van Riesen, S. A. N., Kamp, E. T., Manoli, C. C., Zacharia, Z. C., & Tsourlidaki, E. (2015). Phases of inquiry-based learning: Definitions and the inquiry cycle. *Educational Research Review*, *14*, 47–61. https://doi.org/10.1016/j.edurev.2015.02.003

Renkl, A. (2014). Toward an instructionally oriented theory of example-based learning. *Cognitive Science*, *38*(1), 1–37. https://doi.org/10.1111/cogs.12086

Riga, F., Winterbottom, M., Harris, E., & Newby, L. (2017). Inquiry-based science education. In K. S. Taber & B. Akpan (Eds.), *Science education: An international course companion* (pp. 247–261). SensePublishers. https://doi.org/10.1007/978-94-6300-749-8_19

Sandoval, W. A. (2005). Understanding students' practical epistemologies and their influence on learning through inquiry. *Science Education*, *89*(4), 634–656. https://doi.org/10.1002/sce.20065

Seeratan, K. L., McElhaney, K. W., Mislevy, J., McGhee Jr, R., Conger, D., & Long, M. C. (2020). Measuring students' ability to engage in scientific inquiry: A new instrument to assess data analysis, explanation, and argumentation. *Educational Assessment*, *25*(2), 112–135. https://doi.org/10.1080/10627197.2020.1756253

Shaffer, J. F., Dang, J. V., Lee, A. K., Dacanay, S. J., Alam, U., Wong, H. Y., Richards, G. J., Kadandale, P., & Sato, B. K. (2016). A familiar(ity) problem: assessing the impact of prerequisites and content familiarity on student learning. *PLoS One*, *11*(1), e0148051. https://doi.org/10.1371/journal.pone.0148051

Shanks, R. A., Robertson, C. L., Haygood, C. S., Herdliksa, A. M., Herdliska, H. R., & Lloyd, S. A. (2017). Measuring and advancing experimental design ability in an Introductory course without altering existing lab curriculum. *Journal of microbiology & biology education*, *18*(1), 1–8. https://doi.org/10.1128/jmbe.v18i1.1194

Schäfer, J., Reuter, T., Karbach, J., & Leuchter, M. (2024). Domain-specific knowledge and domain-general abilities in children's science problem-solving. *British Journal of Educational Psychology*, *94*(2), 346–366. https://doi.org/10.1111/bjep.12649

Schwichow, M., Brandenburger, M., & Wilbers, J. (2022). Analysis of experimental design errors in elementary school: how do students identify, interpret, and justify controlled and confounded experiments? *International Journal of Science Education*, *44*(1), 91–114. https://doi.org/10.1080/09500693.2021.2015544

Sirum, K., & Humburg, J. (2011). The experimental design ability test (EDAT). *Bioscene: Journal of College Biology Teaching*, *37(1)*, 8–16. http://files.eric.ed.gov/fulltext/EJ943887.pdf

Sjøberg, S. (2018). The power and paradoxes of PISA: should inquiry-based science education be sacrificed to climb on the rankings? *Nordic Studies in Science Education*, *14*(2), 186–202. https://journals.uio.no/nordina/article/view/6185/5249

Strat, T. T. S., Henriksen, K. E., & Jegstad, K. M. (2024). Inquiry-based science education in science teacher education: a systematic review. *Studies in Science Education*, *60*(2), 191–249. https://doi.org/10.1080/03057267.2023.2207148

Tairab, H. H. (2015). Assessing students' understanding of control of variables across three grade levels and gender. *International Education Studies*, *9*(1), 44–54. https://doi.org/10.5539/ies.v9n1p44

Teig, N. (2024). Uncovering student strategies for solving scientific inquiry tasks: insights from student process data in PISA. *Research in Science Education*, *54*(2), 205–224. https://doi.org/10.1007/s11165-023-10134-5

van Uum, M. S. J., Verhoeff, R. P., & Peeters, M. (2016). Inquiry-based science education: towards a pedagogical framework for primary school teachers. *International Journal of Science Education*, *38*(3), 450–469. https://doi.org/10.1080/09500693.2016.1147660

van Uum, M. S. J., Verhoeff, R. P., & Peeters, M. (2017). Inquiry-based science education: scaffolding pupils' self-directed learning in open inquiry. *International Journal of Science Education*, *39*(18), 2461–2481. https://doi.org/10.1080/09500693.2017.1388940

Vojíř, K. (2021). What tasks are included in chemistry textbooks for lower- secondary schools: a qualitative view. In M. Rusek, M. Tóthová, & K. Vojíř (Eds.), *Project-based education and other activating strategies in science education XVIII*. https://pages.pedf.cuni.cz/pbe/files/2021/05/ProceedingsPBE2020_final.pdf

Volkmann, M., & Abell, S. (2003). Rethinking laboratories. *Science Teacher*, *70*(6), 38–41. https://www.jstor.org/stable/24156087

World Economic Forum. (2026). *The global risks report 2026*. World Economic Forum. https://reports.weforum.org/docs/WEF_Global_Risks_Report_2026.pdf

Windschitl, M., Thompson, J., & Braaten, M. (2008). Beyond the scientific method: model-based inquiry as a new paradigm of preference for school science investigations. *Science Education*, *92*(5), 941–967. https://doi.org/10.1002/sce.20259

Zetterqvist, A., & Bach, F. (2023). Epistemic knowledge – a vital part of scientific literacy? *International Journal of Science Education*, *45*(6), 484–501. https://doi.org/10.1080/09500693.2023.2166372

Zimmerman, C. (2007). The development of scientific thinking skills in elementary and middle school. *Developmental Review*, *27*(2), 172–223. https://doi.org/10.1016/j.dr.2006.12.001