European Journal of Science and Mathematics Education, 2022, **10**(4), 529-546
ISSN: 2301-251X (Online)
OPEN ACCESS
Research Article

An Inquiry Perspective on Statistics in Lower Secondary School in Denmark and Japan–An Elaboration and Modelling of the Anthropological Theory of the Didactic Through Two Statistics Classrooms

Camilla Hellsten Østergaard [1*]
0000-0001-8049-8969

[1] Department of Science Education, University of Copenhagen, Copenhagen, DENMARK
[*] Corresponding author: camilla.h.oestergaard@ind.ku.dk

Citation: Østergaard, C. H. (2022). An Inquiry Perspective on Statistics in Lower Secondary School in Denmark and Japan–An Elaboration and Modelling of the Anthropological Theory of the Didactic Through Two Statistics Classrooms. *European Journal of Science and Mathematics Education, 10*(4), 529-546. https://doi.org/10.30935/scimath/12401

ARTICLE INFO

Received: 28 May 2021
Accepted: 3 Jul 2022

ABSTRACT

We present a detailed analysis of two statistics lessons in lower secondary school, one in Denmark and the other in Japan. The aim of the study is to better understand how inquiry perspectives are implemented in statistics education and what cultural factors shape them. In particular, we draw on the theoretical framework of the Anthropological Theory of the Didactic. The first impressions of the lessons changed during the analysis, and the main differences became explicit in the two question and answer diagrams and the praxeological analysis of the lessons. The two lessons involved cases comprising an "experimental activity with many questions" (Denmark) and "structured problem-solving" (Japan), which differed on a number of points that we shall demonstrate in the analysis. We shall also discuss hypotheses on the possible causes of these differences. We emphasize that our use of question and answer diagrams offers a new way in which to distinguish various kinds of "inquiry-oriented" lessons.

Keywords: inquiry, statistical education, question and answer diagrams, praxeologies, paradidactic infrastructure

INTRODUCTION

One of the major research themes in statistics education in lower secondary school is how to educate independent and statistically literate learners so that they can make data-based decisions and become critical citizens. The main rationale is that through the use of inquiry processes (e.g. Tukey, 1977; Wild & Pfannkuch, 1999), students can learn and understand real-world phenomena (Lehrer & English, 2018). Statistical inquiry is a complex and dynamic system, composed of multiple factors, such as tasks, data, exploration of statistical concepts, digital tools, engagement with informal statistical modelling (Pfannkuch et al., 2018) and classroom culture, including modes of discourse and argumentation amongst students and teachers (Ben-Zvi et al., 2018). In statistical inquiry, it is not enough to manipulate and calculate given statistical representations or descriptors. Instead, students need to have a more holistic approach to data inquiry. This means that students need to shuttle amongst the data, model and context, make sense of the situation and how to construct models or visualizations and think about what questions and conjectures to pose and what variables to explore (Pfannkuch et al., 2018).

In Denmark and Japan, there is a growing interest in statistical inquiry, which entails a deeper understanding of what this type of inquiry requires in order to better exhaust its potential. The national curriculums of both countries recommend an extensive focus on inquiry activities. However, looking at the Danish and Japanese classrooms, we get mixed impressions. In the Danish case, 20 grade eight students were actively engaged in collecting data and worked in small groups. Some groups were in the classroom, while others were in the hallway. The majority of the students did push-ups on the floor, counting along the way; others were out of breath, having reached their maximum performance level. They collected their own data to investigate the following problem: "Is there a connection between how many push-ups and how many jumping squats you can do?" The students were asked to use their (own) personal computers to process the data. In the Japanese case, 33 grade seven students were seated at their desks, while the teacher stood in front of the blackboard with information about two buses. The teacher orchestrated a whole-class discussion about the following problem: "which bus is better to ride?"

In this article, we present a thorough analysis of the two cases from lower secondary school. Our aim is not only to describe the clear differences in the scripts but also analyze the cases from an institutional perspective in order to understand how cultural factors shaped the implementation of the inquiry.

To explore the cases, we used the theoretical framework of the Anthropological Theory of the Didactic. We present the development of questions and answers by the teachers and students in the two cases, model the praxeological process and discuss the differences based on the paradidactic infrastructure and how it influences different conditions in the classroom.

Fundamental Elements of the Anthropological Theory of the Didactic

The Anthropological Theory of the Didactic (ATD) models the evolution of practice and knowledge in educational situations by considering them as situated in an institutional ecology determined by a cultural and societal context (Chevallard, 1992). The analysis of this article relies on three tools from ATD:

(1) question and answer (Q&A) diagrams,

(2) praxeological modelling, and

(3) the paradidactic infrastructure.

The dialectics of questions and answers are the driving force of knowledge development (Bosch & Winsløw, 2016). The dynamics of questions and answers can unfold in various ways during teaching. Questions may be posed by both students and teachers, and the organization of their study and research may be more or less directed by the teacher and correspond to the requirement of more or less autonomy and effort from students. Concretely, a didactical process may begin with a (small or big) question (Q_0). It can be a simple exercise that the students can readily solve (based on their answers or the application of praxeologies–further explained below). Otherwise, an elaborate study and inquiry process would be required. More generally, the analysis of how questions and answers evolve during a didactical process can be used to characterize the didactical approach. One can map the questions (Q_k) and answers (A_i) in a tree diagram to model and visualize the possible paths that the teacher and students will follow. The Q&A diagram stresses the dialectics of the questions and answers, the students' participation and autonomy, the outflow and the way in which the questions and possible answers unfold over time (Jessen, 2014; Winsløw et al., 2013). Q&A diagrams are used to analyze inquiry processes and visualize situations where students consult existing knowledge.

The term *question* does not necessarily or simply refer to a phrase which is grammatically constructed as a question; it also has to be understood semantically, as a discourse which represents a problem to someone in a didactical setting, such as students. For instance, the phrase "What do you think?" is a grammatical question, but it is not a problem in any common sense (Bosch & Winsløw, 2016). Faced with a challenging question to be answered, we can sometimes produce hypotheses based on what we believe, while giving answers that are more complete, may require us to extend our practical and theoretical knowledge – in short, *praxeologies.* In this sense, *answers* represent preliminary, partial or final results from our study and inquiry of a question. In didactical processes, answers are inseparable from the development of students' praxeologies as questions are often deliberately given to students in view of them developing certain praxeologies.

In ATD, practice and knowledge are modelled as *praxeologies* (Chevallard, 1999). A praxeology entails a *practice block* (or know-how) and a *logos block* (or know-why). The practice block is formed by a *type of task* and a *technique*, which can be either given or part of an inquiry process. The logos block is formed by a *technology* (discourse on the practice block, often relating it to other practice blocks) and a theory (a more general discourse) that explains and justifies the technology; in higher mathematics, we may think of "definitions" and "theorems" as typical elements of a theory, while informal mathematical notions and assumptions are often more important in the mathematical theories that are common in school institutions. Praxeologies takes place in didactic systems which, in a classroom context, is formed by a group of students and a teacher.

In order to discuss two didactic systems, such as those alluded to in the introduction, we must also consider the wider institutional context in which they occur and the didactic praxeologies developed by the teacher outside the system (before and after class), for example, the preparation or evaluation of lessons, reflections on students' learning and participation in professional development courses. The praxeologies that the teacher develops outside the classroom are called paradidactic praxeologies and require an infrastructure that goes beyond the necessities of classroom practice–a paradidactic infrastructure (Miyakawa & Winsløw, 2019; Winsløw, 2011). This infrastructure gathers "apparently unrelated factors as a coherent whole, which conditions and constrains a particular set of praxeologies, without determining them entirely" (Miyakawa & Winsløw, 2019, p. 284). Some of the conditions are visible, such as the time and workspace that teachers can use to prepare their teaching and the resources available for their work (Østergaard & Winsløw, 2021; Winsløw, 2011). Other conditions are less visible and difficult to describe, such as opportunities to share mathematical and didactical praxeologies. The paradidactic infrastructure is crucially related to the institutional system in which the teacher works, and the two must be studied together (Miyakawa & Winsløw, 2019).

In this article, we discuss two statistics lessons, one in Denmark and the other in Japan. Our aim is to describe the clear differences in script and analyze the cases from an institutional perspective in order to understand how cultural factors shape statistical inquiry. We address the following questions:

1. RQ1: What are the dynamics between questions and answers in the Danish and Japanese lessons?
2. RQ2: What statistical praxeologies are developed as answers?
3. RQ3: What are the principal similarities and differences between the lessons?
4. RQ4: How can the principal differences be explained in terms of the paradidactic infrastructure?

METHODOLOGY

The present study is a descriptive case study (Yin, 2013), where we present an in-depth description and analysis of local practices in two specific contexts. In the analysis, we interpret and discuss the cases to generate cross-case themes, patterns and findings.

The data consisted of information from a variety of sources: interviews with teachers, audio and video recordings of lessons, field notes, textbooks and official curricula and contextual information about teachers, schools and countries. The diverse sources made up the raw data for constructing each of the cases. The raw data were organized and edited into accessible final cases about the two statistical classroom practices, with information about the different institutions and their particular classroom practice. Both the Japanese and Danish data were translated into English, and the Japanese data was translated with help from a Japanese colleague.

The two selected schools were comparable in the sense that both functioned as "model schools", and both lessons were taught by teachers who were considered "experts" in their system. The Danish school is a so-called mathematics resource center in a large municipality, meaning that a group of mathematics teachers teach and guide mathematics teachers from other schools in the municipality. The Japanese school is a so-called Fuzoku school that is, attached to a university and its teacher education; the teachers are experienced and engaged in development projects and lesson studies. At the same time, the two selected lessons are representative of the two schools in the sense of being similar in "script" (cf. Stigler & Hiebert, 1998, p. 1) to the other lessons taught there. However, the small-scale study presented cannot generalize to all mathematics classrooms in Denmark and Japan.

The Japanese case was constructed through video observations of a grade seven class comprising 33 students as part of a lesson study. The data further consisted of several teachers' reflections about the lesson. The Danish case was constructed through audio observations and lesson notes from a grade eight class consisting of 20 students.

In the analysis of the didactical process, we used three theoretical tools from ATD. We first constructed a Q&A diagram (Winsløw et al., 2013) displaying the development of questions and answers by the teacher and students in the two classes. To do so, we analyzed the whole-class discussions semantically and, in particular, categorized significant contributions by the teacher and students to the dialogue as questions and answers, both related to an overall question (Q_0) proposed by the teacher. In the analysis, the contributions were categorized as questions if they called for statistical praxeologies that were not developed in the dialogue, hypotheses and suggestions requiring further mathematical support and contributions directing the students' development of praxeologies. Students' opinions such as "we need to ask more people" and "higgledy-piggledy" were categorized as questions because they call for statistical praxeologies and need further mathematical support. Answers comprised statistical praxeologies that could help answer the questions and which the students normally learned through the teacher, workbooks or the Internet. The contributions were categorized as answers if they included "official" knowledge and considered an aspect of statistical techniques, technologies or theories.

We then modelled the answers developed in the two processes in terms of praxeologies. The praxeological modelling process took its point of departure from the Q&A diagram, where the types of tasks were identified by analysing the semantical questions. The remaining praxeologies were found in the semantical answers (or when these were missing).

In the third part of the analysis, we discuss the two lessons and explain the main differences identified on the basis of the paradidactic infrastructure. This means that we point to conditions that appear to support or hinder the development of statistical praxeologies in the two contexts. In our analysis, we present several hypotheses to understand the differences in the two lessons.

THE CASE OF "EXPERIMENTAL ACTIVITY WITH MANY QUESTIONS"

The Danish school was designated among several schools as the only mathematical resource center in a large municipality in Denmark. As a resource center, it receives external funding for a group of mathematics teachers who are expected to be role models and first movers as well as guide and supervise other mathematics teachers at other schools in the municipality and design professional development courses. The teacher, Mr. Dan, is an experienced mathematics teacher with a special interest in digital tools and inquiry. He also teaches biology, physics and chemistry.

Twenty grade eight students are seated around six tables, each with a personal computer in front of them. Dan presents the statistical question of the lesson–"find the correlation between jumping squats and push-ups"–and asks one student to prepare a shared Google Sheet so that the all students could later fill in their data: name, jumping squats and push-ups. Thereafter, Dan writes a list on the blackboard, which is an instruction about what the students are expected to do in the lesson: "Do this…", "make this…", "write…" (**Figure 1**).

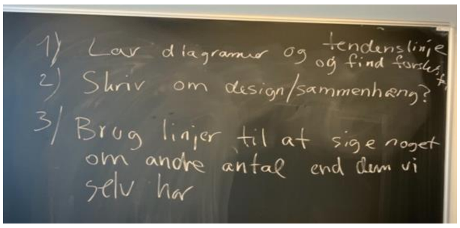


Figure 1. Teacher's writing on the blackboard, translated into English: (i) draw diagrams and trend lines and find the functional equation; (ii) write about design/correlation? and (iii) use lines to say something about quantities other than those we have

The students form smaller groups, leave the classroom and start collecting data. Over the next 11 minutes, they jump, count, perform push-ups, and count again. They return to the classroom, some of them a bit breathless. During the students' data gathering, Dan circulates among the groups comments on the students' work. Back in the classroom, the students type their data in the shared Google Sheet and start drawing diagrams, some in Google Sheets, and others in GeoGebra. A few students succeed in constructing a scatterplot and trend line (**Figure 2**). After finding a trend line and a linear regression function, the students and teacher elaborate on the function and discuss the quality of the data. They also discuss whether there is a connection, correlation and causality between the two data sets. Dan presents the task for the next week: "conduct your own statistical inquiry", including the aim: "students can pose questions, design inquiry processes and implement design in practice, including the use of different statistical descriptors…"

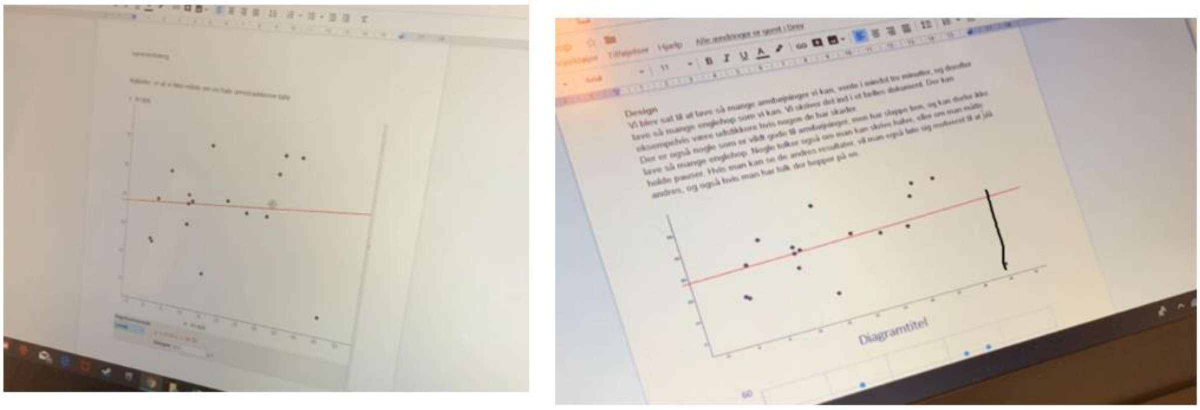


Figure 2. Two students' scatterplots and trend lines

The Dynamics between Questions and Answers in Whole-Class Dialogues

The dynamics between questions and answers in the Danish lesson, as visualized in the Q&A diagram (**Figure 3**).

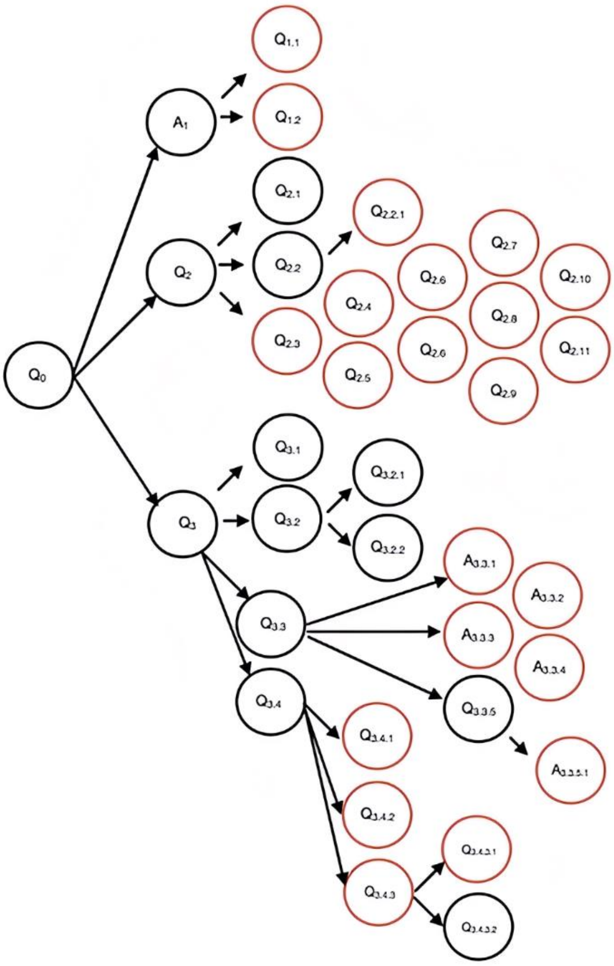


Figure 3. Q&A diagram of a Danish lesson. Contributions from the teacher are in black, while those from the students are in red

Q_0: “Today, you have to find the connection between how many push-ups and jumping squats you can do…”

A_1: “Draw diagrams, draw a trend line and find the linear regression function…”

$Q_{1.1}$: “How do I draw a trend line in Google Sheets?”

$Q_{1.2}$: “Google “how to find trend line in Google Sheets””

Q_2: “Describe the design of the data collection”

$Q_{2.1}$: “How can we make the experiment trustworthy?”

$Q_{2.2}$: “… you think we should use data from just one group or the whole class?”

$Q_{2.2.1}$: “We will get a better result if we ask the whole class”

$Q_{2.3}$: “We need to ask more people”

$Q_{2.4}$: “We cannot compare a person who cannot do the exercise and a person who is an expert”

$Q_{2.5}$: “I think that the order of the exercises is important; you might be tired after doing the push-ups at first”

$Q_{2.6}$: “And the way you do the exercises, are you doing them properly?”

$Q_{2.7}$: “You have to make sure that everybody is doing their best”

$Q_{2.8}$: “Cheering! You might get a better score if your friends are cheering”

$Q_{2.9}$: “I think that some of us gave up before we even got started”

$Q_{2.10}$: “I don’t think bicycle racers can do push-ups”

$Q_{.2.11}$: “I wrote: There is no real connection between push-ups and jumping squats, maybe because we haven’t been serious enough”

Q_3: “Describe the connection in words”

$Q_{3.1}$: “What kind of connection are we searching for?”

$Q_{3.2}$: “You have to explore whether there is a mathematical connection?”

$Q_{3.2.1}$: “Take a look at the dots in your diagrams; is that a connection?”

$Q_{3.2.2}$: “How many of you see a connection?”

$Q_{3.3}$: “How many of you have found the function that shows the connection?”

$A_{3.3.1}$: “y=0.13x+28”

$A_{3.3.2}$: “28.76: you need two digits”

$A_{3.3.3}$: “Push-ups, the slope, how much it increases and the intersection with the y axis

$A_{3.3.4}$: “0 push-ups equal 29 jumping squats”

$Q_{3.3.5}$: “Is the slope 0.13 big or small?”

$A_{3.3.5.1}$: "You will only increase your jumping squats by 13"

$Q_{3.4}$: "How do you think the dots are placed?"

$Q_{3.4.1}$: "Higgledy-piggledy"

$Q_{3.4.2}$: "Yes, it might not fit if we take a random person"

$Q_{3.4.3}$: "We do not argue for a relation; we see no relation between arms and legs"

$Q_{3.4.3.1}$: "We use different muscles in the different exercises"

$Q_{3.4.3.2}$: "In my head, it makes sense. I see causality"

In the above Q&A diagram (**Figure 3**), we see how the teacher poses the question of the day (Q_0). This question is immediately followed by the teacher's formulation of the main answer of the lesson (A_1), namely, the method to be followed to answer the question: "draw diagrams, draw a trend line and find the linear regression function". In A_1, the techniques are provided, which the students have to carry out. In the diagram, we see how a question ($Q_{1.1}$) is derived from A_1 when a student questions the instrumental techniques. The student receives no response from the teacher; instead, she Googles the answer to her question. The data do not provide as explanation regarding the teacher's refusal to respond to the question posed.

Two questions (Q_2 and Q_3) were derived from the Q&A diagram. Q_2 deals with the data collection and leads to many of the new questions, most of which are posed by the students. The new questions are important as they problematize different principal choices pertaining to how to conduct a valid data collection, but the questions remain unanswered as the students are not asked to elaborate or address the questions posed. In question Q_3, the students are asked to present and comment on the correlations between the data $Q_{3.1...}$ $Q_{3.4}$, for example, comment "in words" on the linear regression function and describe the scatterplot. $Q_{3.1}$ and $Q_{3.2}$ are left unanswered. Question $Q_{3.3}$–How many of you have found the function that shows the connection? refers to instrumental techniques designated in A_1. A student answers ($A_{3.3.1}$) with the model y=0.13x+28, and the students' subsequent answers describe the function in relation to the context ($A_{3.3.3...}$ $A_{3.3.5.1}$). In the discussion around the linear regression function, we see that none of the students refer to the connection between the scatterplot and function or to the correlation coefficient. Question $Q_{3.4}$ focuses mainly on the informal degree of the correlation seen in the scatterplot, and a student describes the scatterplot as "Higgledy-piggledy" ($A_{3.4.1}$), while another proposes that there is no mathematical or real-world connection between the observations of push-ups and jumping squats ($Q_{3.4.2}$ and $Q_{3.4.3}$). However, the teacher does not address the students' objections and ends the discussion with $Q_{3.4.3.2}$: "In my head, it makes sense. I see causality". This question points back to A_1, which is never justified.

More generally, we see that many of the questions are not pursued further and that only a few answers are proposed mostly by the teacher. The classroom dialogue is almost exclusively oral in nature, and the blackboard is only used once in the lesson when the teacher formulates A_1.

Statistical Praxeologies

The aim of the lesson was not to find an answer to Q_0. Q_0 is the context of the inquiry. We can identify two statistical praxeologies developed by the students and teacher. The types of tasks are as follows:

1. T_1: Conduct an experiment to investigate the correlation between two (more or less strictly defined) observables.
2. T_2: Determine how two given data sets (two vectors of equal dimension) are (cor)related.

The task (T_1) is about producing data or investigating the correlation between two observables, including how the data are produced and how much data are needed. To inquire into T_1, the teacher poses several questions (Q_2, $Q_{2.1}$ and $Q_{2.2}$). The techniques include the students' data collection, the areas in which the students have first-hand experiences, which are questioned in the classroom ($Q_{2.2.1}$, $Q_{2.1}$.... $Q_{2.11}$). The technology developed by the students about the quality of the data is critical, citing factors such as poor execution of the exercises, a lack of explicit consensus on how to execute the exercises, etc. The questioning of the techniques employed for the data collection (T_1) mostly results in single hypotheses, where the students try to make sense of the quality of the data. However, the hypotheses are not discussed in detail and are not formulated in formal or semi-formal statistical terms, let alone further investigated using statistical techniques or theories.

The task (T_2) is an analysis of the correlation between the data found during the initial experiment. In the lesson, there is a clear but implicit expectation that the students know and use an instrumental technique to produce a trend line based on the collected data and that they use different tools such as GeoGebra, Excel and Google Sheets to carry out the method proposed by the teacher (A_1). A first step towards solving the task is to design a table to collect and display all the students' observations; one student is explicitly asked to construct this table and digitally share it with the rest of the class. The second and third techniques are instrumental and "transform" the table into a scatterplot and a trend line. The fourth technique (also instrumental) exhibits the equation. It is visible on the screen next to the trend line, and the students only need to copy the equation from the screen; in that sense, the two instrumental techniques solve all three subtasks. The techniques provide models of the data: a table, a scatterplot, a trend line and a function accompanied by technologies to make sense of the models, going back and forth between the four models and the real-world context and communicating what has been learned. The students discuss the equation ($A_{3.3.1}$ ... $A_{3.3.4.}$) and question whether there is a connection and whether the equation says anything about the (real-world) data at all ($Q_{3.4.1}$... $Q_{3.4.3}$). The technologies developed by the students are mainly about elaborating on the equation by describing the slope and intersection with the *y*-axis. In the discussion about correlations, the students do not arrive at any answers; instead, they pose several important questions about the visual image. The different models: tables, scatterplots and linear regression function are not compared or discussed in detail; the reliability is questioned by some students, but the teacher's authority shuts down the students' questioning about the assumption of the correlation. The teacher presents no statistical explanation or reasoning in favor of that assumption but, instead, refers to his own subjective judgment: "In my head..." The dialogue about T_2 continues at a concrete level, but further theoretical considerations about the validity of the trend line are absent.

CASE OF "STRUCTURED PROBLEM-SOLVING"

The Japanese school is a *Fuzoku* school, that is, it is attached to a university and its teacher education. The school has a high academic standard, and there is considerable competition around admission. Fuzoku schools often host open lessons for teachers and other stakeholders in the educational sector, and they are first movers in preparing and implementing new curricula. The teacher, Mr. Akamoto, is an experienced mathematics teacher. He is especially interested in problem-solving and blackboard design. In preparation for his lessons, he designs models of the expected blackboard, including students' expected answers. Akamoto is very interested in the development of a new Japanese curriculum and frequently participates in lesson studies around Japan. Akamoto considers it "very important that we explore new ways of teaching statistics".

Thirty-three grade seven students are seated at individual tables, closely arranged in four rows. The students have little previous knowledge about statistical analysis. Akamoto draws a table on the blackboard (**Figure 4**), and explains the problem of today's lesson: "which bus is better?" The students have to compare bus numbers 17 and 18. A piece of paper hides the data at the beginning of the lesson, and Akamoto lets the students see the table for two seconds. The students reflect on the brief observations and formulate a hypothesis about which bus is better. They are then allowed to see the table for as long as they like, and some of them change their guesses. The problem of comparing the two buses is not straightforward because we have a different number of total observations of the two buses. Akamoto puts the question into a wider perspective: how to compare two sets of data with a different number of observations? To do this, the students extend their box of statistical descriptors to include relative frequency and cumulative relative frequency.

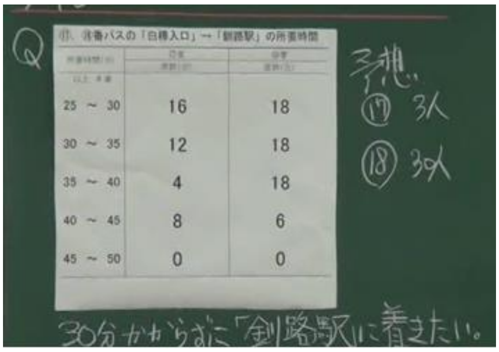


Figure 4. Frequency distribution of bus times

The lesson ends as Akamoto poses a question derived from the main problem and introduces new data represented in a table and a frequency line graph (**Figure 5** to the right). With the new data, the students have to decide which bus to ride (they have a maximum time of 60 minutes) and evaluate the two models. Akamoto presents another table (**Figure 5** lower right corner) to establish the need to develop additional statistical tools to solve the problem–cumulative relative frequency. During the lesson, several students are invited to share their thoughts with the rest of the class. The blackboard is used to produce a cumulative account of the students' guesses, calculations and arguments. The students write additional notes in their notebooks, including copying from the blackboard.

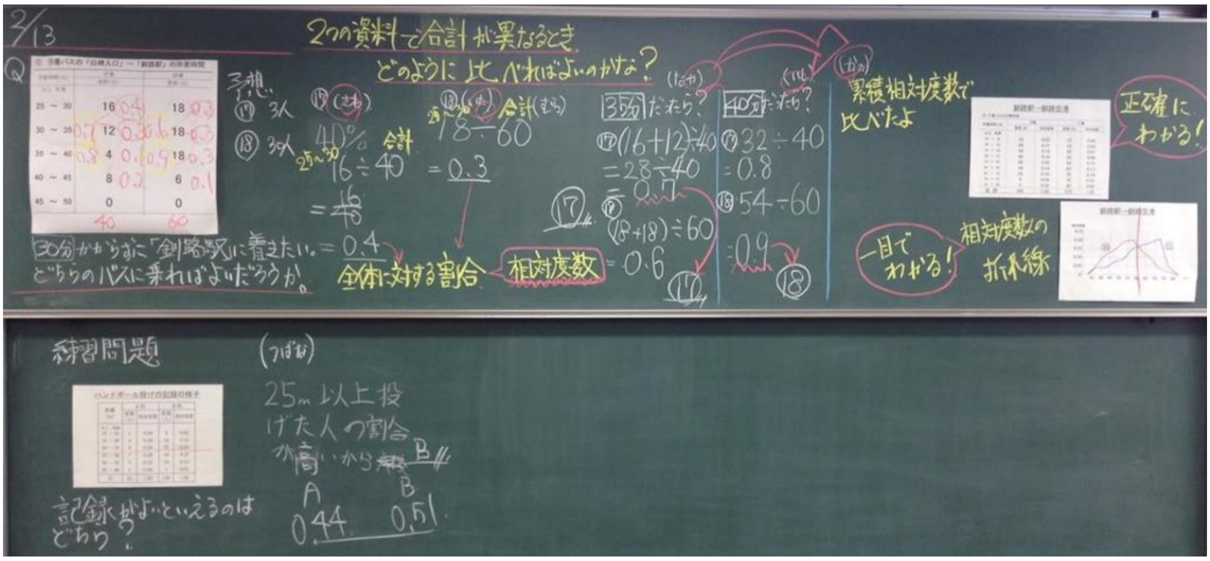


Figure 5. Blackboard with students' and teacher's questions and answer

The Dynamics between Questions and Answers in Whole-Class Dialogues

Q_0: "Which bus is better for us to ride?" (**Figure 6**).

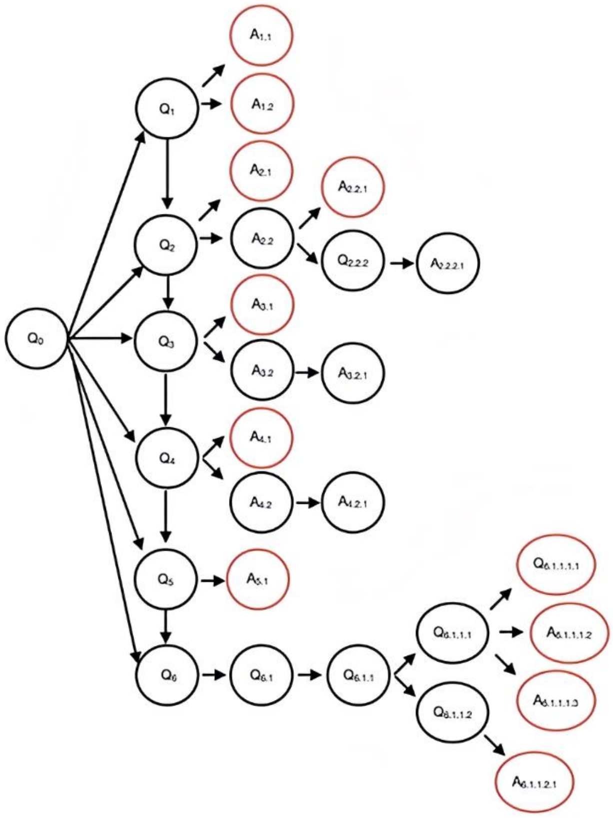


Figure 6. Q&A diagram showing questions and answers by the teacher (black) and students (red) in the Japanese classroom

Q_1: "Which bus is better to ride (if we only have two seconds to interpret the data)?"

$A_{1.1}$ "bus 17"–"three students"

$A_{1.2}$ "bus 18"–"thirty students"

Q_2: "Which bus is better (if it is possible to study the data further)?"

$A_{2.1}$: Some students change their guess (the students' preliminary answers are challenged)

$Q_{2.2}$: "How do we compare two sets of data with a different number of observations?"

$A_{2.2.1}$: "We have 16 buses, which have a driving time of less than 30 minutes. This means that we have 16 out of 40. That is equal to 0.4"

$Q_{2.2.2}$: "What is 18 out of all the observations?"

$A_{2.2.2.1}$: "18:60, 18 is the number of observations with a driving time between 25 and 30 minutes; 60 is the total number of observations of bus 18. Equals 0.3"

Q_3: "Which bus is better and why?"

$A_{3.1}$: "We cannot make a choice only from the number of buses with a driving time between 25 and 30 minutes; we also have to think about the total amount of observations"

$A_{3.2}$: "(repeats) We cannot make a choice based solely on the number of buses with a driving time between 25 and 30 minutes. We must also consider the total number of observations. The ratio is what we call the relative frequency of buses with a driving time between 20 and 30 minutes"

$A_{3.2.1}$: "I want you to read and discuss the definition of relative frequency in your textbooks"

Q_4: "What will happen if we change our bus ride to less than 35 minutes?"

$A_{4.1}$: "Bus 17 (the teacher writes on the blackboard and explains the calculation, relative frequency) (16+12):60=28, erases 60, (16+12):40=28:40=0.7"

$Q_{4.2}$: "What will happen if we change the starting point to within 40 minutes?"

$A_{4.2.1}$: "Bus 18 (the student writes on the blackboard and explains the calculation; the teacher illustrates their calculations by drawing on the bus timetable) 32:40 = 0.8"

Q_5: Do you have other ways to find relative frequency?"

$A_{5.1}$: "Find all the relative numbers (frequency) in each of the intervals and calculate the relative frequency (summarized)"

Q_6: "Compare the driving time of the two buses using relative frequency; which bus is better?"

$Q_{6.1}$: "[What if] we need to be at the airport within an hour?"

$Q_{6.1.1}$: "I have drawn two new diagrams (**Figure 5**)"

$Q_{6.1.1.1}$: "Which diagrams do you think make the most sense?"

$Q_{6.1.1.1.1}$: "It is difficult to get an overview of the bus timetable"

$A_{6.1.1.1.2}$: "You can see an intersection of the two graphs"

$A_{6.1.1.1.3}$: "In the bus timetable, you will get a lot of details. But in the graphs, you will get a picture of it all"

$Q_{6.1.1.2}$: "Do you know the name of the graphs?"

$A_{6.1.1.2.1}$: "Graph of frequency"

In the above Q&A diagram, we see how the teacher poses the problem of the day (Q_0), which leads to a series of sub-questions (Q_1… Q_6), in turn guiding the students to further develop more refined models to analyze the data. Each question requires one or several answers, which generate new questions, themselves requiring answers. There is a clear dynamic continuity between the questions derived (Q_1… Q_5) from finding the best bus, without clear criteria or reasons, to arguing for solutions under specific conditions and, finally, developing new statistical descriptors to serve in these arguments. Q_0 and Q_2 are the same question, but the milieu is different. Q_6 is a special case of Q_0, where Q_0 is reformulated to "What if we need to be at the airport within an hour?"

Question $Q_{2.2}$ draws on the students' established statistical techniques about frequency. These techniques consist of finding frequencies in the diagram. However, because the two sets of data do not have the same number of observations, the students need a new statistical technique–comparing relative frequency, followed by cumulative relative frequency. These new techniques are the main focus of the lesson, and we see several students as well as the teacher contributing towards explaining and illustrating the techniques (e.g. A2.2.1, Q2.2.2, and A2.2.2.1).

The tree diagram visualizes a teacher-guided inquiry lesson, where students meet statistical needs to develop their current statistical praxeologies. In particular, new techniques are discussed and illustrated in question-driven interactions between the students and teacher. In the lesson, the teacher writes questions (posed by him and the students) on the blackboard, along with the main elements of the students' answers, and the dialogue is represented by notes and illustrations (**Figure 5**).

Statistical Praxeologies

Like in the Danish case, Q_0 is the context of the inquiry. We can identify one main and two smaller statistical praxeologies of the students. The types of tasks are, as follows:

1. T: Compare two data sets based on grouped frequency tables
   a. T_1: Develop "new" techniques to compare two (grouped) data sets
   b. T_2: Evaluate statistical potentials in different statistical representations

The technique to answer T_1 was to first interpret the table and draw out the right information in the bus table in order to compare the frequencies. The students drew out the information (**Figure 4**), for example, on the interval between 25 and 30 minutes, 16 observations of bus numbers 17 and 18 and observations of bus number 18 ($A_{1.1}$ and $A_{1.2}$). They then concluded that bus number 18 was faster, probably because the frequency of the bus number was larger in the interval between 25 and 30 minutes than that of bus number 17. In the validation and justification of the technique (technology), the students and teacher concluded that the notion of frequency was not suitable to compare two data sets when the total number of observations of the two buses differed ($A_{2.2.1}$). The students further concluded (theory) that in order to use frequencies to compare the data sets, then the data sets would have the same number of observations ($A_{3.1}$). The later technique was an extension of the first part: to develop and validate new and more advanced techniques to compare the data sets (e.g. $A_{4.1}$). In the lesson, the students used ratio and proportion techniques and developed new techniques to

1. calculate the relative frequency: (relative frequency)=(frequency of a class) or (total frequency) and later (**Figure 5** lower right corner)
2. calculate the cumulative relative frequency: (cumulative relative frequency)=(frequency of several classes up to one class) or (total frequency).

In the lesson, several students were at the blackboard calculating and describing the process of finding relative and cumulative relative frequencies (e.g. $A_{2.2.1}$ and $A_{.4.1.1}$).

To solve the second task (T_2), the students explored new data visualized as a table and a frequency line graph (**Figure 5** to the right). They had no formal techniques to apply, so their techniques were mainly interpretations of what they "saw" in the two models, for example, the details in the table and an overall picture of the situation in the frequency line graph. The techniques used to solve T_2 was not developed further, and the students did not justify their answers, remaining at the practice level.

Though Q_0 questioned the everyday context, the task (T) was far more than finding the fastest bus. It was about the use of techniques, technology and theory when we compare data sets, validate comparisons and, further, use an appropriate representation when we want to visualize data.

DISCUSSING THE TWO CASES

We have analyzed two statistics classrooms, one from Denmark and the other from Japan. At first sight, the Danish class sported a student-oriented inquiry activity, while the Japanese case was a teacher-led chalk and blackboard lesson. Analyzing the cases in terms of question and answer dynamics and the statistical praxeologies that develop in these dynamics, we now engage in a more nuanced discussion. The two cases differed markedly in five ways as shown in **Table 1**.

Table 1. The main differences between the two cases

	Case of "experimental activity with many questions"	Case of "structured problem-solving"
1. Distribution of questions and answers	Lots of derived questions by students No final answers	Final answers produced by students, aided by teacher's sub-questions
2. Balance between generic aims and content items	Experimental activity, physical active students, outside classroom activity	Data are given Common analysis of the data
3. Statistical techniques	The techniques are given by the teacher Instrumented routine techniques	Analogue techniques developed by students
4. Oral and written support for technology	Oral dialogue	Oral dialogue and writing at blackboard and in notebooks
5. Relation to statistical theory	No statistical theories	Development of statistical descriptors supported by statistical theories

The first difference was the structure of questions and answers (**Figure 3** and **Figure 6**): In the Danish case, the students posed many questions but did not reach final answers. In the Japanese case, the dialogue was structured, and the students produced several answers, aided by the teacher's sub-questions. The second difference was the balance between generic aims and content aims. The Danish students engaged in an inquiry activity, were physically active and collected data outside the classroom. The Japanese students engaged in a common analysis of a set of constructed data. The third difference concerned how the students arrived at the techniques. In the Danish case, the techniques were provided by the teacher from the outset and included routine instrumental techniques. In the Japanese case, the students, aided by the teacher, developed new (analogue) techniques. The fourth difference concerned oral and written support for technology (discourse about techniques). In the Danish case, the statistical analysis was mainly oral. The students (likely) had individual written notes on their computers, but the notes were not shared. In the Japanese lesson, oral dialogue was sustained by the students' and teacher's notes on the blackboard and the students' notes in their notebooks. The fifth difference was the relation to statistical theory. In the Danish lesson, there was no formal statistical theory. The students questioned the validity of the trend line, and the teacher simply answered, "In my head, it makes sense. I see causality". In the Japanese lesson, the newly developed techniques were justified by the students, who further studied the formal definition of the technique in their textbook.

To become statistically literate, we cannot point only to the Danish or Japanese case as both cases have very different elements–inquiry-based elements. In an inquiry perspective, both posing questions and finding answers constitute qualities, and students' gathering of their own data and reflections about quality is as valuable as the analysis process. With this in mind, the question arises as to how to explain these differences. Our analysis is based on the paradidactic infrastructure that appears to influence the two classroom situations. We present five hypotheses, each of which is related to one of the above-mentioned differences.

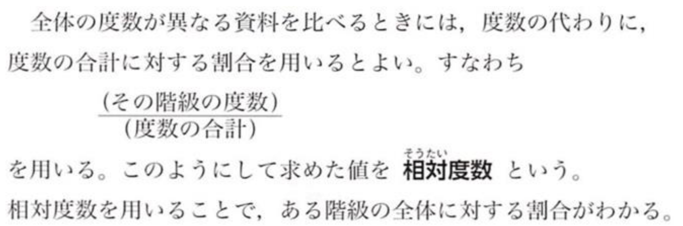
Hypotheses

In the Danish case, question Q_0 was a *genuine* question (Chevallard, 2015): The teacher did not know the answer in advance; the answer could not be found in a textbook; and the students had the opportunity to reflect rather than reproduce the thoughts of others. The present genuine question presented difficulties for the teacher, who could not predict what data the students would produce or how they would handle the data. The teacher might have been surprised by the data and the need to cope with the students' uncertainty, so he provided techniques to solve the task. In the Japanese case, question Q_0 and the data were provided for the students to determine the need to develop new statistical techniques, a technology and theory. Before the lesson, the teacher developed a detailed lesson plan in writing, in which he described the progression of the lesson, possible questions to pose and assumptions on the students' reasoning and answers. The plan framed the lesson, not as a rigid script but as support for the teacher, even in the face of unexpected situations. The practice of making and sharing lesson plans is a regular part of the Japanese paradidactic infrastructure (Jessen et al., 2019; Miyakawa & Winsløw, 2019), and such plans are shared with other teachers, for example, through national teacher journals (Miyakawa & Winsløw, 2019).

In the Danish case, in order to collect the data, the students spent nearly one-third of the lesson doing push-ups and jumping squats outside the classroom. The priority of collecting data could be explained by recommendations in the mathematical curriculum, such as "the student can explore connections in sets of data found in their everyday life". It was also likely to be a product of the generic aim to implement *physical activities* in all school subjects: The Danish school law states "that a school day must be organized in such a way that students are physically active for 45 minutes (average per day)" (Ministry of Education, 2017). Finally, it was also related to a Danish pedagogical trend, whereby *teaching outside the classroom* has been considered progressive (Bentsen & Jensen, 2012; Rea & Waite, 2009). Teaching outside the classroom is not mentioned in the Danish curriculum but represents a growing educational practice implemented in one-fifth of all schools (Barfod et al., 2016). The main point is that the teacher can tick off several generic aims and a pedagogical trend, in addition to the broad statistical aim of the curriculum. In the Japanese case, the focus was on the data analysis. The proceeding in the Japanese lesson was consistent with the curriculum recommendations: "…organize data according to their purpose, then read trend in the data", "organize data purposefully, using various tables and graphs and examining average distributions" and "help students understand the meaning and necessity of representative values and histograms and to be able to interpret trends in data by identifying and explaining them through these ideas" (Isoda, 2010, p. 21).

The techniques in the Danish and Japanese lessons differed in various ways. In the Danish case, the techniques were provided by the teacher and were mainly instrumental in nature. The instrumental techniques were the means of finding one possible answer to Q_0, namely, a linear correlation. In the Japanese case, the development of new analogue techniques was the objective of the lesson, and finding an answer to Q_0 was only a stepping-stone towards solving the real task T: Compare two (grouped) data sets. In the Danish case, the teacher had a strong personal interpretation of the curriculum goal, stating that "The student has knowledge about methods to elaborate connections between sets of data, including the use of digital tools" (Ministry of Education, 2019, p. 8). The teacher designed Q_0 with no references to a textbook. In fact, the textbook used by the teacher (Holte et al., 2009) contained no tasks about correlations or trend lines. In the Japanese case, the textbook strongly supported the teacher's interpretation of the curriculum. Textbooks are authorized for use by the Ministry of Education, and the authorization guarantees that the textbook follows the national curriculum and is considered of good quality by the Ministry of Education. The overall question (Q_0) from the Japanese lesson was strongly inspired by the textbook (Fujii et al., 2016), especially the page on relative frequency (**Figure 8**). This was evident in task T_2: Evaluating the statistical potentials of different statistical representations is comparable to questions and representations in the textbook (**Figure 7** and **Figure 8**).

Figure 7. Textbook, New mathematics 1 (Fujii et al., 2016, p. 206). Introducing the problem: "Which bus should we ride?"


全体の度数が異なる資料を比べるときには，度数の代わりに，
度数の合計に対する割合を用いるとよい。すなわち

$$\frac{(\text{その階級の度数})}{(\text{度数の合計})}$$

を用いる。このようにして求めた値を **相対度数** という。
相対度数を用いることで，ある階級の全体に対する割合がわかる。

Figure 8. Textbook, New mathematics 1 (Fujii et al., 2016, p. 211). Explaining relative frequency. Relative frequency=frequency of class/total frequency

The Japanese textbook in this case did not involve the use of digital tools, which potentially explains why all the techniques were analogue. Another explanation might be the different national examinations and entrance tests in the two countries: The Danish test requires students to use of digital tools, while the Japanese lower secondary school test relies exclusively on analogue techniques.

In Japan, the blackboard (**Figure 5**) plays a central role in the process of recording students' ideas in writing and visualizing the didactical process of the lesson. *Bansho* is a Japanese word meaning "(teachers') blackboard organization" and is a vital part of lesson planning (Tan et al., 2018). Most teachers in Japan take into account the teaching material used, the content of the lesson and predicted responses from students; they include these aspects in their Bansho during the lesson. Bansho is not used only as a means to display the teacher's knowledge; it is also used to share and connect students' inquiries and ideas. The teacher's work with Bansho is supported by a strong paradidactic infrastructure, which has led to a significant number of books and studies on Bansho techniques (e.g. Ikeno, 2013; Okamoto, 2018). Also, textbooks guide how to visualize possible Bansho (blackboard designs) for lessons (University of Tsukuba, Attached Elementary School Mathematics Department, 2016). In fact, students' mathematical writing is considered very important by Japanese teachers, not only on the blackboard but also in students' notebooks. In Japanese textbooks, we find chapters on "how to take notes", and the notebook works as a continuous memory for students. They are a clear part of the didactical infrastructure supporting students' mathematical writing. In the Danish case, there are some reminiscences from Grundtvig (1783-1872), a theologian-philosopher who has durably influenced the Danish culture of education. Grundtvig believed that teaching should not be based on books but on the living spoken word, with an extensive focus on stories and dialogue (Winter-Jensen, 2004). In Danish lower secondary schools, blackboards are often considered *old-fashioned* and associated with lecturing. In Denmark, there is no tradition, no available literature, where teachers can find help on using the blackboard.

Both the Danish and Japanese lessons can be compared with a *national script of lessons*. In the Danish lesson, the main focus was on students' inquiry processes and how the students, individually or in small groups, collected and explored the data in the research phase; also, the phase in which the students discussed their reflections and findings was more in the realm of "show and tell" than dialogic. The focus on processes and students' individual work is in line with other *European scripts* (e.g. Clivaz & Miyakawa, 2020). In the Japanese case, most of the lesson was a whole-class discussion where the students developed and discussed theory. In Japan, instructions do not prioritize particular students; the whole class is the focus. This practice is in line with *Japanese scripts* where the whole-class discussion is often the teacher's way of managing more than 30 students in the classroom (Clivaz & Miyakawa, 2020).

DISCUSSION AND CONCLUSION

Analyzing the two cases enabled an explication of certain institutional conditions, along with didactical and pedagogical theories, in two school systems and the prospect of at least partially understanding the strikingly different practices observed in lessons in the two systems.

In both the Danish and Japanese lessons, the students' activity took place at the center of didactical practice. Both classes explored a non-routine problem, made conjectures and experimented with and evaluated their answers. The lessons could thus be characterized as inquiry-based mathematics teaching in a broad sense (Artigue & Blomhøj, 2013). However, a closer look at the dialectics of the questions and answers reveals strong differences between the cases.

In the Danish lesson, the students spent nearly one-third of the time collecting data. They were physically active and got to go outside the classroom; they were critical about the collected data and posed many questions derived from this activity. However, the key statistical techniques (draw a table showing ordered pairs of values for each student, then plot and find a trend line) were provided by the teacher, and following an oral dialogue about the lines produced, the students ended up with few or no justified answers. In the Japanese lesson, more than half of the time was spend sharing and discussing answers. The data were deliberately designed to force the students to question and change their initial techniques, and the statistical knowledge (aimed at a definition and meaning of relative frequency) allowed them to provide justified answers to the questions posed by the teacher. The dialogue was structured around the students' statistical techniques, technology and theories and was framed by the teacher's sub-questions to the main question. In the lesson, the oral dialogue was accompanied by the students' and teacher's notes on the blackboard and the students' notes in their notebooks. The students developed answers (including new justified techniques) in the lesson, while the questions were mainly posed by the teacher.

The Q&A diagrams (**Figure 3** and **Figure 6**) visualized these differences, further outlined in **Table 1**. The Japanese lesson appeared as strongly structured by the teacher's main question and sub-questions, and nothing seemed to surprise the teacher: The students developed new techniques, along with technology, more or less according to the lesson plan. In the Danish lesson, the structure was much more open at first, almost "free play": The teacher posed the generating question along with a set of statistical techniques aimed at seeking answers. The students posed many ensuing questions, and the "answer" planned by the teacher (a regression line) was at best left unjustified.

We consider that Q&A diagrams could become an important analytical tool to obtain a more nuanced view of classroom situations that are considered, in some rough sense, "inquiry based". The main contribution of this paper is methodological in nature, the semantical analysis of the question and answer dialectics in such situations. We formulated explicit principles for the detection of questions and answers and demonstrated how the resulting Q&A diagrams can be used to visualize how "inquiry" can be variously conceived and evolve in a different manner.

Other differences emerged from the praxeological analysis regarding the plausible didactical justifications of the choice of tasks for the students, including the different balances between generic aims and content-specific aims. To an outsider of mathematics education, the generic aims of mathematics teaching (e.g. students posing critical questions or engaging in physical exercise) may seem more important than students thinking about technical details related to grouped and relative frequencies. It is also likely that blackboard writing and handwritten notes could be considered by outsiders as less progressive than the use of computer tools and a lively oral discussion. In the end, while the outcomes of the lessons were not directly comparable, they were categorically different, both in terms of the realized statistical praxeologies and the didactical techniques deployed.

One of the forces of ATD is its "verticality", that is, its focus on analyzing not just praxeological phenomena but also the institutional conditions that produce them. In this case, we did not stop at identifying differences in the dialectics of questions and answers or among the observed mathematical and didactical praxeologies; we proceeded to formulate five hypotheses about institutional conditions that may explain the differences.

Funding: The author received no financial support for the research and/or authorship of this article.

Acknowledgements: The author would like to thank to Japanese and Swedish colleges: Koji Otaki of the Hokkaido University of Education and Yukiko Asami Johanson of Gävle University. They were indispensable in the translation of Japanese texts, interview situations and classroom episodes into English. The author would also like to acknowledge Professor Carl Winsløw, University of Copenhagen, for being a great source of inspiration and many discussions and valuable comments.

Declaration of interest: The author declares no competing interest.

Data availability: Data generated or analyzed during this study are available from the author on request.

REFERENCES

Artigue, M., & Blomhøj, M. (2013). Conceptualizing inquiry-based education in mathematics. *ZDM*, *45*(6), 797-810. https://doi.org/10.1007/s11858-013-0506-6

Barfod, K., Ejbye-Ernst, N., Mygind, L., & Bentsen, P. (2016). Increased provision of udeskole in Danish schools: An updated national population survey. *Urban Forestry & Urban Greening*, *20*, 277-281. https://doi.org/10.1016/j.ufug.2016.09.012

Bentsen, P., & Jensen, F. S. (2012). The nature of udeskole: Outdoor learning theory and practice in Danish schools. *Journal of Adventure Education & Outdoor Learning*, *12*(3), 199-219. https://doi.org/10.1080/14729679.2012.699806

Ben-Zvi, D., Gravemeijer, K., & Ainley, J. (2018). Design of statistics learning environments. In D. Ben-Zvi, K.Makar, & J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 473-502). Springer. https://doi.org/10.1007/978-3-319-66195-7

Bosch, M., & Winsløw, C. (2016). Linking problem solving and learning contents: The challenge of self-sustained study and research processes. *Recherches en Didactique des Mathématiques* [*Research in Didactics of Mathematics*], *35*(2), 357-399.

Chevallard, Y. (1992). Fundamental concepts in didactics: Perspectives provided by an anthropological approach. *Research in Didactique of Mathematics: Selected Papers*, 131-168.

Chevallard, Y. (1999). L'analyse des pratiques enseignantes en théorie anthropologique du didactique [The analysis of teaching practices from the anthropological theory of the didactic]. *Recherches en Didactique des Mathématiques* [*Research in Didactics of Mathematics*], *19*(2), 221-265.

Chevallard, Y. (2015). Teaching mathematics in tomorrow's society: A case for an oncoming counter paradigm. In *Proceedings of the 12th International Congress on Mathematical Education* (pp. 173-187). Springer. https://doi.org/10.1007/978-3-319-12688-3_13

Clivaz, S., & Miyakawa, T. (2020). The effects of culture on mathematics lessons: An international comparative study of a collaboratively designed lesson. *Educational Studies in Mathematics*, *105*(1), 53-70. https://doi.org/10.1007/s10649-020-09980-1

Fujii, T., Matano, H., Amano, H, Ichikawa, S., Ohta, S., Ohtake, K., Ohtani, M. Okada, H. Ginnan, Y, Kudo, M., Kodera, T., Sakaori, F., Sasa, H., Shimizu, Y., Seino, T., Soeda, Y., Takahashi, H., Takahashi, H, Takeuchi, A. … Tokyo, S. (2012) *Atarashii Suugaku 1* [New Mathematics 1]. Tokyo Shoseki.

Holte, H., Thomsen, H., Weng, P., Andersen, M. W., & Lindhardt B. (2009). *KonteXt 8, kernebog*. Alinea.

Ikeno, N. (2013). Tairitsu to sentaku kara kanousei wo miidasu [The discovery of possibilities through confrontation and choice]. *Shakaika Kyouiku* [*Social Education*], *12*, 54-55.

Isoda, M. (2010). Junior high school teaching guide for the Japanese course of study: Mathematics (Grade 7-9). *CRICED, University of Tsukuba*. http://www.criced.tsukuba.ac.jp/math/apec/ICME12/Lesson_Study_set/Junior_high_school-teaching-guideMathmatics-EN.pdf

Jessen, B. E. (2014). How can study and research paths contribute to the teaching of mathematics in an interdisciplinary settings? *Annales de Didactiques et de Sciences Cognitives* [*Annals of Didactics and Cognitive Sciences*], *19*, 199-224.

Jessen, B., Otaki, K., Miyakawa, T., Hamanaka, H., Mizoguchi, T., Shinno, Y., & Winsløw, C. (2019). The ecology of study and research paths in upper secondary school. In M. Bosch, Y. Chevallard, F. Carcía, & J. Monaghan (Eds.), *Working with the anthropological theory of the didactic in mathematics education: A comprehensive casebook* (pp. 118). https://doi.org/10.4324/9780429198168-8

Lehrer, R., & English, L. (2018). Introducing children to modeling variability. In D. BenZvi, K. Makar, & J. Garfield (Eds.), *International handbook of research in statistics educatio*n (pp. 229-260). Springer. https://doi.org/10.1007/978-3-319-66195-7_7

Ministry of Education. (2017). *Bekendtgørelse af lov om folkeskolen § 15* [*Promulgation of section 15 of the primary and lower secondary school act*]. https://emu.dk/grundskole/varieret-undervisning/bevaegelse/bevaegelse-lovgivning-og-rammer

Ministry of Education. (2019). *Matematik, fælles mål* [*Mathematics, common goals*]. https://emu.dk/sites/default/files/2020-09/GSK_FællesMål_Matematik.pdf

Miyakawa, T., & Winsløw, C. (2019). Paradidactic infrastructure for sharing and documenting mathematics teacher knowledge: A case study of "practice research" in Japan. *Journal of Mathematics Teacher Education*, *22*(3), 281-303. https://doi.org/10.1007/s10857-017-9394-y

Okamoto, M. (2018) *Bansho, note-shido no kihon to idea* [*Basic idea of guidance for Bansho*]. Natsumesha.

Østergaard, C. H., & Winsløw, C. (2021). Sustainable lesson study and paradidactic infrastructure: The case of Denmark. *Hiroshima Journal of Mathematics Education, 14*, 1-15.

Pfannkuch, M., Ben-Zvi, D., & Budgett, S. (2018). Innovations in statistical modeling to connect data, chance and context. *ZDM, 50*(7), 1113-1123. https://doi.org/10.1007/s11858-018-0989-2

Rea, T., & Waite, S. (2009). International perspectives on outdoor and experiential learning. *International Journal of Primary, Elementary and Early Years Education. International Perspectives on Outdoor and Experiential Learning, 37*(1), 1-4. https://doi.org/10.1080/03004270802291699

Stigler, J. W., & Hiebert, J. (1998). Teaching is a cultural activity. *American Educator*, *22*(4), 4-11.

Tan, S. Fukaya, K., & Nozaki, S. (2018). Development of bansho (board writing) analysis as a research method to improve observation an analysis of instruction in lesson study. *International Journal for Lesson and Learning Studies, 7*(3), 230-247. https://doi.org/10.1108/IJLLS-09-2021-0076

Tukey, J. W. (1977). *Exploratory data analysis.* Addison Wealey.

University of Tsukuba, Attached Elementary School Mathematics Department. (2016) *Bansho de miru zentangen, zenjikan no jugyo no subete* [*Looking at the lessons through bansho*]. Toyokan Shuppansha.

Wild, C. J., & 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

Winsløw, C. (2011). A comparative perspective on teacher collaboration: The cases of lesson study in Japan and of multidisciplinary teaching in Denmark. In G. Gueudet, B. Pepin, & L. Trouche (Eds.), *From text to "lived" resources* (pp. 291-304). Springer. https://doi.org/10.1007/978-94-007-1966-8_15

Winsløw, C., Matheron, Y., & Mercier, A. (2013). Study and research courses as an epistemological model for didactics. *Educational Studies in Mathematics, 83*(2), 267-284. http://doi.org/10.1007/s10649-012-9453-3

Winter-Jensen, T. (2004). Comparative education in a world of globalization. *World Studies in Education, 5*(2), 81-94. https://doi.org/10.7459/wse/05.2.06

Yin, R. K. (2013). Validity and generalization in future case study evaluations. *Evaluation*, *19*(3), 321-332. https://doi.org/10.1177/1356389013497081