LUMAT
International Journal on Math, Science and Technology Education

Research article

Data-Information-Evidence-Claim (DIEC): A model for designing and analysing teaching to support students in formulating scientific claims

Per Anderhag[1,2], Jenny Ljung[3], Andre Bürgers[4], Johanna Lundström[4], Mattias W Hugerth[1,2], Linda Klaesson[4] and Kristin Bratsberg[4]

[1] Stockholms universitet, Sweden
[2] Stockholms stad, Sweden
[3] Kunskapsskolan, Sweden
[4] Nacka kommun, Sweden

Abstract: Research in science education consistently shows that students often do not draw on the data they collect during inquiry activities when formulating scientifically founded claims. To address this, we developed the Data–Information–Evidence–Claim (DIEC) model in collaboration between researchers and teachers, making explicit the interpretive step between data and evidence. The study examined the model as both a tool for lesson design and an analytical framework for students' reasoning. The model was used in two iterative research lessons in upper secondary science classrooms where the students worked with a simulation on natural selection. The students' written responses were analysed with a DIEC-based rubric, and their recorded discussions were analysed thematically. The findings show that, despite the use of the DIEC model, students rarely used the data they had collected when constructing their written claims and instead relied mainly on theoretical reasoning. In contrast, their oral discussions involved extensive interpretation of data through trends, relations, and comparisons. The study thus identifies a discrepancy between process and product: interpretive work was present in students' talk but largely absent from their written answers. While the DIEC model showed limited value as a short-term instructional scaffold, it proved useful for lesson planning and for analysing students' scientific reasoning.

Keywords: scientific reasoning, data interpretation, inquiry, instructional model

Correspondence: per.anderhag@su.se

LUMAT Vol 4 No 12 (2026), 16. https://doi.org/10.31129/LUMAT.14.2.3488
Received 7 July 2026 / Accepted 31 October 2026 / Published 17 November 2026
Published by the University of Helsinki, Finland and LUMA Centre Finland.

1 Introduction

This study investigates how students use collected data to formulate scientific claims, and how an educational model may support this process. The study was developed collaboratively by researchers and teachers, grounded in challenges identified both in science education research and in our classroom practices. A recurring problem, observed by the teachers and highlighted in the literature, is that students have difficulties in transforming empirical data into evidence, and how evidence connects to broader scientific explanations (Brown et al., 2010; Kuhn & Modrek, 2022; Osborne, 2013). Instead of seeing evidence as the result of interpretation, students often treat data as self-evident rather than as something that needs to be interpreted and used (Chen et al., 2011; Hardy et al., 2010). This interpretive step involves identifying patterns and relationships in data and considering how these can support claims. Understanding how students engage in this step is therefore important. This also raises questions about the relationship between argumentation as a process and argument as a product (Ryu & Sandoval, 2012). Examining students´ spoken interaction during inquiry alongside their written responses may reveal how they interpret and use data that are not apparent in their written claims alone. This is also the background to this study, in which we investigate a step often treated implicitly in both practice and research, namely how students recognize and articulate patterns in data when doing inquiry (Brown et al., 2010; Hardy et al., 2010; Lee & Irving, 2018). The aim of the study is to develop and explore an educational model intended to support students formulating scientifically well-founded claims based on their collected data and examine its usefulness for instructional design and analysis of students' reasoning. The following three overarching questions guided the study:

1. How do students use data from their investigations to construct scientifically well-founded claims in their written responses?
2. How do students transform collected data into information during their investigations?
3. What opportunities and limitations emerge when using the model for instructional design and analysis of students' reasoning and written claims?

This article is organized as follows: In section 2, we summarize previous research concerning students' ability to formulate scientific claims. In section 3, we present our own instructional model data-information-evidence-claim (DIEC). This model was tested in two research lessons, which are presented in section 4. This section also describes how we collected data, which in turn is analysed in section 5. We conclude with a discussion in section 6.

2 Previous research

Scientific reasoning or argumentation - students' ability to use evidence to construct and evaluate scientific claims - has consistently been emphasized in science education policy and reform as a key learning goal (AAAS, 1993; National Research Council, 2012; OECD, 2006, Swedish National Agency for Education, 2022). Correspondingly, understanding how students engage in evidence-based reasoning continues to be a central research interest within science education (Allchin & Zemplén, 2020; Cavagnetto, 2010; Engelmann et al., 2018; Schwartz et al., 2023; Tosun, 2024; Weiss et al., 2022). Over time, this research field has shifted from an early emphasis on cognitive argument structures toward broader concerns with epistemic practices, classroom discourse, and scientific literacy (Ergunt & Yilmaz, 2026). Typically, scientific reasoning in classrooms involves the transformation of data into evidence, and evidence into explanations and claims, forming a continuous interpretive process that is argued to be central to scientific learning and understanding (Chinn & Iordanou, 2023; Hardy et al., 2010; Manz et al., 2020; Osborne, 2013; Schwartz et al., 2023). This process is not solely cognitive or procedural, but epistemic and social, requiring students to engage in the justification, evaluation, and communication of claims in line with disciplinary norms and practices (Driver et al., 2000; Erduran, 2013; Governor et al., 2021; Osborne, 2013; Schwartz et al., 2023).

A common obstacle to developing scientific reasoning is the confusion between key components of scientific argumentative products, particularly data, evidence, explanation, and theory. Recent research continues to show that students, but also teachers, may conflate these elements, leading to difficulties in understanding how data become evidence, or how evidence is used to justify claims (Kan et al., 2026; Mesci & Erdaş Kartal, 2021; Gyllenpalm et al., 2021). This challenge has been consistently documented in science education research (Brown et al., 2010; Gott & Duggan, 1996; Kuhn, 1993; Osborne, 2013).

2.1 Challenges in using evidence to support claims in science education

A well-documented challenge in science education is students' difficulty in supporting claims with appropriate evidence (e.g. Kuhn & Modrek, 2022; Novak & Treagust, 2018). Research across several decades has shown that students often make unsubstantiated claims or base their assertions on anecdotal information rather than empirical data (Bell & Linn, 2000; Brown et al., 2010; Driver et al., 2000; Gott & Duggan, 1996; Kuhn & Lerman, 2021; McNeill, 2011). Based on an analysis of eighth-grade students' reasoning across seven different forms of evidence, Kuhn and Lerman (2021) argue that a central challenge in scientific argumentation is establishing the relation between evidence and claims, rather than attending to only one of them in isolation. This challenge is for example visible when students describe what they observe without using those observations to support their claims (McNeill, 2011; Brown et al., 2010) or when they rely on pre-existing beliefs or isolated data points rather than making inferences from patterns in data (Hardy

et al., 2010; Chen et al., 2011). Data may thus be treated as self-evident, rather than as something that needs to be interpreted and used (Hardy et al., 2010; Chen et al., 2011). Sandoval's (2003) study of students' work with natural selection illustrates this problem; even when students had access to relevant data, they rarely used it explicitly to support their explanations, instead relying on plausible causal accounts not directly tied to the data. Sandoval interprets this as a gap in students' epistemic understanding of how data should be used to justify scientific claims. Relatedly, Kuhn and Modrek (2022), studying college students, found that even students with adequate analytical skills often failed to identify what constituted relevant and sufficient evidence for supporting a claim. They argue that this ability, choosing one's evidence, is an important but often overlooked aspect of scientific reasoning.

These challenges are also shaped by how practical work is approached in classrooms. Practical work does not automatically foster evidentiary understanding unless tasks are explicitly designed to develop and apply such understanding (Gott and Duggan, 1996; Samarapungavan, 2018). When procedures and outcomes are emphasized over reasoning and justification, students may be left without adequate scaffolding to distinguish between observed phenomena, patterns in those observations, and the scientific ideas they are meant to support (Gericke et al., 2022; Newton et al., 1999; Brown et al., 2010). Overcoming this confusion may require explicit teaching of the epistemic functions of data, evidence, and theory (e.g. Kuhn & Modrek, 2022). It may also require classroom practices that support students in recognizing that evidence is not simply data, but data interpreted with purpose, context, and justification. Dialogic practices such as asking clarifying questions, weighing alternatives, and articulating uncertainty can therefore help students engage more meaningfully with data (Duschl & Osborne, 2002; Governor et al., 2021). Research further suggests that students need repeated opportunities to construct and revise explanations over time in order to learn how to use scientific ideas as evidence. Through iterative work with explanations, students develop more integrated understandings and become better able to apply new knowledge in new situations (Novak & Treagust, 2022).

2.2 Instructional models for supporting students' scientific reasoning

Instructional models differ in their emphasis - some focus on the structure of arguments, instructional design, or examine how reasoning develops through discourse. Taken together, they offer insights into both the challenges and possibilities for helping students move from observation to evidence and build meaningful scientific explanations.

The Evidence-Based Reasoning (EBR) framework developed by Brown et al. (2010) structures scientific reasoning as a transformation across three stages: analysis, interpretation, and application. The model emphasizes the conceptual and epistemic work involved in moving from empirical observations to generalizable claims. This framework was adapted by Furtak et al. (2010) into a coding scheme for analysing reasoning in classroom discourse. Their version identifies how students and teachers jointly construct

arguments during whole-class discussions by distinguishing between claims, data, evidence, and rules. The framework also includes teacher prompts and contributions as part of the analytic lens, highlighting the interplay between instructional support and the quality of reasoning. Like the original EBR model, it sees reasoning as a continuum from unsupported to rule-based, with a focus on how data is generalized into evidence through dialogue. Similarly, Hardy et al. (2010) applied the EBR framework to classroom talk and found that while most reasoning stayed at an observational level, relational and rule-based reasoning appeared when scaffolding was strong.

The Claim-Evidence-Reasoning (CER) model has been used as both a structural tool and instructional scaffold. McNeill (2011) implemented CER in a year-long study aimed at supporting elementary students in constructing written and oral arguments. The framework emphasizes the coordination of three components - claim, evidence, and reasoning - and was shown to help students produce more complete arguments when teachers explicitly modelled scientific thinking. However, while students referred to evidence, the process of identifying patterns in data or transforming observations into generalized claims received less emphasis. McNeill and Martin (2011) used CER to explore how instructional strategies and classroom talk support students' written argumentation. While students often included measurements in their writing, they struggled to link these data to scientific principles such as work or force, indicating a need for continued support in generalizing from data and applying core scientific ideas.

Other models explicitly link reasoning with instructional design. Sampson and Grooms (2009) introduced the Generate an Argument model, in which students collaboratively build scientific arguments using provided datasets and theoretical ideas. Although students do not collect their own data, they analyse tabulated information to identify trends and construct claims supported by evidence and reasoning. The revised Argument-Driven Inquiry (rADI) model developed by Songsil et al. (2019) adapts Toulmin's (1958) structure to include counterarguments and peer feedback. Implemented in classrooms, the model required students to construct claims using empirical data and provided scaffolding through teacher support and peer interaction. Results showed improved use of evidence in arguments, although students continued to struggle with certain elements, such as justifying claims. The findings suggest that while data use was supported within the instructional design, students continued to struggle with certain components of argumentation, particularly in providing sufficient justification for their claims.

Taken together, these studies show that the process of moving from data to pattern - transforming observations into meaningful information - is an essential part of scientific reasoning. While not always made explicit, this step is reflected in frameworks using terms such as relational reasoning (Hardy et al., 2010), analysis and interpretation (Brown et al., 2010), or patterns from data (Lee & Irving, 2018). In instructional models like Generate an Argument identifying trends across data points is a central element, even if not always coded separately. Across these models, a shared challenge becomes evident: even if students can observe and describe phenomena, they often need explicit support to

recognize patterns and construct intermediate representations that connect data to scientific concepts.

3 Developing the DIEC model

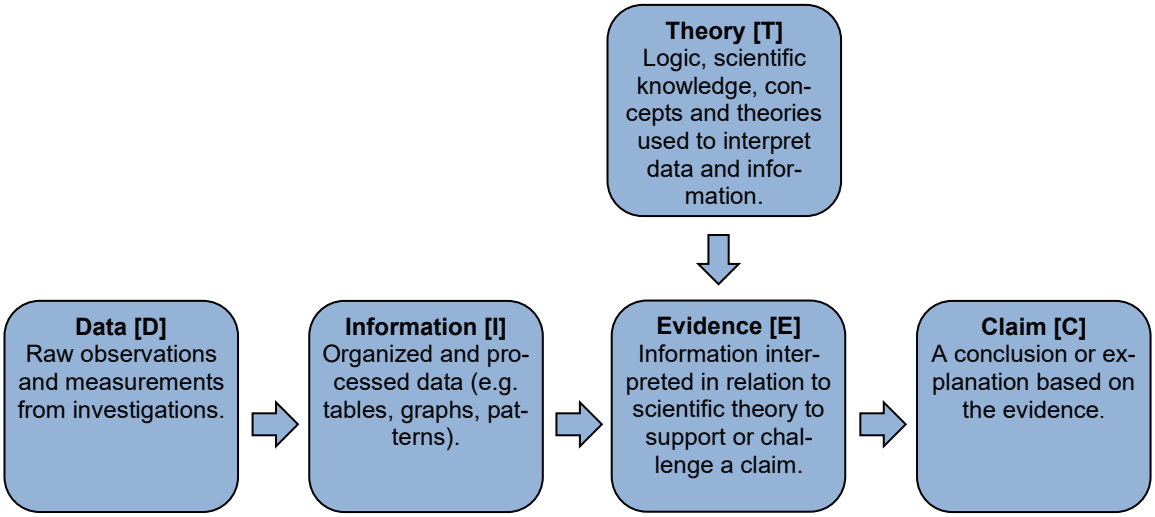
The research project presented here was a collaboration between researchers at Stockholm University and teachers from different schools in the Stockholm region in Sweden. The DIEC model was developed in the beginning of this project through discussions between researchers and teachers, as illustrated in Figure 2 below. Our aim was to create a model that would help students to identify the crucial steps in formulating an answer to their research questions, hoping that this would help them to formulate their own research answers.

The model draws inspiration from existing frameworks such as the Claim-Evidence-Reasoning (CER) model (McNeill et al., 2006) and the Evidence-Based Reasoning (EBR) framework (Brown et al., 2010), which have been used both to scaffold instruction and to analyse student reasoning. At the same time, previous research has pointed to a recurring difficulty among students: identifying and interpreting patterns in data before using it to support a claim (Hardy et al., 2010; Gericke et al., 2022; Lee & Irving, 2018). Although models such as CER include "evidence" as a central element, the interpretive work involved in turning raw data into usable evidence is often treated implicitly in both research and teaching.

In response to this, the DIEC model introduces Information as a distinct phase in the reasoning process (cf. Schildkamp & Kuiper, 2010) - positioned between the collection of Data and the construction of Evidence to justify a Claim. This addition reflects findings from studies that highlight students' difficulties in distinguishing between raw observations and data that have been organized or contextualized for explanatory purposes (Osborne, 2013; Driver et al., 2000). By making this interpretive step explicit, the model seeks to highlight a process that is often present in classroom activity, but not always visible in students' final products or instructional design. Within this framework, *data* refers to unprocessed observations - such as measurements, numbers, or visual representations generated during experiments or simulations. *Information* refers to data that have been interpreted and structured by students, identifying trends, differences, or other relevant patterns, and perhaps visualizing the information in diagrams, tables and graphs. *Evidence*, in turn, is the combination of information or data with theoretical concepts or logical reasoning that gives it argumentative power. In this sense, evidence is not separate from data or information but builds upon them through conceptual and epistemic justification. This distinction aligns with the view that scientific argumentation requires both the interpretation of empirical observations and the application of disciplinary knowledge (Osborne et al., 2004; Sampson et al., 2011). A *claim* refers to a statement or conclusion about a phenomenon - what it is like or how it works.

In the first version of this model, *theory* was not included explicitly. However, when analysing the research lessons, we noted that students continuously referred to theory when formulating their claims. We therefore inserted *theory* to visualize this phenomenon.

Figure 1. The DIEC model



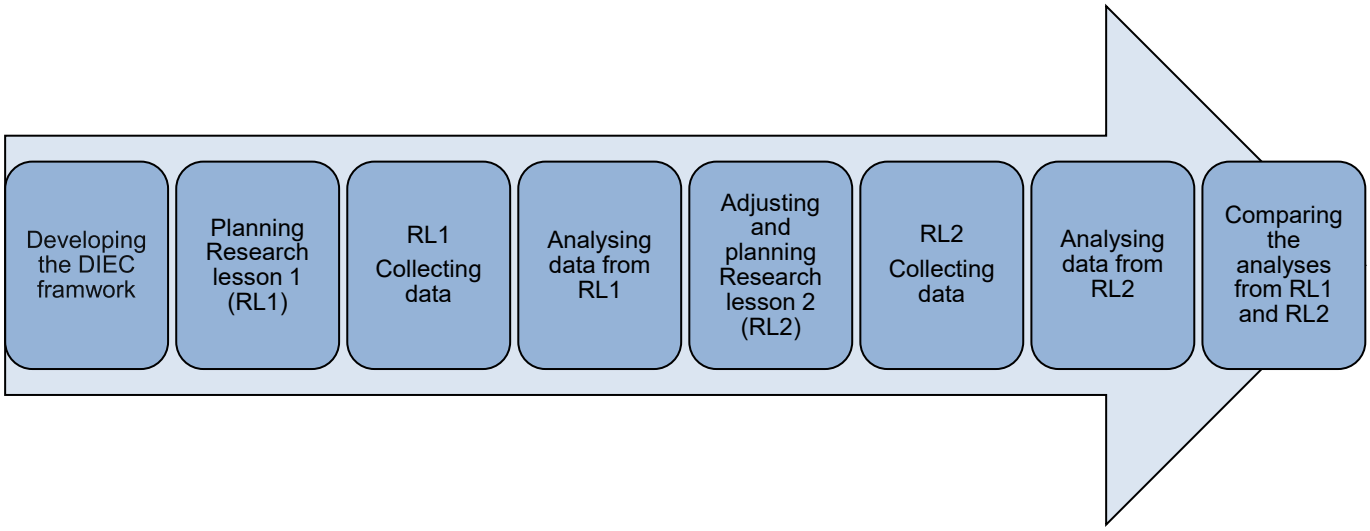
Note. The model illustrates the process of scientific reasoning by distinguishing between four components: Data, Information, Evidence, and Claim. Each box represents a product of student activity, while the arrows indicate the interpretive processes that transform one component into the next.

The DIEC model was used in two ways in this study. First, it served as an instructional tool in lesson planning, aiming to provide opportunities for students to engage with each component of scientific reasoning. When talking with the students, the teachers stressed that the purpose of the model was to guide them in identifying central components of a scientific argument. Second, it was used as an analytical framework for evaluating both the students' written responses and their recorded conversations while performing their task, as will be explained in the following section.

4 Methods

When the DIEC framework was developed, the study's research lessons were planned based on two of the participating teachers' ongoing teaching. The study followed an iterative process in which a research lesson was planned, conducted, analysed and revised in two cycles, as illustrated in Figure 2.

Figure 2. Chronological illustration for the research.



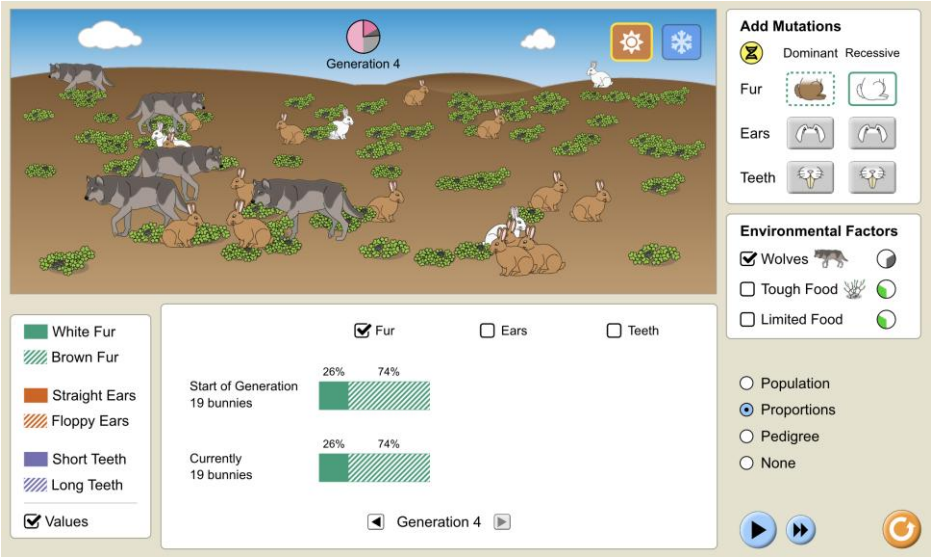
Note. This illustration shows the chronological work for the research starting from the development of the DIEC model as a framework for planning and analysing. Data was collected from RL 1 and RL 2.

4.1 Study context

The study was conducted at a Swedish upper secondary school (ages 16–18) in the Stockholm region, where the first iteration was conducted in science class (5 groups) and the second in Biology class (6 groups). In total, 24 students participated in the study, 11 in research lesson 1 and 13 in research lesson 2.

Students worked with a teaching material provided by the Swedish Center for School Biology, utilising data simulations of natural selection developed at the University of Colorado in Boulder (PhET Interactive Simulations n.d.). In the simulation a rabbit population is followed. The user can select different aspects of climate (temperate or arctic), presence of predators (wolves) as well as different properties of the rabbits (fur colour, form of the ears and length of the teeth). The physical properties follow independent Mendelian genetics, but the user is able to choose whether a trait is dominant or recessive.

Figure 3. Screenshot of the PhET Simulation



Note. Screenshot of generation four, data display in percentages when selecting white fur recessive and with wolves.

In our case, students were asked to observe the population dynamics of the rabbit population dependent on which fur colour (white or brown) was dominant or recessive, respectively, and whether predators were present or not. The underlying theory is natural selection which aligned with the teachers' original planning.

During the research lessons, students used tablets to run the simulation. Their work was recorded using the tablets' built-in screen recorder, capturing both screen activity and audio. Due to technical problems in lesson 2, audio was not recorded in three of the groups. Thus, audio from three groups (7 pupils) were analysed from lesson 2. In addition, one of the groups in lesson 2 did not answer on the final task, which means that the written answers of five groups were analysed.

4.2 Lesson design and data collection

Cycle 1

In Cycle 1 we used an existing lesson plan provided by the Swedish Centre for School Biology together with the simulation. The student guide of the lesson plan contained instructions and fields and tables where students could write down data and answers to questions during the activity. The guide included several questions, culminating in the final task of responding to the question: "Does it make a difference whether the mutation is dominant or recessive for the course of evolution?" To this question we inserted the prompt: "Formulate a well-founded claim."

Before the research lesson, the teacher conducted a preparatory session designed to support students in constructing scientific explanations. The session began with a whole-class discussion about the characteristics of a well-founded claim, followed by an

assignment adapted from McNeill and Krajcik (2008), in which students were asked to explain a phenomenon based on a provided data table (see Appendix A in McNeill & Krajcik, 2008). After completing the task, the teacher introduced the key concepts of the DIEC model. She exemplified data, information, evidence and claim, by the use of test scores for grading, a context which is highly familiar to students. No additional DIEC-specific instructional materials were provided to the students. She explained each component and emphasizing how a well-founded claim often is developed through this sequence. This step corresponded to common student difficulties in distinguishing between key stages in the scientific reasoning process (Brown et al., 2010; Osborne, 2013). At this point the DIEC model was not conceptualized or illustrated as a successive chain as in its final graphic form (Fig. 1). The class then applied the model to analyse a sample response to the same assignment they had just completed. The lesson ended with a discussion and a summary of their analysis.

The main assignment, carried out during the research lesson, consisted of two parts and was based on the visual simulation of natural selection described above. In Part 1, students familiarized themselves with the simulation by testing various combinations with potential relevance for rabbit survival such as food availability, fur colour, and the presence of predators. They were instructed to observe outcomes and identify, discuss, and answer questions regarding traits, resources, and factors such as predation or competition that may be important for the survival rate of rabbits.

In Part 2, students focused on the role of fur colour in the evolutionary process. They explored how survival was affected depending on whether the brown fur trait was dominant or recessive, and whether predators were present. The students investigated four different scenarios: 1–2 brown fur recessive, white fur dominant without and with wolves in the habitat, and 3–4 brown fur dominant, white fur recessive without and with wolves in the habitat. They documented percentage distributions of fur colour (e.g., proportion of white and brown rabbits) for generations 3, 6, 9, 12, and 15 in a table. After collecting data from scenarios 1–2, they were prompted to discuss and answer the following questions:

1. What happens to the proportion of rabbits with different fur colours when there are no wolves in the area compared to when wolves are present? Use the table values and describe the simulation results in words. Describe any patterns you can see (NOTE: do not try to explain here – only describe what the data shows!).
2. Provide a possible explanation for the patterns you described – why does the proportion of rabbits with different fur colours look the way it does in the population with and without wolves?
3. The simulation is a simplified model. What other factors in nature, not shown in the model, could influence the evolution of rabbits and wolves in reality?

After the students had investigated scenario 3–4 they discussed and formulated a written response to the following question:

4. Does it make any difference for the course of evolution whether the mutation is dominant or recessive? Formulate a well-founded claim.

Following the investigation, students exchanged their written reports with peers. As part of the peer assessment, they were instructed to identify the elements of Data, Information, Evidence, and Claim in their classmates' responses and evaluate whether the claim was well-founded. The lesson ended with the teacher, who supported and interacted with her students during the lesson as she normally does, giving a short summary of the activity.

Cycle 2

After an initial analysis of Cycle 1, we made the following adjustments in Cycle 2:

- In research lesson 1, the activity involving identification of Data, Information, Evidence, and Claim in peer reports did not work as intended. Students had difficulties locating and agreeing on the different components. We therefore decided to omit this part in Cycle 2.
- Also, the preparatory session was modified. The DIEC-model was never introduced explicitly to the students. Instead, they read texts that they had written for previous inquiry activity and discussed what characterizes a well-founded claim. The teacher then summarized their conclusions and highlighted key aspects considered important.
- Students began directly with Part 2 of the research lesson (see above). In Cycle 1, the students had spent considerable time testing and discussing features that did not focus their attention on formulating a claim (although they did discuss content relevant to their understanding of natural selection and Mendelian genetics). To enhance reasoning and reflective pausing (Gericke et al., 2022), the following instructions were included: "Compare the data in the tables. What patterns do you see? Describe any similarities and differences."
- The students were instructed to collect data from a greater number of generations: 2, 4, 6, 8, 10, 12, 14, and 16. The rationale was that more generations might make patterns more notable. They also had their tables on a separate sheet of paper. This was intended to encourage the students to take a more active role in organizing and processing the data, and to ensure that they had the tables in front of them throughout the task. In research lesson 1, pre-made tables were placed on different pages of the student instructions.
- Before formulating their claims, student groups were asked to discuss any patterns they had observed with another group. This activity was intended to provide the students with a clear purpose, helping to focus their attention on identifying patterns in the data.

In addition to the screen recordings, all student materials were collected, including written responses and worksheets. The screen recordings and written group reports were

transferred to the university's appointed storage. The films were then first transcribed using Automatic Speech Recognition software. For the first cycle Whisper large-v1 (Radford et al., 2023) was used on a local computer. For the second cycle a more advanced pipeline was used, namely a slightly modified WhisperX (Bain et al., 2023) which utilises Whisper (at that point large-v2), Pyannote (Bredin et al., 2020; Bredin & Laurent, 2021) and a language specific model which in this case was KBlab's Wav2vec 2.0 large VoxRex Swedish (Malmsten et al., 2022). The first version gave a rough draft for the transcription as a large text block, while the second version gave a time-stamped version roughly divided by speaker.

Figure 4. Comparison of research lessons

	Lesson 1	Lesson 2
Aim	Learning to collect and use data to formulate a scientific claim.	
	11 students Science class	13 students Biology class
Data collection	Screen recording with voice recording on iPads Documentation on worksheets	
Material	Simulation on natural selection (PhET) Ready-made material from Swedish Centre of School Biology	Simulation on natural selection (PhET) Updated material – less questions, separate tables with more data points
Simulation	Simulation to investigate the effect of different variables on the fur colour of rabbits • White or brown fur colour • Dominant or recessive allele • With or without predators	
Pre-lesson	Introducing the DIEC-model with examples	Discussion on what constitutes a well-founded claim by examining previous work without introducing the DIEC-model
Main lesson	Simulation (see above) Peer response using the DIEC-model	Simulation (see above) Peer discussions on identified patterns in the collected data

4.3 Data analysis

Two complementary forms of qualitative analysis were conducted. First, students' written responses were analysed to examine how they constructed scientific claims, using the DIEC framework as an analytical tool to identify how data, information, and theory were used in their reasoning. A rubric was developed for analysing written responses (Figure 5), inspired by McNeill et al. (2006), in which each of the four components is assessed separately on a scale from 0 to 3. The components were assessed by several authors independently, and the results discussed. The rubric was then revised, and a new round of analysis conducted, until agreement was reached. It should be noted that claim was assessed whether it was correct or not and thus separated from evidence. The rationale for this decision was twofold. First, analytically and for teachers' assessment of student texts,

separating the claim from the evidence provided a straightforward entry point for evaluating student reports. In some instances, for example, a claim is not formulated or is difficult to understand or is wrong in one or several aspects. Secondly, pedagogically, when teaching – claim is also a productive start for talking about a scientific argument. The students' everyday understandings of what is meant by a claim (or statement, conclusion, etc.) thus serve as a ground for entangling what makes a scientific claim right or wrong, good or bad, etc. For example, a claim could be right but at the same time not well-founded. And the opposite, of course, given the evidence – a claim could initially seem plausible but actually turns out to be, when more data is collected and analysed, wrong. Finally, whether a claim is right or wrong needs to be understood in relation to the specific purposes and contexts of the inquiry activity - in some situations, neither the teacher nor the students know whether the claim is correct. Rather, the issue is the plausibility of the claim given the available evidence.

Figure 5. DIEC rubric for analysing written statements

	0 None	1 Poor	2 Good	3 Strong
Data Scientific data which could be used to support the claim.	Does not present any data.	Presents data which either is incomplete or irrelevant.	Presents relevant data but in an incorrect way.	Presents all relevant data in a correct way in suitable tables.
Information Processed data which can be used for supporting the claim.	Does not present any information.	Presents some information.	Presents more relevant information.	Presents all relevant information based on the collected data.
Evidence Argument which links information and theory to give support for the claim.	Does not present any arguments to support the claim.	Presents arguments grounded on either information or theory alone.	Presents arguments that combine information with theory, but the information is vague or incomplete.	Presents arguments that combine correctly used information with theory referring to the collected data.
Claim A claim or conclusion which gives an answer to the research question.	Does not present any claim.	Presents a false claim.	Presents a correct claim which is either incomplete or unclear.	Presents a claim which is correct.

In the second step of the analysis, students' spoken interaction during the inquiry activity was analysed to capture how reasoning unfolded in the classroom. This dual approach made it possible to investigate both the *product* (written claims) and the *process* (oral meaning-making) of scientific reasoning, and to examine how students moved from data, through information and evidence, to claims in their reasoning.

A qualitative content analysis (Hsieh & Shannon, 2005) was conducted on transcribed student conversations during the inquiry activity. This analysis aimed to explore how students engaged with data in real time, particularly how they identified patterns and began to construct meaning through interaction. The analysis proceeded in two steps. First, a deductive coding scheme was applied based on components of the DIEC model, *Data*, *Information*, and *Evidence*, to categorize different types of student talk. We chose to use a broader meaning of evidence in this phase of the analysis to capture student utterances in which they accounted for, interpreted, or made claims about observed phenomena, not necessarily framed as evidence supporting a claim. This allowed us to include reasoning that was exploratory, tentative, or descriptive in nature. Second, an analysis was conducted to identify recurring patterns in students' talk about data, focusing on how they interpreted and transformed data into information.

In both analyses, the initial coding was conducted independently by author 1—3. Each author coded the same data set separately - written responses in the rubric-based analysis and transcripts in the thematic analysis. The transcripts of the thematic analysis were entered into a spreadsheet, with each row corresponding to a single student utterance, resulting in a total of 2826 lines. Cases where coding differed, or where the categorization of a student utterance was uncertain, were subsequently discussed in detail. Through iterative dialogue, these cases were resolved until full consensus was reached.

4.4 Ethical considerations

The study follows the ethical guidelines of the Swedish Research Council (2017). The participating students were guaranteed anonymity, that their participation was voluntary, that the decision to/not to participate would not affect their grades or learning opportunities, and that they could withdraw their participation whenever they wanted. Participating students took part in and signed a letter of consent in which the purpose and the structure of the study was presented. The data of the study is handled according to the General Data Protection Regulation (GDPR).

5 Results

This section presents the results from both research cycles. In cycle 2 there was a small shift into a higher quality in written claims. The analysis of the transcripts showed no difference in how the students talked during the activity in Cycle 1 and Cycle 2. We begin with the students' written claims, focusing on how they used data, information, and theory to support their conclusions (research question 1). We then present the analysis of students' conversations during the activity, focusing on how they interpreted and discussed data (research question 2). The applicability of the DIEC model (research question 3) is partly addressed in the Method section through the adjustments made between the two cycles, and further discussed in Section 5.3, where we summarize the opportunities and

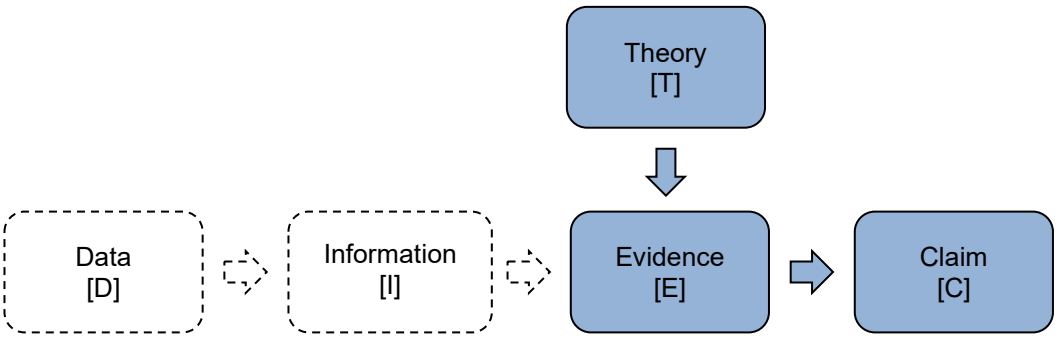
limitations of using the model for instructional design and analysis.

5.1 Students use of data in written responses

Here we present the findings corresponding to research question 1: How do students use data from their investigations to construct scientifically well-founded claims in their written responses?

As described in the methods-section, the student groups followed an instruction guiding them to collect and discuss data from a natural selection simulation. In general, the groups had little difficulties following the instructions and interacting with the simulation. All groups, except for one, formulated a written claim on whether the allele for fur colour (brown or white), dominant or recessive, had (or had not) any effect on the evolutionary process. Only to a lesser extent were these claims grounded in the data they had collected. For most groups, claims (marked by C in the examples below) were instead primarily backed up by theory (marked by T). If any data or information was present, it was marked D or I, respectively. Figure 6 below illustrates this pattern.

Figure 6. A common pattern in students' written responses



Note. All groups formulated a claim, typically supported by evidence grounded in theoretical reasoning. References to data and information were less salient, indicated by the transparent boxes.

The analysis thus showed that while all student groups formulated a claim, their written responses only to a lesser extent included explicit references to the data they had collected or to patterns identified during the activity. The first example illustrates a written response primarily supported by theoretical reasoning.

> Example 1: "A dominant or recessive allele obviously plays a role"
>
> **A dominant or recessive allele obviously plays a role [C]**, as it **makes it easier for the dominant allele to spread in a larger population [T]**. However, the climate plays a greater role in determining which fur **colour blends in better if predators are present [T]**. Another factor is the fur colour of the first rabbits, as an evolution or mutation that results in brown fur must occur. Since it started with white rabbits, which are **recessive and require the gene in both alleles to be white, a mutation is necessary for brown rabbits to appear [T]. Brown rabbits survive the wolves, and even if they produce white rabbits, the white ones are quickly killed by the wolves [I, T]**. (Cycle 1 group 4)

Figure 7. Analysis of students written response, using the DIEC rubric.

Cycle 1 Group 4	0 None	1 Poor	2 Good	3 Strong
Data			X	
Information		X		
Evidence		X		
Claim			X	

Note. Analysis for the written response given in Example 1, using the DIEC rubric as shown in Figure 5. The evaluation of "Data" refers to the tables the students presented in their worksheet, which are not presented here.

In the students' written answers, the evidence used for arguing that dominant or recessive allele matter for the evolutionary process was grounded in Mendelian genetics and natural selection. For example, the advantage of a fur colour that makes the rabbit more difficult to detect for a wolf or the spread of dominant or recessive alleles in the population were typically used as backing for their claims. Data, which they had collected and inserted in tables, were rarely referred to and when information was presented it was usually vague with phrasings like "larger population" or "white ones are quickly killed" potentially implying what they had observed in data. Example 2 further exemplifies this, here the students referred to an observed relationship between fur colour and predators.

> Example 2: "There was no difference (except for the values)"
>
> There was **no difference (except for the values [I]) between dominant and recessive traits [C]** for fur colours when it comes to the evolutionary process. White rabbits **dominate regardless of D or R traits when there are wolves [I]**. The same goes for brown rabbits. **They dominate regardless of D or R traits when wolves are present [I]**. Why? It does not depend on whether the fur colour is D or R; the principle is the same. Brown rabbits **"camouflage" themselves from the wolves and survive [T].** White rabbits **cannot hide and get eaten [T].** But what makes white rabbits survive best without wolves (with limited food) even though brown rabbits are dominant? (Cycle 1 group 3)

Figure 8. Analysis of students written response, using the DIEC rubric.

Cycle 1 Group 3	0 None	1 Poor	2 Good	3 Strong
Data				X
Information		X		
Evidence			X	
Claim		X		

Note. Analysis for the written response given in Example 2, using the DIEC rubric as shown in Figure 5. The evaluation of "Data" refers to the tables the students presented in their worksheet, which are not presented here.

The students claim that there is no difference between dominant and recessive traits in terms of their effect on the evolutionary outcome. They suggest that survival is primarily determined by environmental factors, specifically whether predators are present, rather than by whether the trait is dominant or recessive. Brown rabbits will dominate when wolves are present. At the same time, they note that there are differences when one looks at data ("values") which potentially could have resulted in fruitful discussions. For example, in the scenario where white fur is the dominant allele and disappears from the population (as it does in their simulation, with 0% white rabbits in generation 9), it has consequences for the evolutionary process. However, this is not something the students noted or discussed during the activity.

The following example represents what the teachers considered being a high-quality answer, gaining the highest score in the DIEC rubric (Figure 5):

> Example 3: "It does not matter whether the gene is dominant or recessive"
>
> **It does not matter whether the gene is dominant or recessive as long as there are no predators present [C]** since both end at **about 95% [D]**. **But when there are** [predators present] **they play a role [C]** since the colour **disappears faster when the non-advantage colour is recessive [I]**. The white **died out after only two generations when brown was dominant [I]**. They died since they were **not suited for surviving in the forest [T]**. When white was dominant, they were **still present in the eighth-generation [I] since it takes only one rabbit for the trait to survive, but in the end, even they got extinguished [E]**. (Cycle 2 group 6)

Figure 9. Analysis of students' written response, using the DIEC rubric.

Cycle 2 Group 6	0 None	1 Poor	2 Good	3 Strong
Data				X
Information				X
Evidence				X
Claim				X

Note. Analysis for the written response given in Example 3, using the DIEC rubric as shown in Figure 5. The evaluation of "Data" refers to the tables the students presented in their worksheet, which are not presented here.

5.2 Students transforming data into information

Here the results corresponding to research question 2 are presented: How do students transform collected data into information during their investigations?

The lack of an empirical grounding in the students' written answers does not mean that data was not part of the inquiry process. The analysis of student talk during the activity revealed that data played a central role in their discussions. Student talk concerned, in descending order of prevalence, (a) observing and recording allele percentages to be entered into the table, followed by (b) observations of notable

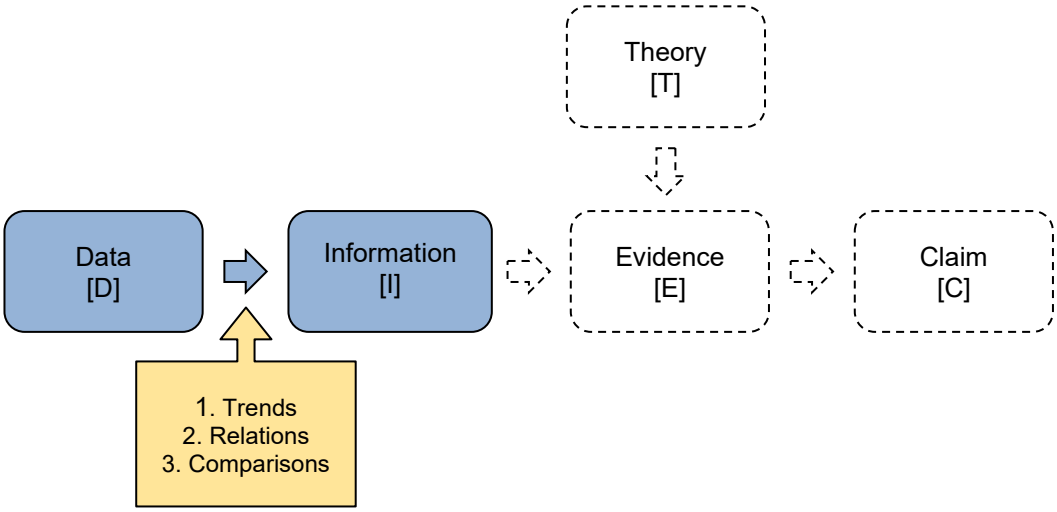
differences or anomalies in the data, or (c) discussions about the final values in the simulations - that is, the values from the last generation. The following short excerpt demonstrates a typical student discussion. The students were preparing to formulate a claim and were commenting on the table containing their collected data, as seen in turns 1–2 (category (a) above).

Example 4: “And then two, ninety-eight. Whoa. How come?”

1. S3: Hundred percent. On the brown ones.
2. S1: And then Hundred. Hundred percent.
3. S3: And then two, ninety-eight. Whoa. How come?
4. S1: Okay, hey. Okay, guys.
5. S2: If it was a hundred percent, you can't even change it.
   (Cycle 1, group 3)

In turn 3, Student 3 noticed something unexpected in their data: in a preceding generation, white rabbits suddenly reappear (two percent of the population) after having previously disappeared (category (b) above). Such instances sometimes resulted in more elaborate discussions in which students noted certain aspects in the data. In these cases, students began to describe and interpret the data beyond simply stating percentages. Their talk indicates that some transformation of data into information occurred, as they identified patterns and made contextual observations that gave the data meaning within the simulation setting. These instances could be categorized into three main themes, based on the nature of students’ observations: (1) Trends and descriptive statements, (2) Relations between variables, and (3) Comparisons across scenarios. These three themes are described in more detail below, see also Figure 10.

Figure 10. Themes of information in student talk during data collection



Note. Illustration how students interpreted and made sense of their data during data collection: (1) Trends and descriptive statements, (2) Relations between variables, (3) Comparisons across scenarios. Evidence and claim blocks remain transparent since students had not yet started formulating claims.

1. Trends and descriptive statements, in these cases students identified temporal patterns in the data, such as steady increases or decreases in the proportion of certain traits - or, as exemplified in the following excerpt, noting a stable population. Before turn 1, the students had decided to test a scenario where the brown allele was dominant, the food was limited, and no wolves were present in the habitat. Now they were discussing the outcome of the simulation as they looked at their data table.

Example 5: “It doesn't show any major change”.

1. S1: Eighty-six... fourteen... twelve becomes... eighty-seven
2. S2: Um, it would... Ah, exactly... eighty-two... eighty-two... eighty-two... I mean, it stays almost at the same level the whole time.
3. S1: Yeah...
4. S2: It doesn't show any major change.
5. S1: That's true. But during generation twelve and generation nine, it increases, and then it decreases again by generation fifteen. But...
6. S2: I mean, I don’t think it’s a really, really big difference.
7. S1: No, no...
(Cycle 1 group 4)

Although they noted some fluctuations (turn 5) and that “it stays almost at the same level the whole time” (turn 2) they jointly concluded that it is a stable population under these specific conditions. Based on this conclusion, they then decided to test a new scenario by introducing wolves into the habitat.

2. Relations between variables, in these cases students pointed to possible relationships or links between variables in the simulation. In the following excerpt, the students summarized their observations and identified a relationship between fur colour and the presence of wolves in the habitat (turns 1–2).

Example 6: “The one that is best adapted survives”

1. S1: Yes, and when brown was dominant it still started with more white, but without wolves the white were always more. Even though brown was dominant.
2. S2: Mm...
3. S1: Oh, when there were wolves like this, then there were a lot more brown rabbits. The white rabbits went extinct almost immediately.
4. S2: Oh, that’s kind of it.
5. S3: Mm. The one that is best adapted survives.
(Cycle 2 group 6)

The first aspect, the effect of predators, was something all student groups identified as a clear pattern in their data. The second aspect, the role of dominant and recessive alleles, was more challenging for students to interpret. All groups were aware that the recessive

allele could remain hidden between generations, and several groups observed that the dominant allele appeared to have little effect on survival when predators were present. However, students rarely formulated conclusions based on observable patterns related to allele behaviour over time. This may be linked to their tendency to focus on the final values - that is, the data from the last generation - rather than on changes across generations.

3. Comparisons across scenarios, in these cases the students compared data or information from different investigations, simulation conditions, or trait combinations. In the following example, the students were discussing the effect of dominant brown allele. They noted that the frequency of brown rabbits was greater than in the previous scenario (turns 1–3), where brown allele had been recessive. They jointly concluded that even if the increase was faster, it did not have any effect on the outcome (turns 5–6).

Example 7: "Here, the increase is faster..."

1. S1: Here, the increase is faster...
2. S2: Yeah.
3. S1: ...then it is... Look. Here the increase in brown ones is larger than it was before, without wolves. Same thing with wolves.
4. S2: Umm...
5. S1: So, I mean, it doesn't really make much of a difference.
6. S2: Okay, but how should we formulate a well-founded claim? (Cycle 1 group 4)

Throughout the activity, students continuously discussed their understanding of what they observed in the simulation and in their data. The students used concepts and theory from Mendelian genetics and evolution, as well as emotional arguments, in order to explain what they observed. This interpretive work appeared to constitute a central part of how students made sense of their observations - both in terms of percentages and through visual representations of population changes in the simulation (rabbits jumping around) - and began to transform it into information.

5.3 Summary of results

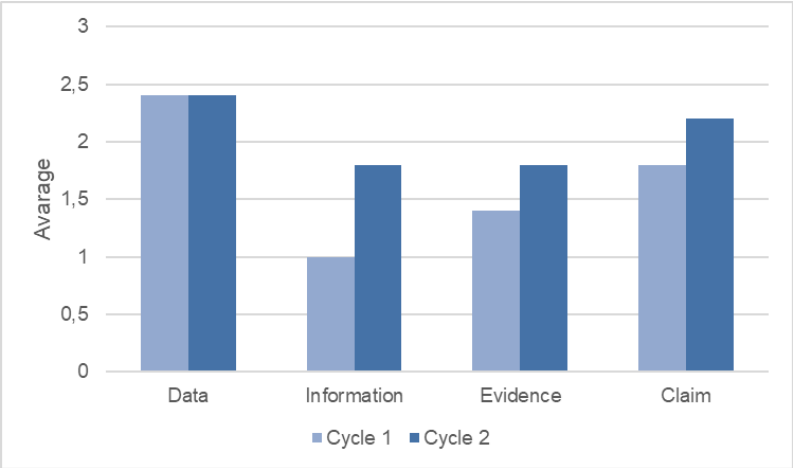
The results from the study above are summarized in this section in relation to each research question.

5.3.1 How do students use data from their investigations to construct scientifically well-founded claims in their written responses?

The students were to answer this final question in their written reports: "Does it make a difference if the mutation is dominant or recessive for the course of evolution?". They were also reminded that they should formulate a well-founded claim to that question. When

analysing the students' written reports using the DIEC rubric there were no differences between cycle 1 and cycle 2 in the students' ability in collecting and presenting data in tables. The students had none or only a few difficulties in collecting the correct data from the simulation. Regarding the use of relevant information from this data as well as using this information for constructing evidence and formulating correct claims, the results from cycle 1 show that students only to a small extent use information in their written reasoning. The students tend to rely on theory rather than their own empirical findings. There was, however, a small shift toward higher quality from cycle 1 to cycle 2 regarding the information presented. Specifically, students were better at presenting relevant information as well as using this information as evidence, see figure 11. Thus, the written claims in cycle 2 were slightly more correct and more well-founded than in cycle 1.

Figure 11. The results of the students' written reports



Note. Showing the results of the students' written reports by using the DIEC rubric (Fig. 5). The diagram shows a quantification of the qualitative analysis by taking the average score, according to the rubric in figure 3, in each category for all groups in cycle 1 and cycle 2, respectively.

5.3.2 How do students transform collected data into information during their investigations?

The lack of data and information in the written answers did not match the students' discourse. When analysing the students' talk during the task, the process of transforming data into information is present and plays a central role when the students are trying to make sense of the data. As shown in the partial transcripts above there were no obvious differences in how students talked about data and information when comparing cycle 1 and 2.

Three types of data interpretations were identified when students transformed data into information: *trends, relations* and *comparisons*. The students identified trends in data as patterns (stable population, increasing or decreasing population), they noticed relations between variables as predators and fur colour and at a smaller extent they compared the different scenarios whether a fur colour was dominant or not.

The results point to a discrepancy between what students write in their reports and their discussions during the laboratory work. There is a clear difference in what occurs during the process when students are trying to make sense of their data and what is presented to teachers as a final written product.

In both the students' written answers as in their oral discussions, we could observe that they tend to focus on the final line in their tables even when noticing irregularities in their data. This is highlighted in example 2 by "There was no difference (except for the values)…".

5.3.3 What opportunities and limitations emerge when using the model for instructional design and analysis of students' reasoning and written claims?

The results for cycle 1 showed that the DIEC model used by students to identify Data, Information, Evidence, and Claim in peer reports, did not help them in writing a scientifically well-founded claim backed by evidence grounded in empirical data. In cycle 2 the DIEC model was therefore excluded from instructions. The students were no longer asked to use the DIEC model to give feedback to their peers. Despite the use of the model to plan for activities such as identifying what characterises a well-founded claim and putting more focus on patterns in data, there was only a small effect on the students' written products.

However, the DIEC model was still used as an analytical tool in two different ways for analysing students' written reports and the transcripts of student talk. Firstly, the model was used to identify the elements of scientific reasoning according to the DIEC model as exemplified in Examples 1 to 3. Subsequently, the quality of each element was assessed according to the DIEC rubric. This provided a possibility to compare the two cycles as seen in Figure 11 above. Secondly, the DIEC model was used to analyse the transcripts of student talk. As data and information were identified as separate entities in students' arguments, we were able to find three different types of transformations from data to information, as discussed above. The DIEC model proved highly functional as a tool for analysis of both written reports and transcripts.

6 Discussion

The aim of the study was to investigate how students use collected data to construct scientifically well-founded claims, how they transform data into information during inquiry, and how the DIEC model can function as a tool for instructional design and analysis.

The study is based on a relatively small number of students working in groups, and some data loss occurred during data collection, resulting in a more limited empirical basis for certain analyses. The empirical material is also drawn from a limited number of lessons and captures students' work during a single inquiry activity. The findings should therefore be understood as a situated account of students' reasoning in a specific situation rather than as evidence of how such reasoning develops over time.

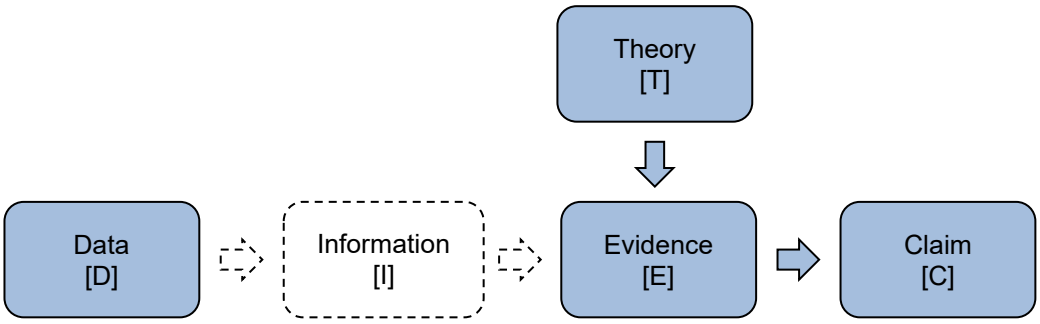
The study demonstrated that the students had little difficulty collecting relevant data or formulating claims about the phenomena under investigation. However, these claims were only to a limited extent grounded in the data they had collected, which is in line with previous research (e.g., Kuhn & Modrek, 2022; Novak & Treagust, 2018; Osborne, 2013). Similar to Sandoval (2003), the students in this study tended to rely on their understanding of natural selection rather than grounding their claims in empirical observations. This suggests that the difficulty may not primarily concern collecting data or producing claims, but rather how data is used as evidence. The findings therefore support previous research demonstrating that students often not construct evidence from data, potentially not fully recognizing that evidence involves the interpretation and justification of data in relation to a claim (Gyllenpalm et al., 2021; Kan et al., 2026; Mesci & Erdaş Kartal, 2021)

At the same time, the analysis of student talk nuances this picture. In contrast to their written responses, students' discussions showed that data had an important role in their reasoning. Throughout the activity, the students were verbally engaged in data interpretation, thus supporting research emphasizing the role of discourse in scientific reasoning (e.g., Duschl & Osborne, 2002; Governor et al., 2021). However, here students' verbal interpretations were not confined to the explicit analytical prompts in the instructions but emerged informally throughout the activity. In these situations, the students identified, discussed and negotiated trends over time, relations between variables such as fur colour and predation, and outcomes across different simulation conditions. These processes correspond to what has been described in earlier research as relational reasoning (Hardy et al., 2010), analysis and interpretation (Brown et al., 2010), or identifying patterns in data (Lee & Irving, 2018). This suggests that important epistemic work may occur during ongoing social interaction during classroom activities. However, as discussed above, the outcome of this work was only to a limited extent reflected in students' written products. When asked to make data-based arguments the students tended to focus on specific data points or on the final outcome, rather than the patterns identified. Even when explicitly asked to comment on patterns, the students still saw the bottom line, or who "wins", as the most relevant information, rather than population changes over time.

One contribution of the study is therefore to point to a discrepancy between scientific reasoning as a process and as a product: while students engage in data interpretation during inquiry, this work is not consistently taken up in their written claims. This may suggest that the activity did not sufficiently support students in recognizing the epistemic role of such patterns in giving weight to a claim, with theory instead becoming the main basis for the argument, as illustrated in Figure 12. The reliance on theory might explain why students at times disregard their own data. For instance, even when students noticed irregularities such as fluctuations in numbers – specifically white rabbits of the recessive phenotype appearing from time to time, rather than reconsidering their understanding, they often interpreted such irregularities as noise. Despite the scaffolding provided during data collection and argument construction, the students might have benefitted from

additional support. One possibility is to introduce writing activities that, as demonstrated by Hand et al. (2004), can help students articulate and develop their understanding.

Figure 12. Typical structure of students' written arguments.



Note. Illustrating, using the DIEC model, how the use of theory becomes a basis for evidence when information derived from data is absent.

The role of theory in students' reasoning appears complex. Even though the lessons at times highlighted describing rather than explaining data patterns, students frequently introduced theoretical concepts such as dominance and camouflage, using these to explain ongoing changes in the data. It remains unclear whether spontaneous intertwining of theory inhibits or supports the process of transforming data into evidence (e.g., Hardy et al., 2010; Chen et al., 2011). Although the consequences of this intertwining are not fully understood, it is evident that the students in the study drew on theory as a resource for making sense of what they observe and for negotiating their interpretations. This suggests that theory, in some activities and situations, plays a central role for interpreting data, which is not surprising but rather something that one could expect and also should support as a teacher. The issue is thus not the use of theory itself, but that theory may become the primary basis for arguments, with information derived from data playing a more peripheral role in the justification of claims.

The design of the inquiry activity may also have influenced how students used theory and empirical data when formulating their claims. Working with a predefined research question and having limited opportunities to discuss or plan the investigation may have made the purpose of collecting and interpreting data less clear, encouraging students to rely on theoretical knowledge rather than empirical data. The adjustments made between Cycle 1 and Cycle 2 were intended to provide clearer ends-in-view, directing students' actions more explicitly towards the ultimate purpose of the lesson (sensu Johansson & Wickman, 2011). An alternative design, allowing students more time to discuss and make sense of the question, consider what data would be needed to answer it and which scenarios to explore in the simulation to generate such data, might have given data a more explicit role in constructing and justifying claims. Another possibility would be to let the students formulate their own questions and theory-based hypotheses which they could test by collecting and interpreting data. Thus, the limited use of data in students' written responses may reflect not only difficulties in reasoning but also how the inquiry activity

was organized. It would be interesting to explore how the DIEC model could be used in such an inquiry design.

Considering this, the students in this study may have needed even more explicit support in recognizing the role of information derived from data within the epistemic context of the task – that is, not only to observe and interpret, but also to use such interpretations as a basis for argument. This aligns with research (Chinn & Iordanou, 2023; Gott & Duggan, 1996; Sandoval, 2003), suggesting that students need more explicit scaffolding in understanding the relevance of data interpretation, identification of patterns, and changes in data over time. Gericke et al. (2022) similarly emphasise the importance of integrating argumentation, writing, and communication throughout the inquiry activity, rather than confining these activities to the final reporting phase. This includes helping students attend not only to the final values but also to variation, trends, and relationships in datasets. It also points to the importance of classroom practices in which the interpretive work of transforming data into information is made visible and discussed. Seda Cetin et al. (2018) illustrate how structured argumentation and peer review can provide opportunities for students to critically examine and revise their arguments. Similarly, Kuhn and Lerman (2021) argue that such discussions require explicit support targeting the cognitive challenges faced by individual students when relating evidence to claims. They demonstrate that even brief activities in which students critically evaluate and discuss different forms of evidence can support this process.

The results suggest that the DIEC model had limited impact as a short-term instructional scaffold for students. Despite being introduced and used in the lesson design, its use did not substantially change how students used data and information in their written claims. It was evident that the steps of the model did not serve as conceptual tools for the students, they neither referred to them in their informal talk nor used them to guide their actions when producing a written claim. It needs therefore to be reconsidered how the model can be used together with students. At the same time, the model was useful as a tool for designing and analysing teaching. It provided a framework for the teachers and researchers to communicate terms and concepts of scientific reasoning. This enabled a systematic examination of both written responses and classroom interaction and made it possible to identify forms of data transformation that might otherwise remain less visible. The model can thus be used by educators to examine, understand, evaluate, and assess different qualities in how students formulate claims. In this sense, the model may provide a conceptual resource for both researchers and teachers.

In summary, the DIEC model allowed us to identify the discrepancy between process (students' discussions) and product (their written claims) regarding the role of data in their reasoning. We therefore suggest that the DIEC model can provide, (a) a model for teachers and researchers to communicate terms and concepts of scientific reasoning when planning a series of lessons, and (b) a tool to analyse student work such as laboratory reports, discussions and presentations. However, more studies are needed to test its usefulness. It would for example be interesting to explore which purposes that need to be introduced in order to make the data transformation step meaningful and interesting for

the students to engage with. Other aspects worth exploring are long-time studies of how to use the DIEC model together with students. McNeill (2011), amongst others, showed that learning to formulate scientifically grounded answers is a skill that needs time to develop. Moreover, according to the study, the students in Cycle 1 had little or no understanding of the key concepts *Data, Information, Evidence, Claim* nor the relationship between them although the model was introduced to them prior to the task. However, if these concepts were introduced to the students in a more explicit way together and reiterating over a longer period of time together with faded scaffolding, we think they could help them develop a more scientific way of reasoning. This would be of interest for future research. We therefore recommend future studies on different aspects of the DIEC model. Finally, by using the DIEC rubric with students before, during and after any laboratory task, the rubric may help students to evaluate their written reports, pointing to the importance of using both data and information as well as theory when making evidence for a claim. Repeated use of the DIEC rubric together with students may clarify and teach students the importance of explicit use of relevant data. Future research could then be focused on a study over time and analysing the progress when using the rubric together with students.

Research ethics

Author contributions

A.P.: conceptualization, methodology, funding acquisition, investigation, data curation, formal analysis, writing—original draft preparation, writing—review and editing
B.A.: investigation, data curation, formal analysis, writing—original draft preparation, writing—review and editing
B.K.: investigation, data curation, reviewing and editing
K.L.: investigation, data curation, reviewing and editing
Lj.J.: investigation, data curation, formal analysis, writing—original draft preparation, writing—review and editing
L.J.: investigation, writing—original draft preparation, writing—review and editing
W.H.M.: investigation, formal analysis, writing—original draft preparation, writing—review and editing

All authors have read and agreed to the published version of the manuscript.

Artificial intelligence

Artificial Intelligence was used for first draft transcriptions of audio recordings, in the form of Automatic Speech Recognition and voice diarisation software as described in section 4.2. Some of the transcript excerpts were translated from Swedish to English using generative AI and then checked manually. There was no other use of AI in the study or authoring of this paper.

Funding

The research was funded through Stockholm Teaching & Learning Studies (STLS) by participating school authorities in collaboration with Stockholm University.

Informed consent statement

Informed consent was obtained from all research participants.

Data availability statement

Due to ethical restrictions, data from this research is not available.

Conflicts of Interest

The authors declare no conflicts of interest

References

AAAS. (1993). *Benchmarks for science literacy*. Oxford University Press.

Allchin, D., & Zemplén, G. Á. (2020). Finding the place of argumentation in science education: Epistemics and Whole Science. *Science Education, 104*(5), 907–933. https://doi.org/10.1002/sce.21589

Bain, M., Huh, J., Han, T., & Zisserman, A. (2023). WhisperX: Time-Accurate Speech Transcription of Long-Form Audio. INTERSPEECH 2023.

Bell, P., & Linn, M. C. (2000). Scientific arguments as learning artifacts: Designing for learning from the web with KIE. *International Journal of Science Education*, *22*(8), 797–817. https://doi.org/10.1080/095006900412284

Bredin, H., Yin, R., Coria, J. M., Gelly, G., Korshunov, P., Lavechin, M., Fustes, D., Titeux, H., Bouaziz, W., & Gill, M.-P. (2020). Pyannote.Audio: Neural Building Blocks for Speaker Diarization. ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 7124–7128. https://doi.org/10.1109/ICASSP40776.2020.9052974

Bredin, H., & Laurent, A. (2021). End-to-end speaker segmentation for overlap-aware resegmentation (No. arXiv:2104.04045). arXiv. https://doi.org/10.48550/arXiv.2104.04045

Brown, N. J. S., Furtak, E. M., Timms, M., Nagashima, S. O., & Wilson, M. (2010). The evidence-based reasoning framework: Assessing scientific reasoning. *Educational Assessment*, *15*(3–4), 123–141. https://doi.org/10.1080/10627197.2010.530551

Cavagnetto, A. R. (2010). Argument to Foster Scientific Literacy: A Review of Argument Interventions in K–12 Science Contexts. *Review of Educational Research, 80*(3), 336–371. https://doi.org/10.3102/0034654310376953

Chen, J. J., Lin, H. S., Hsu, Y. S., & Lee, H. (2011). Data and claim: The refinement of science fair work through argumentation. *International Journal of Science Education, Part B*, *1*(2), 147–164. https://doi.org/10.1080/21548455.2011.582707

Chinn, C. A., & Iordanou, K. (2023). Theories of learning. In N. G. Lederman, D. L. Zeidler, & J. S. Lederman (Eds), *Handbook of Research on Science Education: Volume III* (1st edn, pp. 89–120). Routledge. https://doi.org/10.4324/9780367855758

Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education*, *84*(3), 287–312. https://doi.org/10.1002/(SICI)1098-237X(200005)84:3<287::AID-SCE1>3.0.CO;2-A

Duschl, R. A., & Osborne, J. (2002). Supporting and promoting argumentation discourse in science education. *Studies in Science Education*, *38*(1), 39–72. https://doi.org/10.1080/03057260208560187

Engelmann, K. F., Chinn, C. A., Osborne, J., & Fischer, F. (2018). The Roles of Domain-Specific and Domain-General Knowledge in Scientific Reasoning and Argumentation. An Introduction. In F. Fischer, C. A. Chinn, K. Engelmann, & J. Osborne (Eds), *Scientific Reasoning and Argumentation: The Roles of Domain-Specific and Domain-General Knowledge* (1st edn, pp. 1–8). Routledge. https://doi.org/10.4324/9780203731826

Erduran, S. (2013). Cocktail of Food Science and Argumentation: Shaken or Stirred for Learning?. *LUMAT: International Journal on Math, Science and Technology Education, 1*(2), 151–158. https://doi.org/10.31129/lumat.v1i2.1110

Ergunt, E., & Yilmaz, S. (2026). Argumentation Research in Science Education: Global Publication Trends, Intellectual Structure, and Thematic Transformation (2001–2025). *Journal of Education in Science, Environment and Health, 12*(2), 137–159. https://doi.org/10.55549/jeseh.894

Furtak, E. M., Hardy, I., Beinbrech, C., Shavelson, R. J., & Shemwell, J. T. (2010). A framework for analyzing evidence-based reasoning in science classroom discourse. *Educational Assessment, 15*(3–4), 175–196. https://doi.org/10.1080/10627197.2010.530553

Gericke, N., Högström, P., & Wallin, J. (2022). A systematic review of research on laboratory work in secondary school science education. *Studies in Science Education*, *58*(1), 1–34. https://doi.org/10.1080/03057267.2022.2090125

Gott, R., & Duggan, S. (1996). Practical work: Its role in the understanding of evidence in science. *International Journal of Science Education*, 18(7), 791–806. https://doi.org/10.1080/0950069960180705

Governor, D., Lombardi, D., & Duffield, C. (2021). Negotiations in scientific argumentation: An interpersonal analysis. *Journal of Research in Science Teaching*, *58*(9), 1389–1424. https://doi.org/10.1002/tea.21713

Gyllenpalm, J., Rundgren, C.-J., Lederman, J., & Lederman, N. G. (2021). Views about scientific inquiry: A study of students' understanding of scientific inquiry in grade 7 and 12 in Sweden. *Scandinavian Journal of Educational Research*, *66*(2), 336–354. https://doi.org/10.1080/00313831.2020.1869080

Hand, B., Wallace, C. W., & Yang, E. (2004). Using a Science Writing Heuristic to enhance learning outcomes from laboratory activities in seventh-grade science: Quantitative and qualitative aspects. *International Journal of Science Education*, *26*(2), 131–149. https://doi.org/10.1080/0950069032000070252

Hardy, I., Kloetzer, B., Moeller, K., & Sodian, B. (2010). The Analysis of Classroom Discourse: Elementary School Science Curricula Advancing Reasoning With Evidence. *Educational Assessment*, *15*(3–4), 197–221. https://doi.org/10.1080/10627197.2010.530556

Hsieh, Hsiu-Fang, & Shannon, Sarah E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, *15*(9), 1277-1288. https://doi.org/10.1177/1049732305276687

Johansson, A.-M., & Wickman, P.-O. (2011). A pragmatist understanding of learning progressions. In B. Hudson & M. A. Meyer (Eds.), *Beyond fragmentation: Didactics, learning and teaching* (pp. 47-59). Leverkusen: Barbara Budrich Publishers.

Kan, Z., Tan, A.-L., Pei, X., Wu, Q., & Yao, X. (2026). Exploring science teachers' epistemic beliefs about scientific inquiry and influencing factors. *International Journal of Science Education, 48*(6), 840–863. https://doi.org/10.1080/09500693.2024.2428837

Kuhn, D. (1993). Science as argument: Implications for teaching and learning scientific thinking. *Science Education*, *77*(3), 319–337. https://doi.org/10.1002/sce.3730770306

Kuhn, D., & Lerman, D. (2021). Yes but: Developing a critical stance toward evidence. *International Journal of Science Education, 43*(7), 1036–1053. https://doi.org/10.1080/09500693.2021.1897897

Kuhn, D., & Modrek, A. S. (2022). Choose Your Evidence: Scientific Thinking Where It May Most Count. *Science & Education, 31*(1), 21–31. https://doi.org/10.1007/s11191-021-00209-y

Lee, S. C., & Irving, K. E. (2018). Development of two-dimensional classroom discourse analysis tool (CDAT): Scientific reasoning and dialog patterns in the secondary science classes. *International Journal of STEM Education*, *5*(1), 1–17. https://doi.org/10.1186/s40594-018-0100-0

Malmsten, M., Haffenden, C., & Börjeson, L. (2022). Hearing voices at the national library--a speech corpus and acoustic model for the Swedish language. *arXiv preprint arXiv:2205.03026*.

Manz, E., Lehrer, R., & Schauble, L. (2020). Rethinking the classroom science investigation. J*ournal of Research in Science Teaching, 57*(7), 1148–1174. https://doi.org/10.1002/tea.21625

McNeill, K. L. (2011). Elementary students' views of explanation, argumentation, and evidence, and their abilities to construct arguments over the school year. *Journal of Research in Science Teaching*, *48*(7), 793–823. https://doi.org/10.1002/tea.20430

McNeill, K. L., Lizotte, D. J., Krajcik, J., & Marx, R. W. (2006). Supporting Students' Construction of Scientific Explanations by Fading Scaffolds in Instructional Materials. *Journal of the Learning Sciences*, *15*(2), 153–191. https://doi.org/10.1207/s15327809jls1502_1

McNeill, K. L., & Krajcik, J. (2008). Inquiry and scientific explanations: Helping students use evidence and reasoning. In J. Luft, R. L. Bell, & J. Gess-Newsome (Eds.), *Science as inquiry in the secondary setting* (pp. 121–134). NSTA Press.

McNeill, K. L., & Martin, R. A. (2011). Claims, evidence, and reasoning in science education. *Science and Children, 48*(8), 52–56.

Mesci, G., & Erdaş Kartal, E. (2021). Science Teachers' Views on Nature of Scientific Inquiry. *Bartın University Journal of Faculty of Education, 2021 February, Volume 10*(Issue 1), 69–84. https://doi.org/10.14686/buefad.797246

National Research Council. (2012). A framework for K-12 science education: Practices, crosscutting concepts, and core ideas. Washington, DC: The National Academies Press.

Newton, P., Driver, R., & Osborne, J. (1999). The place of argumentation in the pedagogy of school science. *International Journal of Science Education*, *21*(5), 553–576. https://doi.org/10.1080/095006999290570

Novak, A. M., & Treagust, D. F. (2018). Adjusting claims as new evidence emerges: Do students incorporate new evidence into their scientific explanations? *Journal of Research in Science Teaching, 55*(4), 526–549. https://doi.org/10.1002/tea.21429

Novak, A. M., & Treagust, D. F. (2022). Supporting the development of scientific understanding when constructing an evolving explanation. *Disciplinary and Interdisciplinary Science Education Research, 4*(1), 3. https://doi.org/10.1186/s43031-021-00043-w

OECD. (2006). *Assessing scientific, reading and mathematical literacy: A framework for PISA 2006*. OECD Publishing.

Osborne, J. (2013). The 21st century challenge for science education: Assessing scientific reasoning. *Thinking Skills and Creativity*, 10, 265–279. https://doi.org/10.1016/j.tsc.2013.07.006

Osborne, J., Erduran, S., & Simon, S. (2004). Enhancing the quality of argumentation in school science. Journal of Research in Science Teaching, *41*(10), 994–1020. https://doi.org/10.1002/tea.20035

PhET Interactive Simulations. (n.d.). *Natural selection*. University of Colorado Boulder. https://phet.colorado.edu/en/simulations/natural-selection

Radford, A., Kim, J. W., Xu, T., Brockman, G., Mcleavey, C., & Sutskever, I. (2023). Robust Speech Recognition via Large-Scale Weak Supervision. I A. Krause, E. Brunskill, K. Cho, B. Engelhardt, S. Sabato, & J. Scarlett (Red.), Proceedings of the 40th International Conference on Machine Learning (Vol. 202, s. 28492–28518). PMLR. https://proceedings.mlr.press/v202/radford23a.html

Ryu, S., & Sandoval, W. A. (2012). Improvements to elementary children's epistemic understanding from sustained argumentation. *Science Education*, *96*(3), 488–526. https://doi.org/10.1002/sce.21006

Sampson, V., & Grooms, J. (2009). Promoting and supporting scientific argumentation in the classroom: The evaluate-alternatives instructional model. *Science Scope, 33*(1), 66–73. http://www.jstor.org/stable/43183907

Sampson, V., Grooms, J., & Walker, J. P. (2011). Argument-Driven Inquiry as a way to help students learn how to participate in scientific argumentation and craft written arguments: An exploratory study. *Science Education*, *95*(2), 217-257. https://doi.org/10.1002/sce.20421

Sandoval, W. A. (2003). Conceptual and epistemic aspects of students' scientific explanations. *Journal of the Learning Sciences*, *12*(1), 5–51. https://doi.org/10.1207/S15327809JLS1201_2

Samarapungavan, A. (2018). Constructing scientific evidence. The Role of Disciplinary Knowledge in Reasoning with and about Evidence in Scientific Practice. In F. Fischer, C. A. Chinn, K. F. Engelmann, & J. Osborne (Eds), *Scientific reasoning and argumentation: The roles of domain-specific and domain-general knowledge* (pp. 56–76). Routledge, Taylor & Francis Group. https://doi.org/10.4324/9780203731826

Schildkamp, K., & Kuiper, W. (2010). Data-informed curriculum reform: Which data, what purposes, and promoting and hindering factors. *Teaching and teacher education, 26*(3), 482-496. https://doi.org/10.1016/j.tate.2009.06.007

Schwartz, R. S., Lederman, J. S., & Enderle, P. J. (2023). Scientific Inquiry Literacy. The Missing Link on the Continuum From Science Literacy to Scientific Literacy. In N. G. Lederman, D. L. Zeidler, & J. S. Lederman (Eds), Handbook of Research on Science Education: Volume III (1st edn, pp. 749–782). Routledge. https://doi.org/10.4324/9780367855758

Seda Cetin, P., Eymur, G., Southerland, S. A., Walker, J., & Whittington, K. (2018). Exploring the effectiveness of engagement in a broad range of disciplinary practices on learning of Turkish high-school chemistry students. *International Journal of Science Education*, *40*(5), 473–497. https://doi.org/10.1080/09500693.2018.1432914

Songsil, W., Pongsophon, P., Boonsoong, B., & Clarke, A. (2019). Developing scientific argumentation strategies using revised argument-driven inquiry (rADI) in science classrooms in Thailand. *Asia-Pacific Science Education*, *5*(1), 1-22. https://doi.org/10.1186/s41029-019-0035-x

Swedish National Agency for Education. (2022). Läroplan för grundskolan, förskoleklassen och fritidshemmet [Curriculum for the compulsory school, preschool class and school-age educare.]

Swedish Research Council (2017). God forskningssed [Good Research Practice]. Stockholm: Vetenskapsrådet [Swedish Research Council].

Tosun, C. (2024). Analysis of the Last 40 Years of Science Education Research via Bibliometric Methods. *Science & Education, 33*(2), 451–480. https://doi.org/10.1007/s11191-022-00400-9

Toulmin, S. (1958). *The Uses of Argument*. Cambridge University Press.

Weiss, K. A., McDermott, M. A., & Hand, B. (2022). Characterising immersive argument-based inquiry learning environments in school-based education: A systematic literature review. *Studies in Science Education*, *58*(1), 15–47. https://doi.org/10.1080/03057267.2021.1897931