MODESTUM

Research Paper

Physics Pre-service Teachers' Approaches to Scientific Investigations by Data Exploration

Thomas Schubatzky [1*], Benjamin Bock [1], Claudia Haagen-Schützenhöfer [1]

[1] University of Graz, AUSTRIA

Abstract

This article reports how physics pre-service teachers (PSTs) organize their investigations during an exploratory data analysis scenario, which we call scientific investigations by data exploration. In order to analyze the PSTs' investigations, we developed a learning environment in which learners investigate aspects influencing the particulate matter concentration in an Austrian city. Audio documentation and written learner protocols were analyzed using qualitative content analyses, resulting in flowcharts describing the different types of investigations the PSTs conducted. In this analysis, we differentiate between investigations on a micro-level (a single investigation), and investigations on a macro-level. Findings show that the pre-service teachers follow three different approaches: some always start their investigations with a research question, some switch between exploratory and targeted investigations and a few conducted only exploratory investigations. In this article we provide exploratory insights into the strategies students use.

Keywords: pre-service teachers, data exploration, scientific investigations, exploratory data analysis, TinkerPlots

INTRODUCTION

Due to the increasing importance of information technology, the 21st century is also referred to as the information or data age. Large amounts of data and their processing play an important role in optimizing processes and systems but also in gaining new knowledge, for example through data mining. A call for change, adaptation or further development of the affected systems is often connected to these new opportunities. Just to name two examples, such changes take place in systems like in real-time influenza forecasting (Shaman, Karspeck, Yang, Tamerius, & Lipsitch, 2013) or in marketing via the analysis of data on shopping behavior.

Digitalization helps to collect large amounts of data more easily than ever before. This development has also reached scientific research. In many areas of science, large amounts of data enable new research branches. For example, the global weather forecast model ICON alone requires 14 different atmospheric measurements (temperature, air pressure, …) from a total of 2,949,120 lattice points on the surface of the earth in 89 different heights (Reinert, Prill, Frank, Zängl, & Denhard, 2018). All in all, the atmosphere of the earth is mapped for ICON by a total of $3.7 \cdot 10^9$ individual meteorological data. It is hard to imagine that current research in physics could do without large data bases. Additionally, many research groups and governmental agencies make their research or environmental data accessible via online data repositories or by providing data at official homepages. For example, this can be data from an experiment conducted during research (e.g. from CERN), meteorological data like temperature or humidity but also environmental data like air pollution (particulate matter, NOX, etc.).

However, there are only a few examples (e.g. Irish, Berkowitz, & Harris, 2019; Makar & Confrey, 2014) of how data exploration techniques using authentic environmental data found their way into science education. To tackle this shortcoming, we developed a learning environment that combines the use of large amounts of data from an online data repository with the application of the innovative software TinkerPlots

Contribution to the literature

- We developed a theoretical framework for scientific investigations based on already existing frameworks. Our QAIC-cycle allows to investigate students' approaches to exploratory data analysis using already existing real-world data.
- The usage of innovative software and authentic data from repositories allows to analyze single investigations performed by students as well as the sequencing of multiple investigations and how students organize them.
- The findings show that in our exploratory data analysis scenario, some students only perform targeted investigations, while other students switch between targeted and exploratory investigations.

(Konold & Miller, 2005). We have chosen a context-oriented approach, using a real-world scenario that is relevant to students. There, students investigate authentic data (temperature, humidity, particulate matter concentration,…; over 35000 measurements) collected by different meteorological stations in an Austrian city to figure out which factors influence particulate matter concentration (Schubatzky & Haagen-Schützenhöfer, 2019). The goal of this learning environment is to foster students' skills in carrying out scientific correlational investigations, in particular exploratory data analysis. In this article, we report findings about learning processes related to scientific investigations by data exploration. Our focus is on the way how students organize their investigations.

We analyzed the students' approaches on two different levels. The micro-level analysis (in this article referred to as "investigations") focuses on how students organize and approach a single investigation (as operationalized in the QAIC-cycle). In order to analyze the individual phases of a single investigation, we synthesized aspects from different inquiry models and developed the "QAIC-cycle". The macro-level analysis (referred to as "investigation processes") focuses on how students organize multiple related micro-level investigations. In the next section, the theoretical framework of our analysis is described.

THEORETICAL FRAMEWORK AND QAIC-CYCLE

In order to analyze the various micro-level investigations of our sample we synthesized the QAIC-cycle from already existing inquiry models. The QAIC-cycle mirrors the process of scientific investigations by data exploration and describes the relevant phases and steps students can encounter during their investigations throughout our intervention. We argue that such investigations differ from non-science-related exploratory data analysis as well as from scientific investigation without exploratory data analysis.

The QAIC-cycle is based on preliminary models. First, it is partly based on a scientific inquiry competence model developed by Arnold, Boone, Kremer and Mayer (2018) where scientific inquiry is conceptualized as a problem-solving process. Although Arnold et al. (2018) emphasize that the scientific problem-solving process is not a linear process, they identified three key components or sub-competencies specifically for experimentation. These three are:

- Students should be able to formulate scientific questions/generate testable hypotheses.
- Students should be able to design an experiment.
- Students should be able to analyze data.

Additionally, it is based on the PPDAC-cycle developed by Wild and Pfannkuch (1999). This Problem-Plan-Data-Analysis-Conclusion inquiry cycle stems from statistics education and covers all the steps involved when conducting statistical inquiry. The idea of the Problem stage of the PPDAC-cycle is to turn vague ideas of the problem into specific questions, which can be answered using data (Utts & Horton, 2018). The Data step is about obtaining the data, storing and cleaning it. Subsequently, the Analysis and the Conclusion step are about making sense of the data, abstracting and communicating what has been deduced from the data.

Although both models contain relevant elements for the kind of investigation processes we want to describe, none was able to depict the full range of intended phases, steps and learning processes in our intervention. Consequently, we had to merge parts of these models.

In the next four subsections, we describe how we synthesized the main steps in the four phases of the QAIC-cycle and highlight how aspects of the afore mentioned inquiry models contributed to the genesis of the QAIC-cycle. We structure the following subsection according to the four phases within the QAIC-cycle as shown in Figure 1.

Phase One: Generation of Questions & Hypothesis

The operationalization of this first phase in the QAIC-cycle differs slightly from the existing models. Based on them, we have identified three relevant steps for the first phase of scientific investigations by data exploration.

- Students generate research questions. These questions are related to the context and investigable with the data at hand.
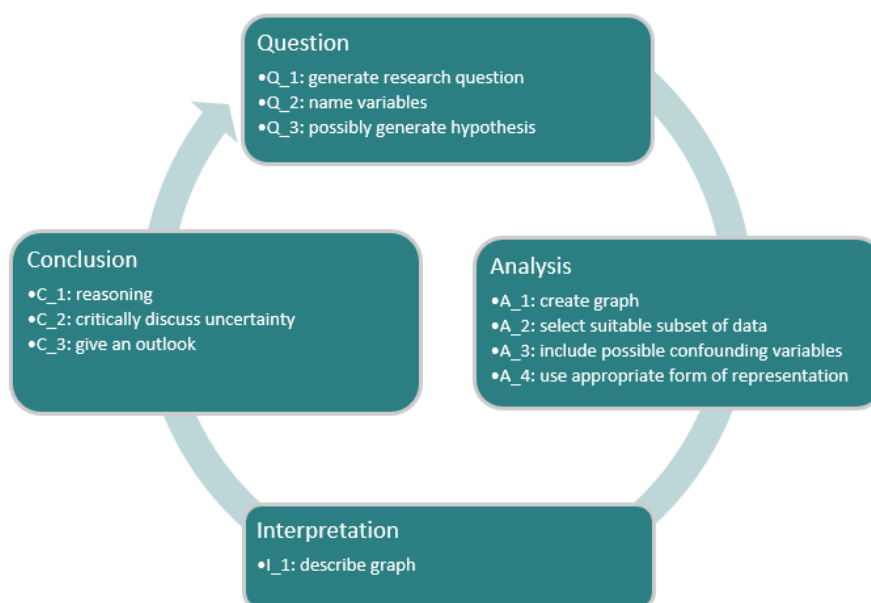
Figure 1. QAIC-cycle (Question – Analysis – Interpretation – Conclusion) as theoretical framework for student's approaches to single investigations by data exploration using already existing data

- Students identify the relevant variables, both dependent and independent.
- Students generate and justify a hypothesis in the form of a presumption regarding the expected results. (optional)

One of the three formulated sub-competencies of scientific inquiry competence after Arnold et al. (2018) is operationalized as "formulate scientific questions" and, if necessary, generating a testable related hypothesis. The dependent and the independent variables are defined as "part of the research question". These steps were integrated in the QAIC-model.

In the PPDAC-cycle, the first step is to conceive the problem of investigation and to generate a research question. In this first process step – plan – of the PPDAC cycle, one determines the essential properties of the system under investigation with regard to the research question, which also refers to the third aspect of the step question within the QAIC-cycle (Pfannkuch, 1999). It shall be mentioned that these variables are not measurands, but properties or factors of the problem itself. This distinction will be further discussed in the next section. As Lederman, Lederman and Antink (2013) however argue, a hypothesis is not necessarily required for all scientific investigations and is therefore optional. Similar to that, the first step of the PPDAC-cycle does not address the generation of a hypothesis. In accordance with that, we think that a hypothesis is an optional part of scientific investigations by data exploration.

Phase Two: Analysis

Regarding the phase analysis, we synthesized three different steps which are relevant for a meaningful analysis of data in an exploratory data analysis scenario:

- Students present the data appropriately.
- Students use special techniques of exploratory data analysis, like the transformation and variation of graphs.
- Students carefully select relevant subsets of data to answer their research questions while also considering relevant confounding variables.

In our learning environments on scientific investigations by data exploration we provide already existing data from repositories or research institutions. Hence, the next phase in the QAIC-cycle is to analyze the existing data by creating suitable representations.

In our learning environment this means the selection of relevant subsets of data. Furthermore, the distinction between variables and measurands is important, since one variable can be operationalized by different measurands. For example, one may ask: "Did the summers in my city became hotter in the last decades?". In this research question, the temperature of a city during summer is the dependent variable. However, one may use the daily mean temperature to investigate this question, or the number of very hot days. These two different approaches may also lead to different conclusions about the same research question.

Further aspects mentioned in the model of scientific inquiry competence are "control confounding variables" and "define replications" (Arnold et al., 2018). In the QAIC-cycle, these types of thinking processes are included in the phase of the creation of suitable graphs.

Phase Three: Interpretation

After the creation of appropriate graphs, students need to interpret them. We understand the term

"interpretation" as the formulation of statements about the data. Therefore, the relevant step of this phase is:

- Students describe graphs and the data presented therein.

It is about the description of patterns, structures and abnormalities of the graph (Tukey, 1980). In terms of exploratory data analysis, such interpretations do not go beyond the available data. If one understands interpretation as such, it does not correspond to "interpretation" in the model of scientific inquiry competence (Arnold et al., 2018). There, interpretation is described as interpreting the data in terms of the hypothesis. In the QAIC-cycle the phase of interpretation has been separated from data analysis, because a created graph represents a result of the "analysis" phase.

The phases "interpretation" and "conclusion" are treated separately in the last step of the PPDAC-cycle (called "conclusion") (Pfannkuch, 1999). However, a graph can lead to multiple interpretations and each interpretation can lead to multiple different conclusions. Thus, these distinctions between data analysis, interpretation and conclusion are necessary to appropriately describe the investigation processes of the students.

Phase Four: Conclusion

The last phase of the investigation is to draw conclusions which base on the interpretation of the graphs. The relevant steps of the conclusion phase in the QAIC-cycle are:

- Students draw a conclusion based on empirical evidence which relates to the research question.
- Students justify their conclusions explicitly, for example by referring to contextual knowledge or the interpretation of the graph.

- Students discuss the uncertainty of their conclusions.
- Students possibly give an outlook to further investigations.

This describes what the QAIC-cycle addresses as 'drawing a conclusion'. According to the concept of informal inferential reasoning (Ben-Zvi, 2005; Garfield et al., 2008; Makar, Bakker, & Ben-Zvi, 2011), to draw a conclusion means to integrate contextual knowledge with the interpretation of the graph and make a claim beyond the data. A conclusion correlates with the context and the problem under investigation, especially the research question. The processing of the interpretations of a graph together with other, consciously included, knowledge elements is understood as a conclusion. As we operationalize this phase, additional knowledge elements can stem from contextual knowledge, but also from knowledge about physics or statistical ideas like location, variation and tendency.

If based on empirical observations or measurements, conclusions should only be expressed with a considered degree of uncertainty. This aspect is mentioned in the model of scientific inquiry competence (Arnold et al., 2018) as well as in the concept of informal inferential reasoning (Makar et al., 2011). Additionally, each conclusion should consider data aggregates and should not be based on individual data points.

Furthermore, the methodology should be critically reflected and an outlook on the following investigations should be given.

Table 1 lists the different types of thinking and reasoning occurring during the QAIC-cycle. After describing the theoretical basis of our learning environment, the overall design will be described shortly.

Table 1. Phases and steps during the QAIC-cycle

| Phase	step	description
Question	Students…	
	generate a research question	*… generate a research question which they want to investigate based on the available data.*
	name variables	*… name the dependent and independent variable.*
	generate a hypothesis	*… generate a hypothesis in the form of a presumption.*
Analysis	Students…	
	create a graph	*… create a graph to help answer the research questions.*
	select a suitable subset of data	*… select a suitable subset of the dataset.*
	include possible confounding variables	*… include possible confounding variables.*
	use appropriate forms of representation	*… use appropriate forms of representations.*
Interpretation	Students…	
	Describe a graph	*… describe a graph or the measured data presented therein.*
Conclusion	Students…	
	conclude	*… draw conclusions based on empirical evidence which relate to the research question.*
	reason	*… justify their conclusion.*
	include uncertainty	*… critically discuss uncertainties of their conclusion.*
	give an outlook	*… give an outlook to further investigations.*
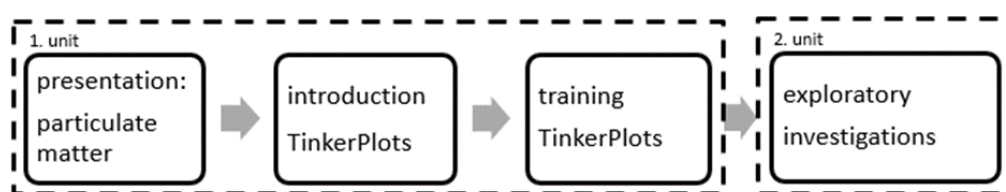
Figure 2. Sequence of the learning environment

DESCRIPTION OF THE LEARNING ENVIRONMENT

In this section we introduce the learning environment which was based on the theoretical ideas discussed in the previous section. In our learning environment physics teacher-students are trained to carry out scientific investigations by data exploration using innovative software like TinkerPlots. In addition, they learn how to elementarize and transfer it to school settings.

The sequencing of the learning environment relevant for this article on the level of physics teacher-students can be seen in Figure 2. The intervention is implemented in two units of a weekly course, which lasts 120 minutes each. In the first unit, students are introduced to the topic of particulate matter in general, especially its sources, influencing factors and about its dispersal. We put a special focus on the situation of our university town Graz, the second largest Austrian city. We discuss the special circumstances of Graz and why particulate matter poses a problem to the citizens.

After this initial introduction, students get introduced to the software TinkerPlots. During this phase, they work with training data from a database about ozone pollution in New York (sample database provided by TinkerPlots) to get familiar with the software.

At the beginning of the second unit, after a short recap of the first unit, we make the physics teacher-students familiar with particulate matter data collected in Graz and introduce the assignment. Additionally, we provide students with scaffolding material regarding contextual knowledge about particulate matter for scaffolding purposes. The physics teacher-students' task is to investigate which variables influence the particulate matter concentration. However, they can choose how to approach this task and which variables they want to investigate. The dataset we use for this assignment originates from an online data repository (the data by the government of Styria is licensed under CC BY 3.0 AT) and contains the following measurements: particulate matter concentration, air temperature, humidity and other variables. For a detailed description of the used dataset and results on an operational level see Schubatzky & Haagen-Schützenhöfer (2019).

After giving an overview of the learning environment, the next section clarifies which research

goals we follow on an educational level when analyzing physics teacher-students' investigations.

PURPOSE OF THIS STUDY

This study aims at exploring physics teacher-students' approaches to scientific investigations by data exploration using real-world data. Due to the use of an innovative software like TinkerPlots, this setting even allows to investigate how students coordinate multiple investigations since the students can perform several investigations to undertake the task in a rather short amount of time.

Following these ideas, we formulated two research questions that guided our study:

(1) What specific procedures do physics teacher-students follow during their investigations?

(2) How do physics teacher-students organize their investigations in an exploratory data analysis scenario?

SAMPLE

Our study was conducted as part of a physics pre-service teacher seminar at the University of Graz. The overarching theme of this seminar was digital technology in physics education. The seminar was organized in two seminar groups with 16 physics teacher-students each. Our sample consisted of 32 pre-service secondary school physics teachers (21 male, 11 female) with an average of 6.83 ± 1.36 semesters in their physics teacher education, which is an integrated study of physics, didactics and pedagogy. In Austrian teacher preparation programs, all students need to study two subjects with an overall duration of 8 bachelor-semesters and 4 master-semesters. It must be noted that 13 out of the 32 participants studied Math as their second subject. 17 participants reported that they had never conducted a similar exploratory data analysis.

DATA COLLECTION

During the intervention, students worked in teams of two in order to promote discussions. Additionally, they were separated into two intervention groups consisting of 7 and 9 teams. We assigned a number (1 and 2) for the two intervention groups and colors for the individual teams. Hence, each team can be identified with a number-color combination (e.g. 01_grey stands for the team with the color grey in intervention group 1). The

team-color: 5.12.2018

attention:

- Please clearly describe what you did using whole sentences.
- Indicate the numbers of the graphs and facts from the fact sheet you refer to.

Investigation protocol: particulate matter concentration in Graz

Please check the relevant box:	☐ We want to investigate the following:	☐ We create/describe a graph Graph number: _____	☐ We communicate a (partial) result of the investigation
Please enter a consecutive numbering here.			

Figure 3. This is a section of the protocol students used during their investigation. Students were asked to start a new section for each action they performed during the intervention

data collection for this study was administered during the second unit of the learning environment as shown in Figure 2. While following their investigations, each team filled in a detailed protocol. For illustration, a section of such a protocol is shown in Figure 3. The teams were instructed to document each action they performed during their investigations in the blank space of the investigation protocol (see Figure 3). For every action, they were asked to use a new section of the protocol. Additionally, they were asked to tick one of the boxes shown in the upper-mid section of Figure 3 if their action performed corresponded to the preemptively given descriptions in the investigation protocol. In order to allow a chronological reconstruction of the teams' investigations, they were asked to number their documented actions. Additionally, each team was audio taped in order to allow a triangulation of the data.

DATA ANALYSIS

The physics teacher-students' protocols were analyzed using qualitative content analysis (Mayring, 2014). The protocols were divided into meaningful units and transcribed as such. Following the general content-analytical procedural model (Mayring, 2014), these units were summarized (simplified and possibly shortened) without changing their meaning or order. Subsequently,

these units were categorized according to the developed QAIC-cycle. When it was not possible to deduce the corresponding category from the text in the protocol alone, we further explicated the statements in the protocol. This explication process was supported by the audio documentation we had made during the intervention in the seminar. In the final step (structuring), the explicated units were categorized according to the developed QAIC-cycle.

The protocol of one team did not meet the requirements of this analysis process and was therefore excluded from it. In total, 312 meaningful units were categorized and double coded, reaching an accordance of Cohens Kappa = .87. A contingency table summarizing the categorization can be found in Table 2.

Each category represents one stage of the QAIC cycle, the category X refers to cases that could not be assigned to any category.

In order to represent the investigations of the physics teacher-students on the micro- and macro-level in a feasible way, flowcharts were created. They represent the sequencing of the single categories (phases) according to the phases of the investigation operationalized in the QAIC-cycle. Figure 4 shows an example flowchart representing the investigations of one team. In these flowcharts, the violet square "Q" indicates

Table 2. Contingency table of categorization according to the QAIC-cycle. The letters Q, A, I, C and X refer to the categories Question, Analysis, Interpretation, Conclusion and Other

		categorization: rater 1					
		Q	**A**	**I**	**C**	**X**	
categorization: rater 2	**Q**	43	3			4	50
	A	0	117	1		2	120
	I		3	71	12		86
	C			3	52		55
	X				1		1
		43	123	75	65	6	
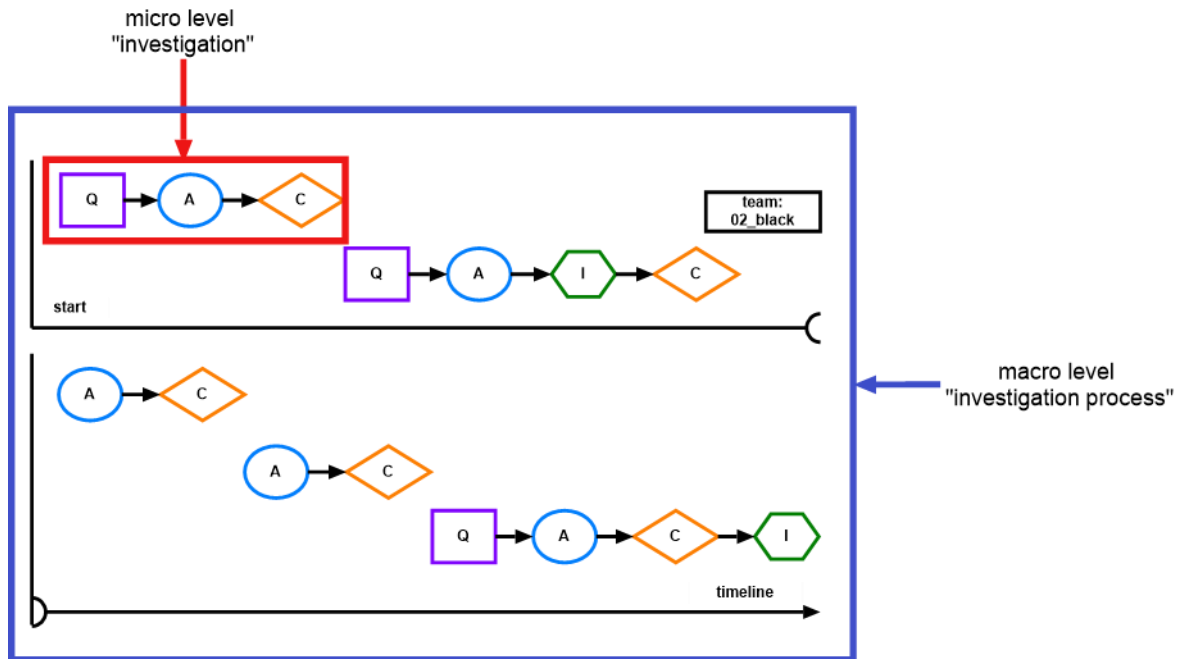
Figure 4. Flowchart of the reconstructed investigation process of team 02_black. An example of an investigation on the micro-level is indicated by a red frame. The blue frame shows an example of an investigation process on a macro-level. The symbols used (purple square, blue ellipse, green comb, orange diamond) represent different phases of the QAIC-cycle (Question – Analysis – Interpretation – Conclusion). These phases are described in the theoretical framework section. The black horizontal and vertical lines indicate the chronological order of the various micro-level investigations, starting from top left to bottom right

that a research question is proposed. The blue circle "A" indicates that a graph is created/modified, the green honeycomb "I" indicates an interpretation of a graph and an orange rhombus "C" indicates that a conclusion based on the graph or interpretation is drawn. The horizontal axis indicates the sequence of the phases as documented in the protocol, or to put it into other words, the timeline of the different phases. The arrows furthermore cluster the phases according to their topic of investigation. For example, the first three phases in Figure 4 refer to the investigation of the relationship between air temperature and particulate matter. Ultimately, these flowcharts allow to simultaneously represent the students' investigations on a micro-level (an example is framed in red in Figure 4), but also the full investigation process on the macro-level (framed in blue in Figure 4).

Additionally, some teams referred to graphs or findings from their previous micro-level investigations.

Such direct references are represented as a dotted line in the flow-charts, as shown in Figure 5.

RESULTS

In the next section, the results will be described on the micro-level first, followed by the results on the macro-level.

Micro-level Investigations

Based on the flowchart in which we translated the students' protocols, we can now treat the research questions of this article. As already mentioned, the categories for the analysis of the protocols had been derived from the QAIC-cycle (Figure 1). The ideal sequence of phases is given by the QAIC-cycle as it is represented in Figure 6. However, we want to emphasize that ideal in this case only means that it ideally represents the QAIC-cycle, not that this sequence is the
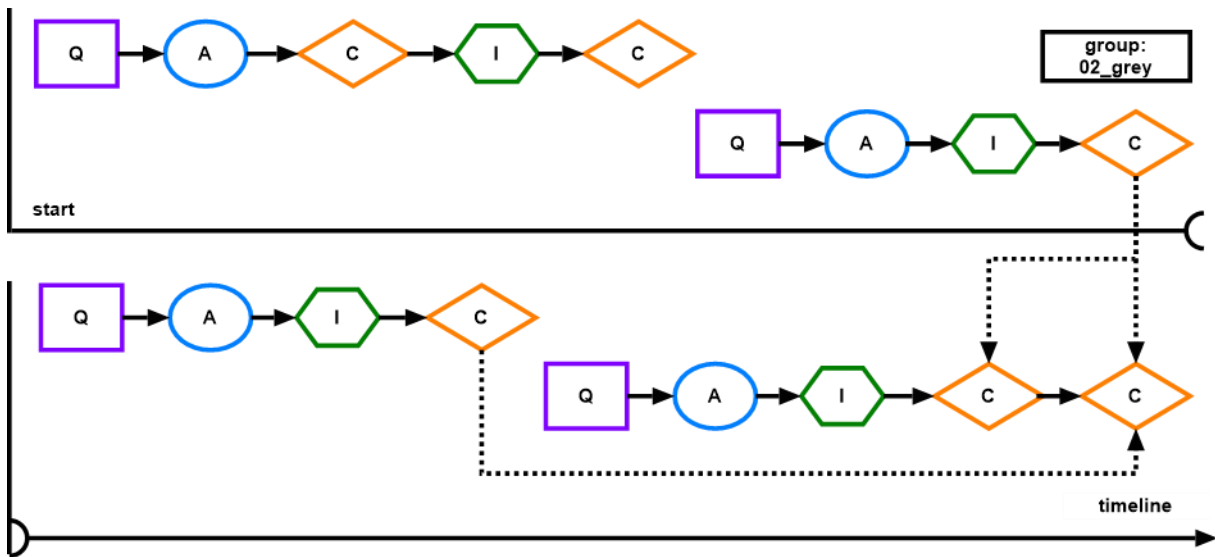
Figure 5. Flowchart of the reconstructed investigation process of team 02_grey. The symbols used (purple square, blue ellipse, green comb, orange diamond) represent the different phases of the QAIC-cycle (Question – Analysis – Interpretation – Conclusion) described in the theoretical framework section. Direct references to other micro-level investigations are represented using a dotted line. The black horizontal and vertical lines indicate the chronological order of the various micro-level investigations, starting from top left to bottom right
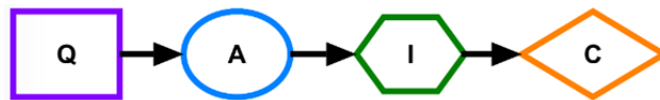

Figure 6. Sequence of phases of an ideal targeted investigation according to the QAIC-cycle. The symbols used (purple square, blue ellipse, green comb, orange diamond) represent the different phases of the QAIC-cycle (Question – Analysis – Interpretation – Conclusion) described in the theoretical framework section
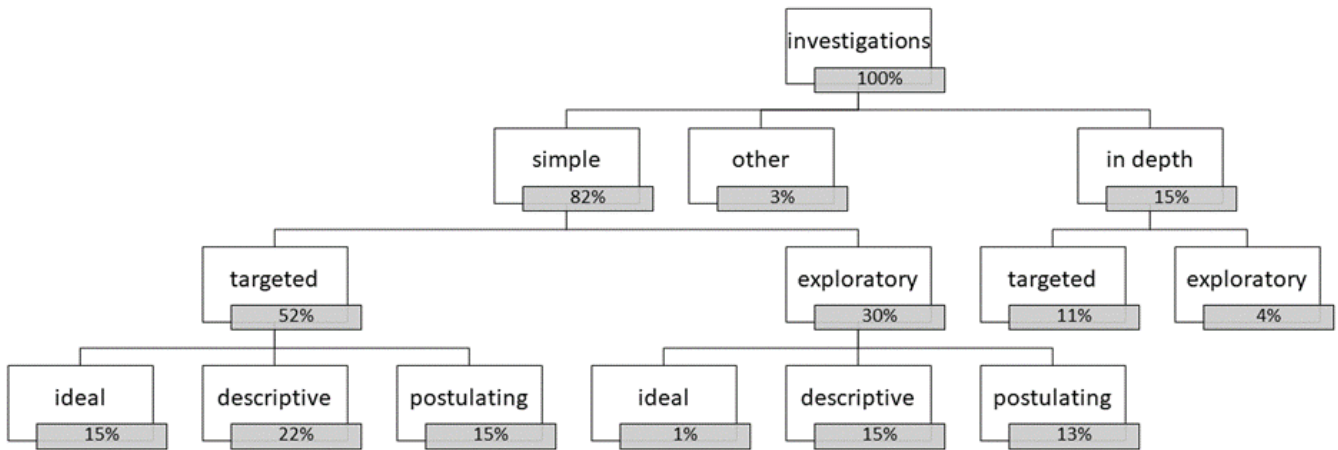

Figure 7. Tree diagram of the different types of investigations. The percentage-values refer to the percentage of all investigations

ideal set of phases involved in every type of data exploration.

The sequence of the different phases gives insight into the micro-level investigations of the teams. All investigations were analyzed and split up in different phases as described in the QAIC-cycle. We divided the investigations into "simple investigations" and "in-depth investigations" from a substantive point of view, but also to get a better overview. We deductively defined simple investigations as investigations that consist of at least two phases. One of these two different phases is either an interpretation or a conclusion. In-depth investigations are investigations, where more than one interpretation and/or conclusion was made based on the graph created. The percentages of the different types of micro-level investigations is shown in Figure 7.

Table 3. N_{abs} indicates the total number of occurrences of this type of micro-level investigations, p_{abs} the percentage of all micro-level investigations. N_t indicates the total number of teams that performed this type of investigations, p_t the percentage of all teams. The symbols used (purple square, blue ellipse, green comb, orange diamond) represent the different phases of the QAIC-cycle (Question – Analysis – Interpretation – Conclusion) described in the theoretical framework section

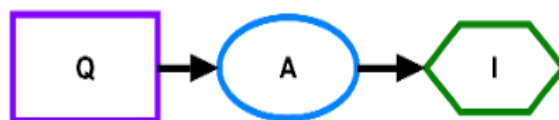| simple investigations						
	title	representation	N_{abs}	p_{abs}	N_t	p_t
targeted investigation	ideal targeted investigation		10	0,15	6	0,4
	descriptive investigation		15	0,22	10	0,6
	postulating investigation		10	0,15	8	0,53
exploratory investigations	exploratory descriptive investigation		10	0,15	4	0,27
	exploratory postulating investigation		9	0,13	4	0,27
	ideal exploratory investigation		1	0,01	1	0,07
	$\sum$		55	0,82	16	1

Figure 8. A simple targeted descriptive investigation, including the phases Question, Analysis and Interpretation. The symbols used (purple square, blue ellipse, green comb) represent the different phases of the QAIC-cycle (Question – Analysis – Interpretation) described in the theoretical framework section

Simple investigations

Depending on the actual phases carried out we can distinguish different subtypes of simple investigations. The flowcharts show six different types of simple investigations that meet these criteria as shown in Table 3.

A crucial distinction was made between targeted and exploratory investigations. While targeted investigations always begin with a research question, exploratory investigations begin with the representation of a graph. Exploratory and targeted investigations can occur in three different forms, as shown in Table 3.

Targeted investigations: In the following section we describe simple investigations that are targeted, since they start with a question. Targeted investigations can be divided into different subtypes which we call descriptive, postulating or ideal. The distinctive element is the sequence of the investigative phases that follow the question phase.

First, we treat targeted investigations following the sequence of phases of the QAIC cycle (so called ideal targeted investigations, as shown in Figure 8).

The second type of targeted investigations is labelled as descriptive investigations, as shown in Figure 8. Teams conducting this type of investigation formulated a research question and created a corresponding graph. Once this graph was interpreted, however, this type of investigation ended. In these descriptive investigations, there are not any statements that go beyond existing data as outlined in the section describing the theoretical framework. This type of investigation is also characterized by the fact, that conclusions are not drawn from the data.

What yet remains unclear is: Did the teams in these cases regard the interpretation of a graph as an appropriate result of an investigation, or were they unable to gain additional insight from the graph that allowed them to make a conclusion going beyond the data? The first suggestion is supported by the analysis of the working process of team 02_white, as shown in Table 4.

In total, 15% of all micro-level investigations are ideal targeted investigations. Six out of 15 teams performed such an investigation at least once. This number may appear to be small. However, one should bear in mind that the learning environment was designed in such a

Table 4. Extract of the protocol of team 02_white, displaying a descriptive investigation. The protocol sections were carefully translated into English by the authors, while trying to stay at the original formulation as much as possible. The right column indicates the categorization of the protocol sections according to the QAIC-cycle

| protocol	categorization
We investigate whether the PM10 value depends on the days of the week	Q
We create graph 1. X = PM10 values; Y= days of the week.	A
We create a Boxplot and set the median.	A
The median is lowest on Sunday.	I
On Friday, the median is highest.	I
The mean 50% of the PM10 values showed a lower variability on Sundays and Saturdays compared to weekdays.	I
The highest value is 157,2µg/m³.	I
This value refers to the 1st of January.	I
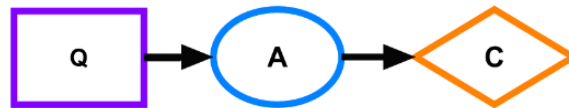
Figure 9. A simple targeted postulating investigation, including the phases question, analysis and conclusion. The symbols used (purple square, blue ellipse, orange diamond) represent the different phases of the QAIC-cycle (Question – Analysis – Conclusion) described in the theoretical framework section

Table 5. Extract of the protocol of team 01_orange, displaying a postulating investigation. The protocol sections were carefully translated into English by the authors, while trying to stay at the original formulation as much as possible. The right column indicates the categorization of the protocol sections according to the QAIC-cycle

protocol	categorization
We investigate the air pollution (particulate matter) during the days of the week.	Q
We create graph 4.	A
During weekends, there is less particulate matter than during the week => graph 4.	C

Table 6. Extract of the protocol of team 01_red, displaying a postulating investigation. The right column indicates the categorization of the protocol sections according to the QAIC-cycle. The protocol sections were carefully translated into English by the authors, while trying to stay at the original formulation as much as possible. The right column indicates the categorization of the protocol sections according to the QAIC-cycle

protocol	categorization
We investigate the relationship between particulate matter concentration and inversion weather conditions.	Q
We create graph 1.	A
During inversion weather conditions, the particulate matter pollution is higher than in non-inversion weather conditions.	C

way that it does not prescribe a specific approach for investigations or even only encourages a specific approach. In this respect, the learning environment was designed to offer the learners the greatest possible freedom. Keeping this in mind, the fact that 40% of all teams conducted ideal investigations is noteworthy.

As one can see, these statements only describe what can be seen in the graph, but there is no general conclusion related to the research question.

The third type of targeted investigations is categorized as postulating investigations and its structure is shown in Figure 9. Compared to ideal investigations, postulating investigations do not include the phase of an interpretation. First, a graph is created to answer a question, then a finding that goes beyond the data is postulated directly afterwards. Generally speaking, it is not possible to conclude from a graph without first interpreting it. It is more likely to assume that the graph was interpreted by the learners, but this interpretation was not explicitly formulated, logged or

discussed in the investigation protocol. Comparing data from the protocols with the audio files of the corresponding teams supports this hypothesis. Additionally, groups that carry out a postulating investigation often refer to the corresponding graph in the conclusion as the audio files reveal, but they only sometime include this phase in the written protocols. This is illustrated in the following protocol section of team 01_orange as shown in Table 5. When carrying out postulating investigations, some teams do not refer to a graph at all, for example team 01_red, as shown in Table 6.

It is striking that all eight teams who carried out any postulating investigation also carried out at least one ideal or descriptive investigation. We already argued that the interpretation of a graph, even if it is not noted explicitly interpreted in the written protocol, is necessary to draw a conclusion of it. What is unclear is why this interpretation is not manifested in the protocols in all cases. One hypothesis is that the teams considered the conclusion trivial based on the graph created. Another
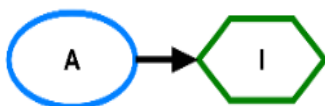
Figure 10. A simple exploratory descriptive investigation, including the phases analysis and interpretation. The symbols used (blue ellipse, green comb) represent the different phases of the QAIC-cycle (Question – Interpretation) described in the theoretical framework section
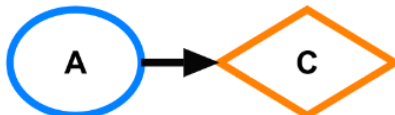

Figure 11. A simple exploratory postulating investigation, including the phases analysis and conclusion. The symbols used (blue ellipse, orange diamond) represent the different phases of the QAIC-cycle (Question – Conclusion) described in the theoretical framework section

Figure 12. A simple ideal exploratory investigation, including the phases analysis, interpretation and conclusion. The symbols used (blue ellipse, green comb, orange diamond) represent the different phases of the QAIC-cycle (Analysis – Interpretation – Conclusion) described in the theoretical framework section

reason could be that the difference between the interpretation of the data and the conclusion based upon an interpretation is not clear for the teams.

Exploratory investigations: As already described above, we categorized an investigation as exploratory investigation when it started with the creation of a graph as the first phase. We have chosen this term since it is in good accordance with the general idea behind exploratory data analysis that "the data should speak for itself". This type of investigation is seen as starting point for the formulation of a hypothesis or new research questions. However, we want to note that it is also possible that the students who performed exploratory investigations had a research question in mind but simply did not discuss it further or write it down. Regarding exploratory investigations, three different types of exploratory investigations have been identified by analyzing our data.

Exploratory descriptive investigations as shown in Figure 10 occur 10 times in total. In three of this ten investigations, it was not clear whether the protocol sections should be categorized as an interpretation or a conclusion. In these cases, the categorization of the first rater was chosen.

A similar situation appeared regarding exploratory postulating investigations, as shown in Figure 11. These divergent categorizations of the units by the two raters clearly indicate that our operationalization of interpretation and conclusion is not able to perfectly distinguish between these two phases in exploratory investigations, although it worked perfectly fine for the categorization of targeted investigations. However, it can also be seen as a hint, that when students do not have

a research question in mind when creating or analyzing a graph, it is much more difficult for them to distinguish between the interpretation and the conclusion based upon this interpretation, resulting in a hybrid-phase with two facets (interpretation and conclusion). This is additionally supported by the fact that only one team performed an ideal exploratory investigation, as shown in Figure 12.

After this overview of the different types of simple investigations, in-depth investigations will be discussed in the next section

In-depth investigations

In-depth investigations differ from simple investigations in the number of interpretations and findings. Overall, investigations by five different teams were carried out that can be categorized as in-depth investigations. Due to this comparably small number of in-depth investigations, all of them are listed in Table 7.

In six out of these eight investigations, only one graph has been created, either at the beginning of the investigation or after a research-question was formulated. This is afterwards followed by a combination of two or three conclusions or interpretations. It must be noted, that in all in-depth investigations which contain two conclusions or two interpretations, the first one is independent from the second one. Especially when two conclusions appear, the second one is either independent from the first one or an enhancement to the first one.

For the remaining two in-depth investigations, the phase analysis is performed twice, both teams substantially changed the first graph. Listening to the

Table 7. Overview of all in-depth investigations. Five teams conducted in-depth investigations while three teams performed more than one in-depth investigation. The symbols used (purple square, blue ellipse, green comb, orange diamond) represent the different phases of the QAIC-cycle (Question – Analysis – Interpretation – Conclusion) described in the theoretical framework section

team	Flow diagram
01_grey	
01_violet	
02_blue	
02_grey	
02_pink	

audio, we found that one team (02_blue) changed their graph because they were not satisfied with the representation of the graph. They changed from a discrete axis (using bins) to a continuous axis (using boxplots) in order to enhance their conclusion. Team 02_pink changed their first graph because they investigated whether a trend they found for the whole period of the sample ranging from 2014 to 2018 (the relationship between air temperature and particulate matter concentration) also holds for the year of 2016.

Macro-level Investigation Processes

The purpose of this article is to highlight in an exploratory way which approaches physics teacher-students show when conducting this type of investigations. We want to show ways how it is possible to examine students' approaches to multiple linked investigations from an educational point of view. Hence, after this description of all types of investigation detected on the micro-level, we now turn to investigation processes which refer to the macro-level as shown in Figure 4. Macro-level investigations are an ensemble of investigations and they consist of at least two simple investigations.

So, the full investigation process of a macro-level investigation is characterized by the sequencing of several micro-level investigations. Hence, we treat all micro-level investigations as one single investigation process per team as described earlier.

The results show that the teams followed three different approaches in their macro-level investigation processes. Seven of the fifteen teams followed the first approach. Thereby, the teams recorded a research question at the beginning of all their investigations. Also, seven teams followed the second approach. These teams switched between targeted and exploratory investigations. However, how they switch between these two types of investigations differs. Most teams seem to switch randomly, the flowcharts in Figure 13 and Figure 14 show examples of such switches. To our knowledge, such changes between targeted and exploratory investigations have not been systematically researched in science education research yet.

Finally, one team followed the third approach found. This team exclusively carried out explorative investigations. This teem investigated the data by just creating graphs on the chance of spotting a pattern, structure or correlation, but without a specific object of investigation in mind.

In the following part, we want to report two selected findings, which we think may be good starting points for future research. The first finding addresses investigation processes in which students incorporated the findings from several simple investigations (during the intervention) on the micro-level for a final conclusion. We illustrated the reference to a result of an earlier investigation that was explicitly mentioned in the conclusion phase of a current investigation by a dotted line in the flow chart (see Figure 15).
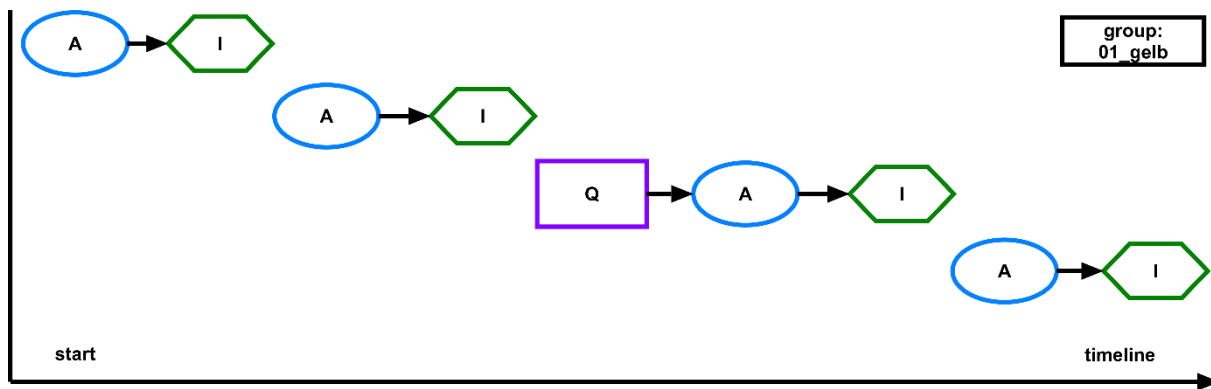
Figure 13. Flowchart of team 01_yellow. Switching between exploratory and targeted investigations. The symbols used (purple square, blue ellipse, green comb) represent the different phases of the QAIC-cycle (Question – Analysis – Interpretation – Conclusion) described in the theoretical framework section. The black horizontal line indicates the chronological order of the various micro-level investigations
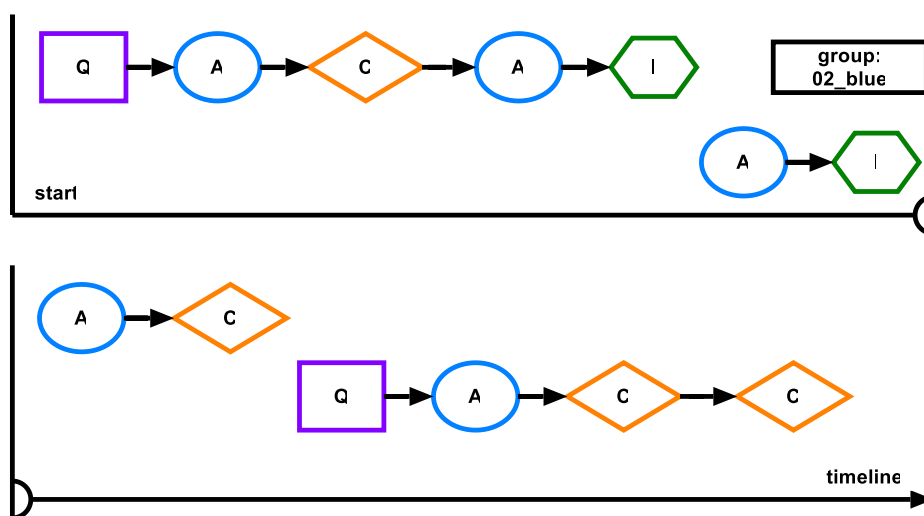

Figure 14. Flowchart of team 02_blue. Switching between exploratory and targeted investigations. The symbols used (purple square, blue ellipse, green comb, orange diamond) represent the different phases of the QAIC-cycle (Question – Analysis – Interpretation – Conclusion) described in the theoretical framework section. The black horizontal and vertical lines indicate the chronological order of the various micro-level investigations, starting from top left to bottom right

Table 8. Extract of the protocol of team 02_grey. The extract was carefully translated into English by the authors, the letter-number combination in the parenthesis indicates that the students referred to a fact about particulate matter, which was written on the scaffolding material the students received. "[..]" indicates a reference to a previous statement of the teams

| protocol	categorization
The high particulate matter levels at low temperature can mainly be attributed to the heating season in winter (F3).	C
The elevated particulate matter levels at low temperature may also be due to increased traffic load (F1) [reference to "heating season in winter", previous statement]	C
~ cut out protocol section ~	
In areas of low particulate matter pollution, no inversion weather conditions occur.	I
The values of highest particulate matter pollution occur almost exclusively at inversion weather conditions.	I
Due to inversion, the air in the "Grazer basin" "stops" and cannot mix with higher air layers (F9.3)	C
Particulate matter can accumulate over Graz if air is not removed, as it is the case during inversion [previous statement]	C
We investigate the particulate matter concentration during the seasons of the year taking temperature inversion into account.	Q
We create graph 4.	A
The maximum values of particulate matter concentration are measured between November and February, the peak month is January.	I
This [previous statement] coincides with the results of graph 2, because of the rather low temperatures prevailing there.	C
The months from November to January show inversion weather conditions.	C
This [previous statement] can be seen in graph 4.	C
The results from the investigations shown in graph 2 & 3 are confirmed here (graph 4).	C
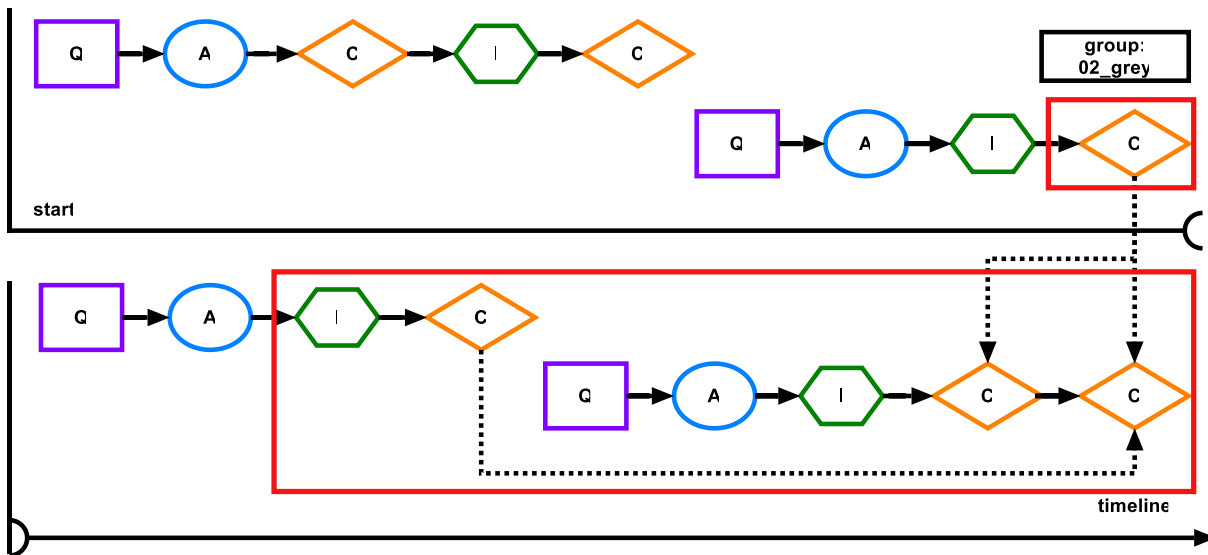
Figure 15. Flowchart of the investigation of team 02_grey. The red framed parts of the investigation process indicate the protocol sections in Table 8. The symbols used (purple square, blue ellipse, green comb, orange diamond) represent the different phases of the QAIC-cycle (Question – Analysis – Interpretation – Conclusion) described in the theoretical framework section. The black horizontal and vertical lines indicate the chronological order of the various micro-level investigations, starting from top left to bottom right
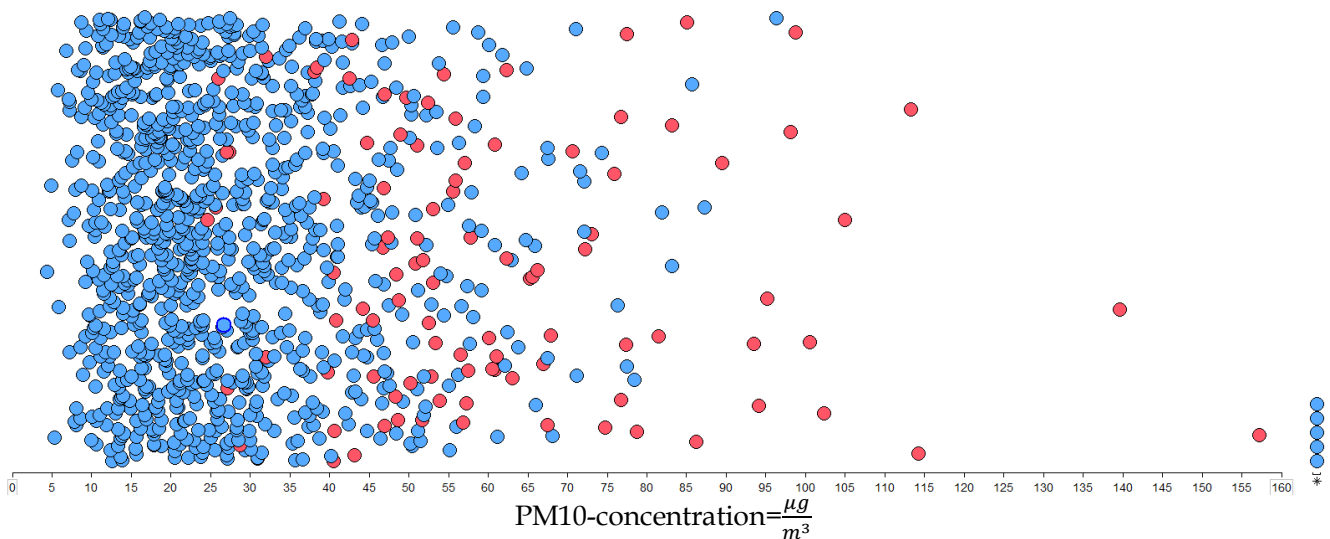

$$PM10\text{-concentration}=\frac{\mu g}{m^3}$$

Figure 16. Graph showing PM10-concentration, color-coded by inversion (red = inversion), created by team 02_grey. Every dot in the figure represents a point of measurement of PM10 at a meteorological station

For the first selected finding, Table 8 shows the relevant extract of the protocol from team 02_grey and its transformation in a flowchart can be seen in Figure 15. The relevant part of the protocol, that is given in Table 8, is highlighted in the red box in the flow chart in Figure 15.

As shown in Table 8, the team first analyzed the relationship between air temperature and particulate matter concentration. As a first conclusion, they stated: *the elevated particulate matter levels at low temperature can mainly be attributed to the heating season in winter* and *the elevated particulate matter levels at low temperature may also be due to increased traffic load*. One can clearly see why this is a conclusion in terms of informal inferential reasoning. First, these statements are claims that integrate

contextual knowledge, in this case the knowledge about the heating season in winter. Furthermore, the use of the terms *can largely be attributed to* as well as *may also be due to* indicates that the conclusion contains a certain degree of uncertainty. Finally, the data displayed in the graph created by this team is used as evidence for the generalizations made.

After this conclusion (smaller red box in the upper right corner of Figure 15), the team investigated the influence of inversion weather conditions on the particulate matter concentration (see Figure 16). They concluded that *Particulate matter can accumulate over Graz if air is not removed, as it is in the case of inversion – see previous statement*. Using these two findings, they proceeded to investigate the relationship between the
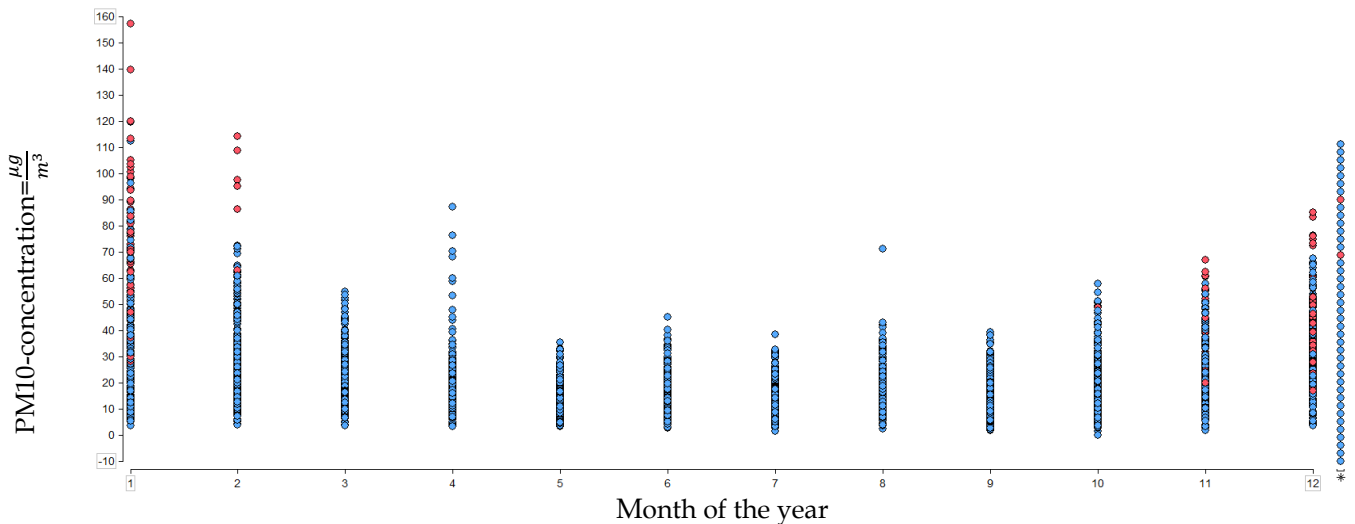
Figure 17. Graph showing PM10-concentration on the y-axis and months of the year on the x-axis, color-coded by inversion (red = inversion), created by team 02_grey. Every dot in the figure represents a point of measurement of PM10 at a meteorological station
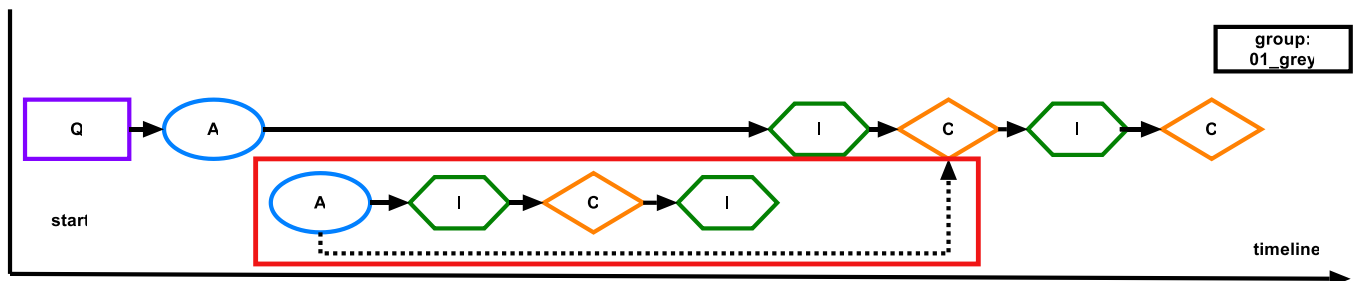

Figure 18. Flow charts of team 01_grey, the investigation of a confounding variable is highlighted in the red box. The symbols used (purple square, blue ellipse, green comb, orange diamond) represent the different phases of the QAIC-cycle (Question – Analysis – Interpretation – Conclusion) described in the theoretical framework section. The black horizontal and vertical lines indicate the chronological order of the various micro-level investigations, starting from top left to bottom right

seasons of the year and the particulate matter concentration. They color-coded the data according to inversion weather conditions, as shown in Figure 17. A red dot indicates inversion. In their final conclusions, they clearly integrated findings from previous investigations, i.e. about the relationship between air temperature and PM10 and the relationship between temperature inversion and PM10. Additionally, they used their last investigation to confirm the two previous ones.

This example of a macro-level investigation process clearly shows that the context of our learning environment seems to allow students to dig deep into the data and supports higher-order thinking. Ben-Zvi and Friedlander (1997) categorize this type of thinking as "meaningful handling of multiple representations". In this so-called mode, students are involved in an ongoing search for meaning and interpretation to achieve sensible results. They make decisions in selecting graphs, consider their contribution to the research question, and make corresponding changes in the data

analysis with a variety of numerical and graphical methods (Ben-Zvi & Friedlander, 1997).

The last selected finding refers to the investigation process of team 01_grey, as shown in Figure 18. Their investigation starts with a research question and the creation of a corresponding graph. This graph shows the relationship between the independent variable of the research question and another variable of the dataset, which could influence the PM10 concentration as well. Hence, this team first investigates the influence of a possible confounding variable on the independent variable before addressing the research question directly, as shown in Figure 18. Subsequently, these previous findings of the team (highlighted in the red box in Figure 18) are incorporated in the investigation of the original research question. Such a structure of an integration, where a confounding variable is treated first, also occurs in the investigation processes of two other teams. These teams follow a very analytic approach in the investigation of these confounding variables. This means that the students consciously depicted which variables may confound their ultimate

findings and checked for possible confoundations in a meaningful way. Additionally, these are three of the few cases of investigations in which particulate matter concentration is not investigated as the dependent variable.

DISCUSSION

In this study, we investigated pre-service teacher-student's approaches to scientific investigations by data exploration when working with data from an online-data repository of a provincial government. In a first step, we synthesized a process model for scientific investigations in the form of exploratory data analysis, the QAIC-cycle. Second, we created a learning environment which allows students to carry out such investigations. The cornerstone for our learning environment was a topic that is relevant for the students but not well known by them and at the same time the access to authentic data that suits such a type of investigation. The topic of particulate matter turned out to fulfill these conditions well: the students had not worked on this topic before, it was familiar to them from media (and other sources) and it was not too tricky to provide them with relevant contextual knowledge so that they could investigate several relevant factors. At the same time, we were able to provide a big set of regional data to work with.

In a third step we developed a tool that helped us to get a written version of the investigation process of the individual teams, the so-called investigation protocols (see Figure 3). The teams were asked to document their investigation processes. Furthermore, we audio documented the teams during the intervention.

After the intervention we analyzed the investigation protocols using qualitative content analysis and a deductive category system based on the QAIC-cycle. Subsequently, the sequencing of the different phases was transformed into corresponding flowcharts. The analysis of these flowcharts revealed different types of simple and in-depth investigations on the micro-level. We found that in a completely open-ended learning scenario, students used both, targeted and exploratory investigations. Some of the graphs created during their investigations were just described by the teams (descriptive investigations). Another type of investigation lacked the interpretation phase, there conclusions were drawn directly from the graphs (postulating investigations). Likewise, ideal targeted investigations were found which correspond to the developed QAIC-cycle (see Figure 1).

The results show that many of the teams' micro-level investigations did not follow the sequence "Analysis – Interpretation – Conclusion". Reasoning about data, especially about the variability in data represents a difficult task for students (Chan & Ismail, 2012; delMas & Liu, 2005) , even up to the tertiary level (Confrey &

Makar, 2002; Hammerman & Rubin, 2004). Reasoning about the variability in data was also identified as a critical aspect when evaluating claims based on data. So, the lack of an interpretation or conclusion phase during the teacher-students' investigations might be due to low statistical reasoning skills.

In addition to these simple investigations consisting of only one interpretation and/or conclusion phase, students also conducted in-depth investigations, in which several interpretations and conclusions were drawn based on the same graph or research question. We think that these results show that learning environments such as the one we developed can allow students to really delve into the data and reason about their findings on a deeper level.

At the macro-level of "investigation processes" (operationalized as a sequence of micro-level investigations) it was shown that students followed three different approaches: Some teams always formulated research questions to start their micro-level investigations, some switched between exploratory and targeted investigations and one team only performed exploratory investigations. So different levels of contextual knowledge, but also different levels of statistical reasoning skills of the teacher-students might be contributing factors to how the teacher-students approach the investigations. Maybe teacher-students with high levels of contextual knowledge only perform targeted investigations since they already know what they want to investigate, due to their contextual knowledge about particulate matter. However, we did not measure teacher-students' contextual knowledge or their reasoning skills, which could be interesting in-future research.

LIMITATIONS

This study had limitations. One limitation is the small sample size. Data were collected from physics teacher-students from two seminars of only two institutions. Hence, we acknowledge that our findings may potentially be unique and cannot be generalized.

Another limitation regards the data collection method used in this study. We used protocols and audio documentation in order to analyze the physics teacher-students' investigations processes. This is significant since it is possible that the students performed various phases of the QAIC-cycle during their investigations but did not document them in the protocol or talk about them.

The fact that we did not measure the teacher-students' contextual knowledge about particulate matter and their statistical reasoning skills represents another limitation. Hence, we could not investigate whether the differences in the students' approaches to micro- but also macro-level investigations were due to these factors.

CONCLUSION

These results represent interesting implications for future research: Is this process of switching between exploratory and targeted investigations a natural emergent effect occurring during an exploratory data analysis scenario? Or is this merely because students struggle to formulate adequate research questions or due to different levels of statistical reasoning skills. Furthermore, the approaches students choose might also depend on the level of contextual knowledge students show, but also how familiar they are with exploratory data analysis. However, in this article we can only provide exploratory insights into the strategies students use. We think that based on our developed model and basic design of our learning environment, this aspect should be target of future research. Additionally, it was shown that students can relate different investigations in order to confirm or revise their previous research. These afore mentioned approaches, which only emerge at the macro-level of the "investigation process", where multiple investigations are looked at, should be investigated more thoroughly.

REFERENCES

Arnold, J., Boone, W. J., Kremer, K., & Mayer, J. (2018). Assessment of competencies in scientific inquiry through the application of Rasch measurement techniques. *Education Sciences, 8*(4), 184. https://doi.org/10.3390/educsci8040184

Ben-Zvi, D. (2005). Reasoning about Data Analysis. In D. Ben-Zvi & J. Garfield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning and Thinking* (Vol. 15, pp. 121–145). Dordrecht: Springer Science + Business Media Inc. https://doi.org/10.1007/1-4020-2278-6_6

Ben-Zvi, D., & Friedlander, A. (1997). Statistical thinking in a technological environment. In J. Garfield & G. Burrill (Eds.), *Research on the Role of Technology in Teaching and Learning Statistics* (pp. 45–55). Voorburg: ISI.

Chan, S. W., & Ismail, Z. (2012). The Role of Information Technology in Developing Students' Statistical Reasoning. *Procedia - Social and Behavioral Sciences, 46*, 3660–3664. https://doi.org/10.1016/j.sbspro.2012.06.123

Confrey, J., & Makar, K. (2002). Developing secondary teachers' statistical inquiry through immersion in high-stakes accountability data. *Proceedings of the Twenty-Fourth Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education PME-NA24,* 3.

delMas, R., & Liu, Y. (2005). Exploring students' conceptions of the standard deviation. *Statistics Education Research Journal, 4*(1), 55–82.

Garfield, J. B., Ben-Zvi, D., Chance, B., Medina, E., Roseth, C., & Zieffler, A. (2008). *Developing Students' Statistical Reasoning: Connecting Research and Teaching Practice.* Dordrecht: Springer Science+Business Media B.V. https://doi.org/10.1007/978-1-4020-8383-9

Hammerman, J. K., & Rubin, A. (2004). Strategies for managing statistical complexity with new software tools. *Statistics Education Research Journal, 3*(2), 17–41.

Irish, T., Berkowitz, A., & Harris, C. (2019). Data Explorations: Secondary Students' Knowledge, Skills and Attitudes Toward Working with Data. *Eurasia Journal of Mathematics, Science and Technology Education, 15*(6). https://doi.org/10.29333/ejmste/103063

Konold, C., & Miller, C. D. (2005). TinkerPlots: Dynamic data exploration: Computer Software. Emeryville, CA: Key Curriculum Press.

Lederman, N. G., Lederman, J. S., & Antink, A. (2013). Nature of Science and Scientific Inquiry as Contexts for the Learning of Science and Achievement of Scientific Literacy. *International Journal of Education in Mathematics, Science and Technology, 1*(3), 138–147.

Makar, K., & Confrey, J. (2014). Wondering, Wandering or Unwavering?: Learners' statistical investigations with fathom. In Thomas Wassong (Ed.), *Mit Werkzeugen Mathematik und Stochastik lernen. Using tools for learning mathematics and statistics* (pp. 351–362). Wiesbaden: Springer Spektrum. https://doi.org/10.1007/978-3-658-03104-6_25

Makar, K., Bakker, A., & Ben-Zvi, D. (2011). The Reasoning behind informal statistical inference. *Mathematical Thinking and Learning, 13*(1-2), 152–173. https://doi.org/10.1080/10986065.2011.538301

Mayring, P. (2014). *Qualitative Content Analysis: theoretical foundation, basic procedures and software solution.* Klagenfurt. Retrieved from http://nbn-resolving.de/urn:nbn:de:0168-ssoar-395173

Pfannkuch, M. (1999). Statistical Thinking in Empirical Enquiry. *International Statistical Review, 67*(3), 223–265. https://doi.org/10.1111/j.1751-5823.1999.tb00442.x

Reinert, D., Prill, F., Frank, H., Zängl, G., & Denhard, M. (2018). *ICON Database Reference Manual: Version 1.2.6.* Retrieved from https://www.dwd.de/SharedDocs/downloads/DE/modelldokumentationen/nwv/icon/icon_dbbeschr_aktuell.pdf?view=nasPublication&nn=369368

Shaman, J., Karspeck, A., Yang, W., Tamerius, J., & Lipsitch, M. (2013). Real-time influenza forecasts during the 2012–2013 season. *Nature Communications, 4*(1), 2837. https://doi.org/10.1038/ncomms3837

Schubatzky, T., & Haagen-Schützenhöfer, C. (2019). Online data repositories as educational resources? A learning environment covering formal and informal inferential statistics ideas in scientific inquiry. *European Journal of Physics, 40*(4), 45802. https://doi.org/10.1088/1361-6404/ab1ffa

Tukey, J. W. (1980). We Need Both Exploratory and Confirmatory. *The American Statistician, 34*(1), 23-25. https://doi.org/10.1080/00031305.1980.10482706

Utts, J. M., & Horton, N. J. (2018). What is Statistics? In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International Handbook of Research in Statistics Education*. Springer.

http://www.ejmste.com