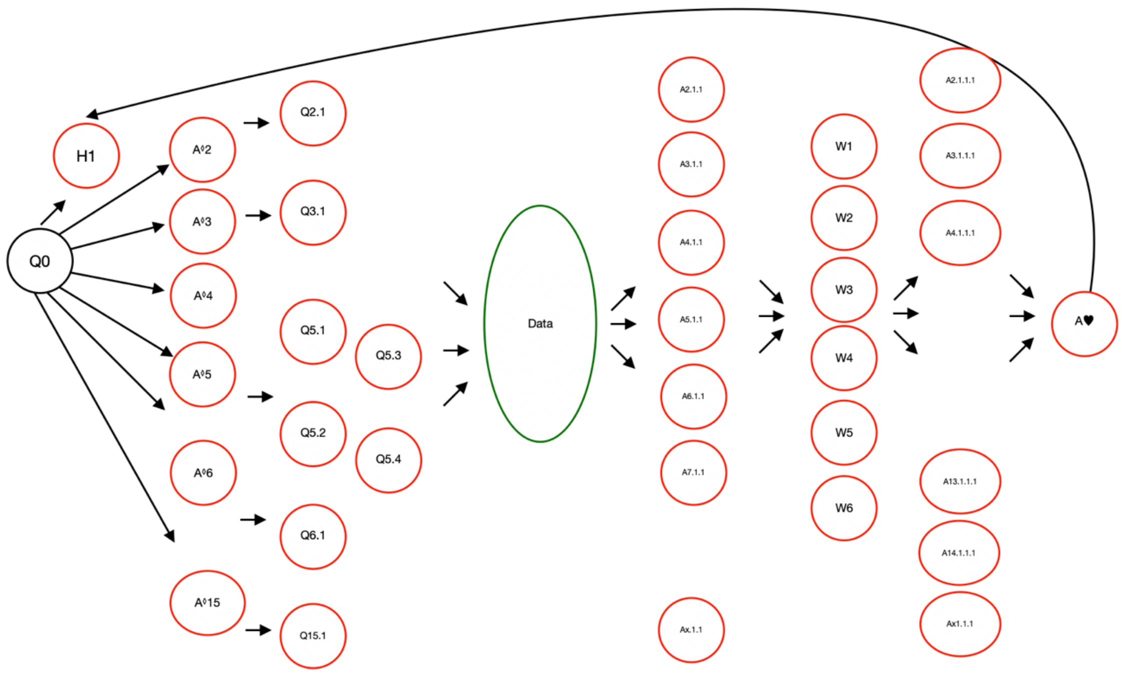
UNIVERSITY OF COPENHAGEN
DEPARTMENT OF SCIENCE EDUCATION



PhD Thesis

Camilla Hellsten Østergaard

Implementing and analysing statistical inquiry approaches in primary and lower secondary school statistics

Professor, Carl Winsløw

Submitted on: 28 January 2023

Implementing and analysing statistical inquiry approaches in primary and lower secondary school statistics

Doctoral Dissertation 2023

Camilla Hellsten Østergaard

Department of Science Education

University of Copenhagen

Index

i Acknowledgements

In my PhD project, I have been privileged to gain several peoples' goodwill.

Thanks to the mathematics teachers and students who kindly agreed to participate in my research projects.

Thanks to my Japanese colleagues, Yukiko Asami-Johansson from the University of Gävle and Koji Otaki from the Hokkaido University of Education for our theoretical discussions, for inviting me to visit Japanese mathematics classrooms, and for introducing me to Japanese teachers. I hope that, in the future, we can complete the research ideas that we started to implement but had to abandon because of the worldwide pandemic.

Thanks to Dorte Moeskær Larsen from the Southern University of Denmark for all kinds of support, for sharing crazy research ideas and for co-authoring papers.

Thanks to the University College Copenhagen, my teacher educator colleagues and my co-supervisors, Charlotte Krog Skott and Klaus Rasmussen, who have supported me in the beginning of this PhD project. A special thanks to Klaus, with whom I co-authored my first paper.

It has been fantastic to be a part of the Department of Science Education at the University of Copenhagen. I want to thank Britta E. Jessen, who always had time to share experiences—the good and bad. Thank you to the PhD group in mathematics education—Louise, Rong Rong, Pia, Mayo and Derya; it has been fun to travel with you.

A very special thanks to my supervisor, Professor Carls Winsløw, for the many, many hours of advice and patience. It has been a pleasure to be supervised by you.

Last, but not least, thanks to Mikael, Roberta, Olga, Dagmar and Solvej; you are more than I could have dreamed of!

ii Abstract

Statistics plays a unique role in how school mathematics contributes to citizenship. Students meet statistical claims and analyses in both school and in society at large, so one aim of school mathematics is to prepare students to relate critically to such claims and even engage themselves in statistical work. This thesis examines a variety of statistical inquiry approaches, realised in primary and lower secondary schools. One of these inquiry approaches is the relatively new approach from the anthropological theory of the didactic, called the *paradigm of questioning the world*. In this approach, learning begins with questions that are in themselves important (to students, to citizens). The main purpose of the study is to develop answers to such questions and also, more generically, to develop an inquiry stance among students.

The thesis draws on several closely related research projects: two large-scale Danish research and developmental projects, a study of paradidactic infrastructure and sustainable lesson study in Denmark, a comparative study of how statistical inquiry was implemented in two lessons taught in Denmark and Japan, and a design research project about how study and research paths (SRP) can contribute to primary school students' development of statistical knowledge. In our research, we present and expand the theoretical idea of paradidactic infrastructure. Paradidactic infrastructure affects the implementation of a statistical inquiry in several ways and, taking these influences explicitly into consideration, is essential when trying to implement new practices for teachers' professional development.

Due to the variety of the research projects, we applied different methodologies, such as case studies and different formats of design research, which involve more or less extensive involvement of teachers' experiences and observations in efforts to improve the SRP design. The data gathered in the projects were qualitative in nature and consisted of classroom observations, field notes, teacher interviews, contextual information about participating teachers, institutions and paradidactic elements (such as curricula, textbooks and teachers' opportunities to participate in professional development activities).

Regarding theoretical frameworks, we mainly drew on two major theoretical paradigms of rather different provenance: the anthropological theory of the didactics from mathematics education research, and cultural-historical activity theory which originates in cognitive

psychology and its adaptations to educational studies. We modelled students' participation in statistical inquiry in question-and-answer diagrams and as statistical praxeologies. We analysed the conditions and constraints that affect statistical inquiry classrooms and the paradidactic infrastructure that supports teachers' professional development. We further modelled teachers' participation in research and developmental projects as interacting activity systems and as a process of expansive learning.

The overall research results indicate how statistical inquiry approaches can support students' development of statistical knowledge. Most notably, the dialectic of students' study and research of statistical knowledge in a digital milieu enables sustained, recursive sequences of statistical inquiry to be developed with primary school students. In our studies, we saw how statistical inquiry within such a dynamic allowed the students to autonomously develop statistical knowledge, and in the SRP approach, we observed that students were encouraged to use and develop theoretical ideas related to statistical practices. The SRP contrasted with common forms of statistics teaching in which statistical techniques are demonstrated and practiced but not further elaborated or questioned.

The ecology of statistical inquiry consists of several conditions and constraints. Among the favourable conditions one can count a mathematical curriculum that endorses and supports statistical inquiry approaches, and also resources, such as textbooks and digital tools, that scaffold implementation of inquiry in the classroom. Unsurprisingly, the role of the teacher is also crucial, for example, in establishing a new didactic contract. However, teachers' autonomy (e.g. to choose textbooks and decide how to organise classroom teaching) can also act as a constraint, for instance when a detailed didactic design for inquiry is to be implemented.

We observed a variety of assumptions about the nature of statistical inquiry that are held by teachers and researchers, as well as the difficulty of achieving truly shared views. We question specific views of interaction between researchers and teachers and investigate the hypothesis that research could work as an integrated part of paradidactic infrastructure, such as in the format of lesson study and other forms of teacher–researcher collaborations. It would be interesting to further investigate how such paradidactically integrated research can lead to new sustainable statistical inquiry practices, and to what extent external support is needed after the research intervention stops.

iii Resume

Statistik spiller en særlig rolle i, hvordan grundskolematematik bidrager til udvikling af demokratiske medborgere. Et af målene i skolens matematikundervisning at forberede elever til arbejdet med statistik, herunder at udvikle statistisk literacy og statistisk ræsonnement, og derved udstyre dem med forudsætninger for at forholde sig til data og udsagn baseret på data, såvel i skolesammenhænge som ude i verden.

Denne ph.d.-afhandling bidrager med viden om metoder til at engagere eleverne i statistisk undersøgelse (*inquiry*) i grundskolen. En af de tilgange, som udforskes, er det relative nye bidrag fra antropologisk teori for didaktik, paradigmet om at stille store 'spørgsmål til verden'. 'Spørgsmål til verden' tager sit udgangspunkt i spørgsmål, som er af generel interesse for alle, og hovedformålet er at udvikle svar og afledte spørgsmål, og fra et bredere perspektiv, at udvikle elevernes dispositioner og forudsætninger for at indtage en spørgende, kritisk og undersøgende position ift. data og databaserede udsagn.

Afhandlingen bygger på flere sammenhængende forskningsprojekter: to danske stor-skala forsknings- og udviklingsprojekter, forskning med fokus på paradidaktisk infrastruktur og levedygtige lektionsstudier i Danmark, komparativ-deskriptiv forskning af hvordan statistisk inquiry blev implementeret i to lektioner i henholdsvis Danmark og Japan, og et designforskningsprojekt om hvordan studie og forskningsforløb (SFF) kan bidrage til grundskoleelevers udvikling af statistik viden.

Vores metodologiske tilgange i forbindelse med disse projekter er forskellige, men omfatter særlig casestudier og designforskning. I vores projekter arbejder vi desuden med forskellige paradigmer for designforskning, både klassisk "implementation" og aktionsforskningsformater, hvor læreres inddragelse, erfaringer og observationer er med til at udvikle et SFF. Vores forskningsdata er kvalitative og består af noter fra klasserumsobservationer, videooptagelser, lærerinterviews, kontekstuel information omkring deltagende lærere, institutioner og paradidaktiske elementer som lærebøger, læreplaner og dokumentation vedr. læreres mulighed for at deltage i efteruddannelsesaktiviteter.

Projektets teoretiske rammer kommer fra to ret forskellige forskningsprogrammer: antropologisk teori om det didaktiske (fra forskningsområdet matematikdidaktik) og kulturhistorisk aktivitetsteori fra grænseområdet mellem psykologi og uddannelsesvidenskab. Derudover trækker vi på viden og begreber fra forskning i statistikkens didaktik.

I vores forskning modellerer vi inquiry processer med spørgsmål-og-svar diagrammer og analyserer, hvilke statistiske praxeologier elever udvikler. Vi undersøger ligeledes betingelser og begrænsninger, som påvirker undervisningen i statistisk, og den paradidaktiske infrastruktur, som understøtter læreres professionelle virke og udvikling. Derudover modellerer vi læreres deltagelse i forsknings- og udviklingsprojekter som interagerende aktivitetssystemer og som expansiv læring.

Desuden præsenterer og udvider vi det teoretiske begreb paradidaktisk infrastruktur. Paradidaktisk infrastruktur påvirker realiseringen af statistisk inquiry på flere måder. Det er vigtigt at tage disse påvirkninger i betragtning, når man overvejer fremtidige strategier for at implementere nye professionelle udviklingsmuligheder for lærere.

Projektets forskningsresultater handler bl.a. om, hvordan de omtalte inquiry tilgange på forskellig vis kan støtte elevers udvikling af statistisk viden. En specifik observation er, at dialektikken mellem elevers studie af medier og deres udforskning af spørgsmål i et digitalt miljø, faktisk kan muliggøre længerevarende og substantielle statistiske undersøgelser i grundskolesammenhæng. I vores forskning observerede vi, hvordan statistik inquiry muliggjorde elevers udvikling af autonom statistisk viden, og i SFF'et så vi, hvordan elever blev opmuntret til at bruge og udvikle statiske ræsonnementer. SFF adskilte sig fra andre undervisningstilgange ved ikke udelukkende at have fokus på statistiske teknikker, men også ræsonnementer, herunder hvordan teknikker bruges, og hvorfor teknikkerne er meningsfulde i den givne kontekst.
Økologien omkring statistisk inquiry består af flere betingelser og begrænsninger. En af de gunstige betingelser er en læreplan i matematik, som eksplicit foreskriver statistiske inquiry tilgange, og desuden resurser, såsom lærebøger og digitale redskaber, der stilladserer realiseringen af inquiry i undervisningen. Ikke overraskende er lærernes rolle afgørende, herunder i etablering af en ny didaktisk kontrakt. Lærernes autonomi viser sig også i nogle

sammenhæng som forhindringer, eksempelvis, når lærernes frihed til selv at vælge lærebøger eller organisering af klasserum støder sammen med detaljerede undervisningsdesigns som skal implementeres.

Vi har observeret en række forskellige måder at planlægge og gennemføre statistisk inquiry i klasserum. I vores forskning sætter vi spørgsmål ved samspillet mellem lærere og forskere, og undersøger særlig den hypotese, at forskning kan integreres i paradidaktiske infrastrukturer, såsom lektionsstudier eller andre former for lærer-forsker-samarbejde. Det vil have forskningsmæssig interesse at undersøge, hvordan sådanne undersøgelser kan medføre mere levedygtige statistiske inquiry praksisser, og i hvilken form ekstern støtte er påkrævet efter endt forskningsmæssig intervention.

iv Short abstract

Statistics plays a special role in how school mathematics contributes to citizenship. This thesis contributes insight into the conditions for realising statistical inquiry in primary and lower secondary schools. We examined students' development of statistical praxeologies within various inquiry approaches, the ecology of statistical inquiry, and the paradidactic infrastructure needed to support teachers' professional development and their implementation of statistical inquiry. A variety of the research projects were drawn on by this thesis, with partly different methodologies, such as case studies and different formats of design research. In terms of theoretical frameworks, we drew on two rather different paradigms: the anthropological theory of the didactic, and cultural-historical activity theory. In our research, we modelled students' participation using question-and-answer diagrams and as statistical praxeologies. We analysed the conditions and constraints that affect statistical inquiry in classrooms and the paradidactic infrastructure that supports teachers' professional development. We modelled teachers' participation in interacting activity systems and as a process of expansive learning. We observed a variety of assumptions about the nature of statistical inquiry held by teachers and researchers. We also demonstrate the potential of study and research paths developed through lesson study or other forms of teacher–researcher collaboration, for leading to sustainable statistical inquiry practices.

Keywords: anthropological theory of the didactic, paradigm of questioning the world, statistical praxeologies, ecology, paradidactic infrastructure

v Kort resume

Statistik spiller en særlig rolle i, hvordan grundskolematematik bidrager til udvikling af demokratiske medborgere. Denne ph.d.-afhandling bidrager med indsigt i hvordan statistiske undersøgelser (inquiry) kan realiseres i grundskolen og de betingelser og begrænsninger, der muliggør dette. Vi undersøger elevers udvikling af statistiske praxeologier, den institutionelle økologi for undersøgende statistik og den paradidaktiske infrastruktur, som er nødvendig for at støtte lærerne i deres professionelle udvikling og i realisering af inquiry. I de forskellige forskningsprojekter anvender vi forskellige metodologier, herunder casestudier og forskellige formater for designforskning. Vi trækker på to teoretiske paradigmer: Den antropologiske teori om det didaktiske, og kulturhistorisk aktivitetsteori. I vores forskning modellerer vi elevers deltagelse i spørgsmål-og-svar diagrammer og som statistiske praxeologier. Vi analyserer betingelser og begrænsninger, som påvirker undervisningen i statistik, og den paradidaktiske infrastruktur, som understøtter læreres professionelle udvikling. Derudover modellerer vi læreres deltagelse i forsknings- og udviklingsprojekter som interagerende aktivitetssystemer og som expansiv læring. Vi har observeret en række forskellige strategier for at realisere undersøgende statistisk i undervisning. Vi foreslår yderligere udvikling af paradidaktisk infrastruktur, såsom lektionsstudier eller andre former for lærer-forsker samarbejde, til etablering og udvikling af undersøgende statistiske praksisser.

Nøgleord: antropologisk teori om det didaktiske, paradigme omhandlende spørgsmål til verden, statistiske praxeologier, økologi, paradidaktisk infrastruktur

vi List of publications

vi.i Papers included in the thesis

Paper I:

Larsen, D. M., Østergaard, C. H., & Rasmussen, K. (2022). Contradicting activity systems–learning from large-scale interventions that fail to change mathematics teaching practice as intended. *Journal of Mathematics Teacher Education*, 1-24.

Accepted, and published online (August 2022)

Paper II:

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

Paper III:

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.

Paper IV:

Østergaard, C. H. & Larsen, D. M. (submitted). Questioning the world – More than inquiry. *Statistics Education Research Journal*

Submitted: October 2022

vi.ii Other papers from the PhD period

Larsen, D. M. & Østergaard (2019). Questions and answers… but no reasoning! In U. T. Jankvist, M. van den Heuvel-Panhuizen & M. Veldhuis (Eds.) *Proceedings of the Eleventh Congress of the European Society for Research in Mathematics Education (CERME11).* Utrecht, the Netherlands: Freudenthal Group & Freudenthal Institute, Utrecht University and ERME.
Available online: https://hal.science/hal-02398122/document

Østergaard, C. H. (2019). Questioning the world – Statistical inquiry. In S. Budgett (Ed.), *Decision Making Based on Data Proceedings of the Satellite conference of the International Association for Statistical Education (IASE)* (pp. 1-6), Kuala Lumpur, Malaysia.
Available online: https://web.archive.org/web/20200724185727id_/https://iase-web.org/documents/papers/sat2019/IASE2019%20Satellite%20186_STERGAAR.pdf?1569666570

Skott, J., Larsen, D. M., & Østergaard, C. H. (2020). Learning to teach to reason: Reasoning and proving in mathematics teacher education. In: S. Zehetmeier, D. Potari & M. Ribeira (Eds.), *Professional Development and Knowledge of Mathematics Teachers* (pp. 44-61). Routledge.

Østergaard, C. H. & Larsen D. M. (submitted 2023). Questioning the world – The dialectic of study and research. *Congress of the European Society for Research in Mathematics Education* (No. 13). ERME.

Larsen, D. M. & Østergaard, C. H. (submitted 2023). Two inquiry approaches to STEM: The role of mathematics. *Congress of the European Society for Research in Mathematics Education* (No. 13). ERME.

vii List of abbreviations

The intention with the list of abbreviations is to support readers and provide reference to frequently used abbreviations.

A	Answer
A$^{\Diamond}$,	Existing answer
A$^{\heartsuit}$	Final answer
ATD	Anthropological theory of the didactics
CHAT	Cultural-historical activity theory
D	Data (e.g. students' own data)
Demo	Demonstration school project
H	Hypothesis
ICT	Information and communication technology
Kid	Quality in Danish and mathematics [Kvalitet i dansk og matematik]
LS	Lesson study
MKT	Mathematical knowledge for teaching
PD	Professional development
PI	Paradidactic infrastructure
PID	Paradidactic infrastructure in Denmark
PQW	Paradigm of questioning the world
PVM	Paradigm of visiting monuments
Q	Question
Q_0	Generating question
Q&A	Question and answer
RCT	Randomized control trial
RME	Realistic mathematics education
RQ	Research question
S	System
SKT	Statistical knowledge for teaching
SRP	Study and research paths
SRP-TE	Study and research paths in teacher education
T	Type of tasks
W	Works (mathematical or statistical content)
X	Students
Y	Teacher

1 Prelude: The subject of this thesis

What is the purpose of general education, and how can mathematics contribute to it? This is certainly a vast and complex question and one that is usually left in the background of mathematics education research. Statistics plays a special role in how contemporary school mathematics contributes to citizenship. The statistics education literature emphasises these general contributions through the use of terminology such as "statistical literacy", which suggests a parallel to literacy in the normal sense, thus underlining the importance of statistics to citizens in modern society. Indeed, students will meet statistical analyses and claims, both in the news media and in other school subjects, so it becomes one of the aims of school mathematics to prepare them to "read" statistical analyses critically and engage in elementary forms of statistical work.

The overall discussion of the goals and means for the general education of citizens is very old. Philosophers from Socrates to Rousseau have criticised the common view of education as mere transmission of established knowledge, advocating for various forms of autonomy and discovery as the necessary elements of education (for a more elaborate discussion of these historical viewpoints, see Bosch & Winsløw, 2016). In more recent research on mathematics education, these general visions are often translated into experiments and curriculum reforms that emphasise "inquiry" or "problem solving" (Artigue & Blomhøj, 2013). A different but related vision is related to mathematical modelling as a central aspect of school mathematics. Mathematical modelling has often been described as "translations between reality and mathematics" (Blum & Fermi, 2009, p. 45), and modelling competence—the corresponding educational goal—as "a person's insightful readiness to autonomously carry through all aspects of a mathematical modelling process in a certain context and to reflect on the modelling process and the use of the model" (Blomhøj & Jensen, 2003, p. 127). This vision of modelling tends to emphasise model building that is driven by theoretical assumptions and methods, such as setting up a system of differential equations to describe the evolution of interacting populations; the model is then confronted with (as confirmed by) data from experiments or observations designed for the model. However, in both the natural and social sciences, modelling can also start with data, and statistics (together with computing technology) are central in generating models from these. This is what Dvir and Ben-Zvi (2018) call *data modelling*; they describe *statistical modelling* and a confrontation of data models with *conjecture models*, with a kind of dialectics between the two (that are continuously compared). The authors contend that "creating more modeling opportunities, and explicitly asking students to compare models, can help them progress within a specific statistical investigation, as well as in developing their conceptual statistical modeling skills" (Dvir & Ben-Zvi, 2018, p. 1196). In other words, engaging students in real dialectics between models built from data and models emerging from conjectures (possibly just related to naïve expectations) could be crucial in developing students' grasp of the full scope of statistical modelling. Naturally, establishing situations in which school children could be involved in some form of "statistical investigation" requires specific tools to characterise, plan and implement lessons that would depart, both in form and content, from more common forms of mathematics teaching.

Chevallard (2006, 2015) proposes an approach to investigation or inquiry in comprehensive schools examining the full question about the role that mathematics (and with it, implicitly, statistics) can and should play in the education of citizens. He names two opposite forms of education: *the paradigm of questioning the world* (PQW) and *the paradigm of visiting monuments* (PVM). The PVM reflects the conception of education as the delivery of established knowledge, "monuments", which are then merely being "visited" without any deeper appreciation of their origin or meaning. These knowledge monuments are answers to questions that are not raised explicitly, are not met by the learners or "visitors" and that may have been forgotten altogether. Conversely, as the starting point for inquiry, PQW begins with questions that are of general importance to citizens. The main purpose of the study and an end in itself is elaborating answers to questions (Barquero & Bosch, 2015) and, from a wider perspective, developing students' aptitude for questioning or their "herbartian attitude" (cf. Corica & Otero, 2015, p. 146). This is in contrast to other so-called "inquiry" approaches where the main aim is to develop generic strategies for "figuring out", rather than to find a good answer to the question (Lampert, 1990), or where the initial question is deliberately constructed in an attempt to make students reach a predefined answer (Brousseau, 1997). In other words, what is called "inquiry" in mathematics education may in fact often be a specific form of PVW, where questions are asked for students to reach and visit a given answer.

Chevallard (2015) suggests that, in the PQW, the curriculum would consist of "big" questions that all citizens should be engaged in investigating and answering rather than mere answers left over from long-forgotten inquiries. These questions should be of general and vital importance to all citizens. To answer them would, in many cases, go beyond the borders of traditional school disciplines. Although mathematics in some form may often be involved, the process of answering big questions may rarely be confined to mathematics. That being said, when mathematics is involved, it will very often also engage statistical methods.

Indeed, statistics, especially when supported by technology, seems to be particularly fit to contribute to contemporary general education according to PQW. Our hypothesis is that, if one wants to experiment and develop PQW in current institutions and involve mathematics teachers centrally in that effort, then questioning involving data analysis could be a very promising avenue. Conversely, considering statistical literacy as an increasingly central aim of general education (Bargagliotti et al., 2020; Franklin et al., 2007) and as the needs for teaching statistical modelling outlined above, the specific tools associated with PQW, especially study and research paths (SRP; see Section 2.2.2), seem promising.

Thus, the research area of the present thesis is situated in the intersection between statistics education and research into the conditions and constraints for engaging students in some form of inquiry, here in the context of comprehensive school mathematics. Both conditions and constraints are highly specific to the institutional and national contexts. In the Danish context, from which the current thesis draws most of its data, very little research has focused on the intersection considered here. An early example of scholarship in this direction is the licentiate thesis by L. Rasmussen (1987), in which he develops what he calls a propaedeutic stochastical research programme for lower secondary school students. Rasmussen classifies three problematic forms of student behaviour in relation to solving school problems in statistics:

- Blind use and application of rules
- Looking for “signal words”
- Ignoring data (Rasmussen, 1987, p. 111)

(quoting here and below from the abovementioned thesis, using the author’s translation from the original text in Danish). Rasmussen further states, “It is absolutely awful that the area of mathematics that is supposed to be a model of good practice [he refers here to probability and statistics, is transformed into a system of rigid rules]…" (ibid., p. 123), and he argues that students do not like systems and endless drills that involve practising statistical procedures on problems that appear unimportant. Rasmussen's programme can be characterised as didactical designs (without any trace of implementation found in the thesis). The proposed collection of teaching activities offers "learning opportunities for students to relate to, and act in stochastic problem solving (ibid., p. 444) and further give the students the opportunity to study meaningful problems in a rreal-lifecontext …" (ibid, p. 439).

In the period from 1985 to 1990, a group of researchers from the Royal Danish School of Education had funding to distribute more than a hundred computers to four different schools, here in relation to a long-term research project studying the "computers' impact in the development of mathematical knowledge and the teaching of mathematics" (Infa.dk, 1985). Parts of the programme focused on statistics education: the development of teaching material heavily drawing on computer programmes developed by the research team. However, no international publications arose from the project, but the materials and software developed in the context of this project remained influential and widely disseminated in Danish schools for decades.

More recently, a larger national development project involving all school subjects was initiated by the Danish government in 2013, with the vision to develop a stronger primary and lower secondary school, where all students learn more, including gaining knowledge about twenty-first-century skills and real-world problem solving (Demonstrationsskoler.dk., 2015a, 2015b). In this project, I took part in a subproject called "Youngsters and ICTs", focusing on dialogue-based teaching and the integration of digital tools for statistics teaching. Together with Charlotte Skott, we designed a sequence of lessons on statistical reasoning in the context of grade 6 mathematics. This is where I began my journey in statistics education, statistical inquiry approaches, implementation projects and systematic teacher development. The lesson design was inspired by Cobb and McClain's (2004) principles of instructional design for statistical reasons, focusing, for instance, on significant statistical ideas, computer-based tools used by the students, and the classroom discourse. In our own project, the central aims were to get teachers to use digital tools as a vehicle for initiating, negotiating and establishing new statistical practices in the classroom and for students to investigate tendencies and patterns in datasets and relate to these in a critical manner to the use of statistics. To achieve these aims, the project established general and lesson-specific prescriptions for the ways in which teachers should engage students in statistical investigations.

A main outcome of our studies concerns specific difficulties that the teachers experienced with establishing the classroom norms and student practices explicitly prescribed in our design, especially with implementing the digital tools according to the design (Skott, 2018; Skott & Østergaard, 2016, 2018). These results and experiences have motivated me to explore in the present thesis how teachers can be involved in more fruitful ways in these interventions and how this that could affect the impact of the intervention on both students' learning and teachers' perspectives of the design and actual uses of it. In other words, we considered that more attention is needed to teachers' role in the design process if the design should lead to them revising their classroom practices (Hegedus & Tall, 2015). With this in mind, we continued to gather data from one "Demo school" after the project had ended, and some of that data have been used in Paper I. At the same time, from 2016 to 2019, another national development project (KiDM.dk, 2017) implemented inquiry approaches, which was partially related to school statistics. Data from that project are also used in Paper I. These and other projects have convinced me that, when investigating the conditions and constraints for realising innovative didactical designs, it is crucial to take into account the conditions and constraints under which teachers operate, including what we will call *paradidactic infrastructure* (see Section 3.3). We must also consider the new conditions and constraints imposed on teachers by the ways in which the new designs are conceived and realised; they could offer opportunities for teachers' professional development and continued use of the designs or merely create a temporary window for implementing the designs.

To sum up, my research project aims to investigate the conditions and constraints for realising statistical inquiry in the context of Danish comprehensive school mathematics while paying special attention to the role and means of teachers. Figure 1 illustrates how the four papers of the present thesis are situated with respect to the six overall research topics (to be further outlined in Section 3 in terms of delimitation and relevant research background). Concerning the theoretical frameworks, I draw mainly on mathematics education research and, more specifically, on two major paradigms: anthropological theory of the didactics and cultural-historical activity theory (to be further described in Section 2).

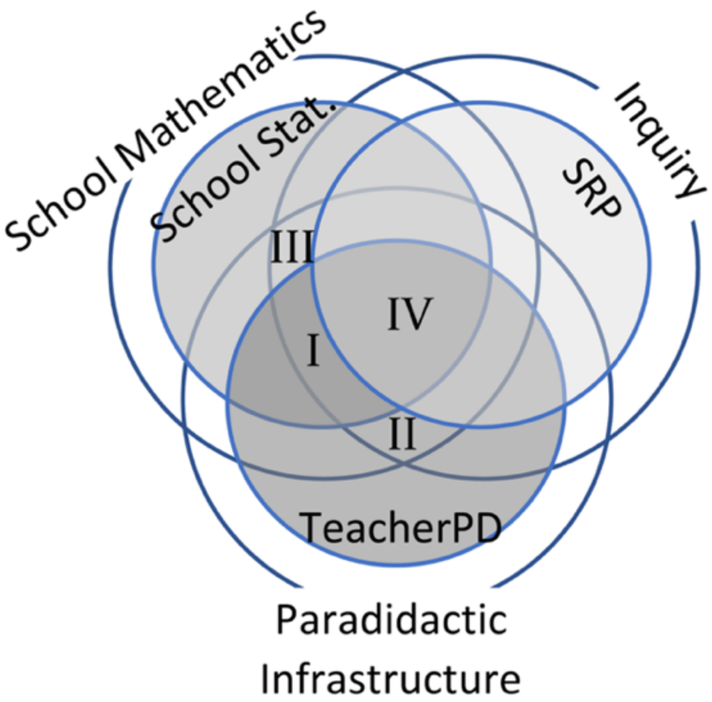


Figure 1: Model illustrating how the four papers are situated with respect to the six overall research topics

The rest of this introduction outlines the theoretical frameworks employed in the present thesis, background from connected research in mathematics and statistics education, the research questions and methodology for the overall project, a discussion of the overall answers emerging from the four papers and a conclusion.

2 Theoretical frameworks

In the papers comprising the core of the present thesis, we use two theoretical frameworks: cultural–historical activity theory (CHAT, used in Paper I) and the anthropological theory of the didactics (ATD, used in Papers II, III and IV). This is mainly a result of the development in my research interests, from an initial focus on the structural aspects of teachers' activity in connection with a large-scale implementation-oriented development project (for which CHAT is used) to more content-oriented studies of the conditions and constraints for inquiry-based teaching of mathematics, especially statistics (based on ATD). In this section, we first describe these two frameworks separately because they function independently both as research programmes and in the papers in the current thesis. Then, we supply a short discussion of their compatibility and potential complementarity.

2.1 Cultural–historical activity theory

CHAT (see Engeström, 2001; Engeström & Sannino, 2010) is a framework for analysing the processes and activities in educational practices. CHAT offers a number of affordances and priorities. First, CHAT offers an object-oriented analysis of individual and collective activity in those activity systems consisting of *nodes* (agents and other elements engaged in the activity) and interactions and other relations between the nodes. Second, CHAT allows us to identify and question the different nodes of the activity and tensions between them. Third, CHAT is a conceptual tool used for understanding multiple perspectives, for instance, by studying a whole network of interacting activity systems (e.g., each corresponding to professional groups), capturing knowledge that spans different settings, and giving us an understanding of the development and boundaries of practices across institutions, for example, school and teacher education. Fourth, CHAT offers a framework to analyse how didactical designs are implemented and how implementation must be understood in relation to a wider network of activity systems than the one in which the implementation of statistics takes place (see Section 3.3).

2.1.1 Activity systems

Activity systems are often represented by a triangular illustration containing six nodes (Figure 2).

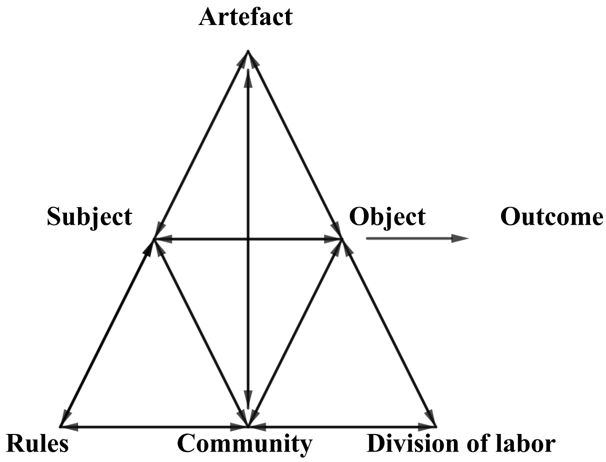


Figure 2: General model of the activity system. (adapted from Engeström, 1987)

The different nodes are as follows:

- *Subject* representing individuals or a group of people participating in an activity, such as a teacher, a teacher educator or a researcher. The subject is the perspective of the analysis.
- *Object* representing the objective of the subject, such as students learning statistics. The outcome represents the transformed object.
- *Artefacts (or instruments)* mediate action (Vygotskij, 1978). Artefacts can be physical (e.g., curricula, textbooks, lesson plans, digital tools and black boards) or mental (e.g., procedures and learning theories). The artefact affects the way the subject carries out an action and how the subject develops the artefact to make it more effective and useful, which again changes the way the subject carries out the action (Kaptelinin et al., 1995).
- *Community* refers to a wider group of individuals who share cultures. The members of the community negotiate the *division of labour* and *rules* for participation. The rules include explicit or implicit regulations and norms that constrain the actions within the activity system (Engeström & Sannino, 2010).

Connections between nodes are visualised with double arrows to indicate a dialectic relationship. *Contradictions* between the nodes are the foundation for innovative attempts to change activities and can result in changes in one node or in more nodes of the activity system (Engeström, 2001).

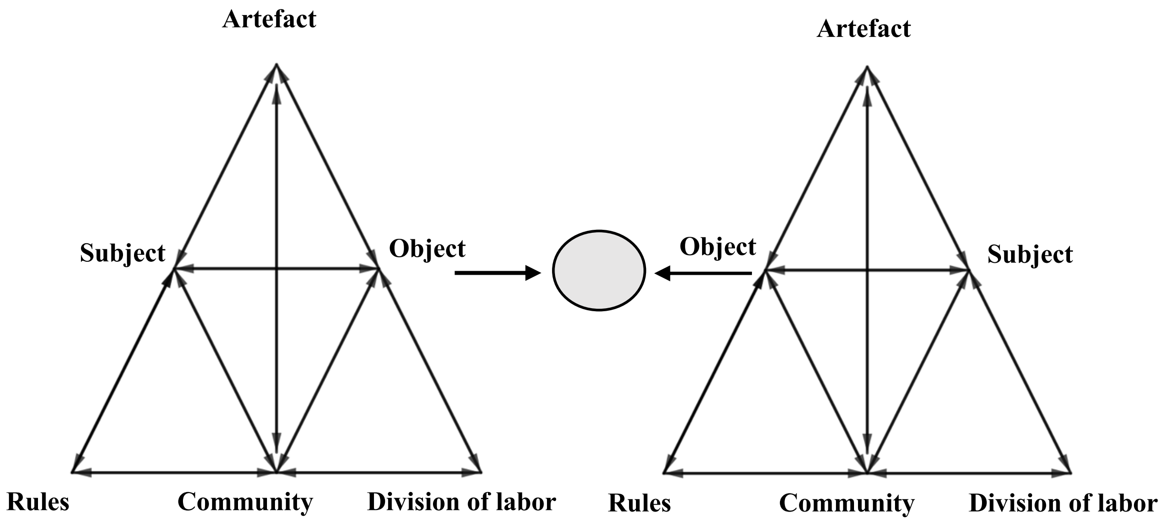


Figure 3: Two interacting activity systems sharing the same object.

2.1.2 Expansive learning

The concept of *expansive learning* focuses on communities as learners and on the transformation and creation of culture and new knowledge, both of which are constructed collectively (Engeström & Sannino, 2010). In expansive learning, learners are integral parts of a wider network. To develop expansive learning, CHAT proposes that the subjects must engage in a cyclic process of seven epistemic actions in which contradictions contribute to the process of learning (Figure 4).

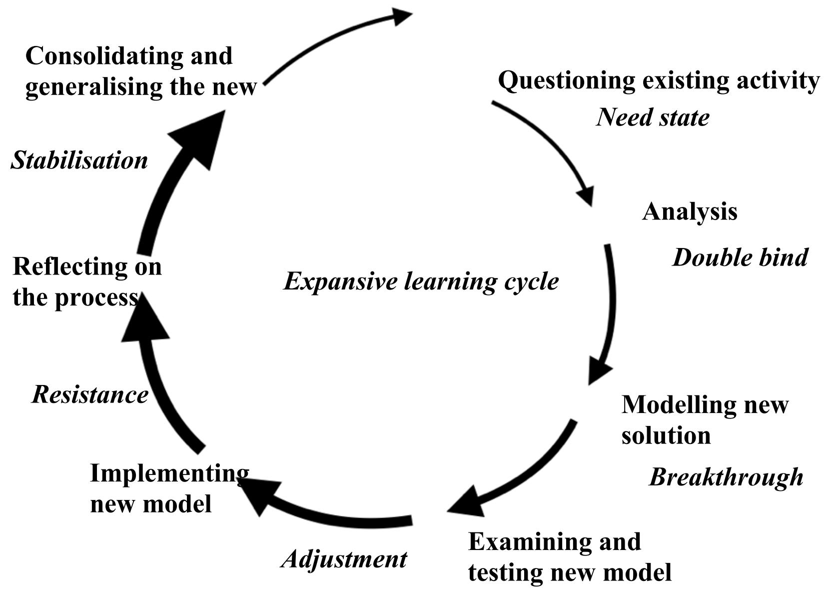


Figure 4: Model of expansive learning (adapted from Engeström & Sannino, 2010)

The expansive learning process of seven actions follows:

- The first action is to question the existing practice, for example, classroom practices and students' learning.
- In achieving the second action, the subjects must analyse the present situation. The objective of the empirical analysis is to construct a model of the relations within the activity system to explain the present situation (Engeström & Sannino, 2010, p. 7).
- The third action is to model a solution by, for example, designing a new intervention or revising an activity.
- The fourth action involves studying the model to understand its potential and limitations.
- The fifth action is to implement the new model, for example, by implementing an inquiry approach in the classroom.
- The sixth action is to reflect on and evaluate the process, including the discussions and rejections of the newly implemented activity.
- The seventh action is the process of making the activity a sustainable part of the activity system.

The expansive model can be a tool to identify and analyse various actions (Foot & Groleau, 2011). Engeström argues that "it seems promising to analyse these cycles in terms of the stepwise formation and resolution of internal contradictions in activity systems" (1999, p. 33). However, only a few studies in mathematics education have examined the process of expansive learning and the *contradictions* within or between activity systems: Chen et al. (2018) have studied the contradictions between a novice mathematics teacher educator and mathematics teachers in an in-service course; Williams and Wake (2007) have studied the contradictions between identifying mathematics in college and the workplace; Wake et al. (2016) analyse lesson studies as an approach of teacher professional development; Wei (2017) investigate professional learning experiences and the complicated trajectory of Chinese school teachers, here viewing the teacher learning process as expansive learning and locating four-level of contradictions into the cycle of expansive learning.

2.2 Anthropological theory of the didactics

The anthropological theory of the didactics (ATD), initiated by Yves Chevallard in the 1980s (Chevallard, 1992), is now a highly developed research programme with multiple contributions, both in terms of its theoretical elements and actual research studies. The aim of the ATD is not just to describe didactical phenomena, but also to question and design didactical practices from different perspectives and analyse the conditions and constraints that enable or hinder the specific evolutions of them.

The ATD offers a number of tools for modelling and designing educational practices. First, the ATD proposes a description of practice and knowledge in terms of praxeologies, so we can model, for instance, statistical praxeologies that appear in some context. Second, the ATD theorises the "context" as institutions (systems of positions relative to a praxeology) and as habitats of praxeologies with "conditions and constraints under which praxeologies start to live, migrate, change, operate, perish, disappear, be reborn, etc. within human groups" (Chevallard, 2007, as translated in Bosch & Ruiz-Munzón, 2022, p. 174). Third, the ATD proposes a change in pedagogical paradigm, introducing an inquiry approach: the paradigm of questioning the world (Chevallard, 2007, 2015).

2.2.1 Paradigm of visiting monuments and the paradigm of questioning the world

The *paradigm of visiting monuments* (PVM) is a pedagogical tradition in which answers (such as theorems or calculation rules) are detached from questions and presented as “monuments” to be acquired and admired in their own right (Bosch, 2018). The curriculum is formulated as a collection of works (in the sense of the Latin word *opera*, or the result of human work) that students have to study one by one as monuments visited during a tour. These works are often small and isolated, such as the notion of mean or relative frequency in descriptive statistics. The works in the curriculum have been “chosen because of their utility in the future life of students, but this utility is more assumed than proved because there is rarely the need in making it explicit” (Bosch, 2018, p. 4007).

The visit to a work takes place in a didactic system. The didactic system S(X;Y;w) is formed by person Y (the teacher), who helps person X (the students) to learn a piece of w (work). The PVM does not necessarily entail a traditional “transmissive” instructional format, where Y knows, organises and presents a piece of work and X practices repetitive drills. The PVM is also compatible with inquiry formats, where the teacher designs problem-based situations, for example, the structured problem-solving situation, where students compare two bus rides to visit the work of the relative frequency (Paper III). The main choice of the visit is not to know more about which bus to ride but the work hidden in the context of the bus problem.

The PVM has “shaped a relation to knowledge as driven by institutional, short-term, and labile motives, which stands away from the functional approach to knowledge based on its real-world utility—to understand a situation, be it mathematical or not, make a decision, or postpone it to allow for further study of the problem addressed” (Chevallard, 2015, p. 176). A consequence is that students study a given work without reflecting on what its uses are, could be, or have been, but if hardly pressed, the students may consider that its sole purpose is to be studied in school.

An alternative paradigm to the PVM is the *paradigm of questioning the world* (PQW, Chevallard, 2007, 2015). In this paradigm, which is largely hypothetical currently, a curriculum is not formulated as a collection of works to visit but as questions about the world in a broad sense (mathematics is part of it), which are considered important and essential to a given public of students. These questions are the central activity of the learning and teaching process. The questions are not an excuse to visit some explicit pre-established organisations of work; the study and elaboration of questions is important and meaningful in and of itself (Barquero & Bosch, 2015).

In the PQW, the study of works remains important, but the works are studied because of their capacity to answer specific questions or generate new derived questions. The new didactic system is formed around question Q in relation to which "it is up to the students led by the teacher to identify the fields of knowledge useful for their inquiry" (Chevallard, 2022, p. 88), thereby developing answer A. The didactic system can be described as S(X;Y;Q).

2.2.2 Study and research paths

Study and research paths (SRP) constitute a didactical design tool closely associated with PQW (Bosch, 2018; Chevallard, 2006, 2015). Here, students must engage in SRP by interacting with a generating question, Q_0 and then follow different paths (Bosch & Winsløw, 2016): the students must pose derived questions $Q_1 \ldots Q_n$, search for established and explore answers $A^{\Diamond}$, study relevant works $w_1 \ldots w_n$, validate answers $A_1 \ldots A_n$ and develop a final answer, $A^{\heartsuit}$. SRP are not only a design tool, where a Q&A (tree) diagram can be used to specify an a priori analysis in terms of the possible paths students can follow; but SRP may also serve as a model for didactic processes that can be used in a posteriori analysis (Jessen, 2014; Winsløw et al., 2013; Papers III and IV here).

2.2.3 Praxeologies

In the ATD, human activity and knowledge are modelled as bipartite entities called *praxeologies* (Chevallard, 1999). A praxeology consists of a *praxis block* (practical knowledge or know-how) and a *logos block* (explicit knowledge or know-that-and-why).

The praxis block is formed by types of tasks and techniques. A type of task is a class of challenges that humans can face and that are somehow similar, for example, collecting data or communicating results; some may have many possible techniques, while others are simply characterised by being solvable. A task can be small and highly dependent on a context, such as the question, "Why is relative frequency a better model than frequency?" (appearing in Paper III), or the questions can be more challenging and in need of being reduced or subdivided to get one or more tasks with a known technique (such as "analyse the data"). A technique is a way of carrying out a task, and for a given technique, we can define a type of task as a task that the technique can solve.

The logos block is formed by technology and theory. A technology is a discourse on techniques that explain and justify them: how and why they work, when to use what technique, and so forth. Theory is a more abstract discourse because it exists to justify and unify one or more technologies, independently of specific praxis blocks.

Praxis and logos do not exist separately, "Praxis thus entails logos which in turn backs up praxis" (Chevallard, 2006, p. 23). In fact, we could consider that even primitive animals possess several praxis blocks, but it is characteristic of human praxeologies to be accompanied by and supported through discourses about praxis and by even general ideas supporting that praxis. Far from being limited to mathematical or other learned praxeologies that would not even exist without a logos block, theory blocks are crucial to all didactic processes that go beyond the mere mimetic learning of praxis (i.e., learning through imitating a praxis, which also exists among animals). Moreover, logos are essential not only to sharing and learning praxis, but also to developing and refining praxis, adapting existing praxeologies to provide answers to slightly new kinds of questions and so forth. The praxeologies are a consequence of continuous processes with intricate dynamics; for example, new praxeologies are elaborated by integrating answers into new techniques (Bosch & Gascón, 2014).

The praxeologies of mathematics teachers constitute a complex structure of mathematical praxeologies, didactical praxeologies (related to teaching tasks and techniques carried out to solve them) and paradidactic praxeologies (tasks that are not directly solved through teaching but nevertheless support didactic praxeologies, such as giving feedback to parents, planning a lesson, etc.). Thus, didactic praxeologies evolve in teaching situations (Bosch & Gascón, 2014) through interactions between the teacher and students. Didactic praxeologies are, in a sense, auxiliary to other praxeologies, namely those that the teaching tasks aim to share with students (e.g., mathematical praxeologies). Similarly, paradidactic praxeologies are auxiliary to didactic praxeologies (as the example of planning a lesson clearly shows). Both didactic and paradidactic praxeologies are carried out under institutional conditions, which can be described as didactic and paradidactic infrastructure, here broadly defined as the conditions for teachers' work inside and outside of the classroom (Miyakawa & Winsløw, 2019). For example, the systems for teachers' professional development (courses and activities) are elements of the paradidactic infrastructure; more examples appear in Paper II.

2.3 ATD and CHAT: Complementary but independent frameworks

Because I have employed two theoretical frameworks consecutively during my years of PhD study, I include a short discussion of the differences and similarities between them, along with how I consider them to be complementary and, in some sense, compatible (or at least nonconflicting).

CHAT is a theoretical framework that aims to understand activity, taking into account affect and other expressions of human psychology (e.g., intention and feelings). CHAT focuses on the structured aspect of subjects' interactions, with the notions of roles, rules, division of labour and (re)definition of goals of the activity. CHAT provides a lens to situate, for example, the individual teacher's or educational researcher's work into an activity system, analysing interaction with resources, roles, rules and division of labour.

- The activity system is used to analyse the teacher's role and labour division.
- The activity system is taken as the prime unit of analysis, especially with a focus on the subject, object and artefacts.
- The activity system is not alone in illustrating what is happening in classrooms, but the division of labour generates different positions for the participants.
- The activity system takes shape and is transformed over periods of time.
- In the activity system, the contradictions between nodes have a central role of change and development.

The ATD was developed primarily from an epistemological perspective in the research on mathematics education. The objective is to explore mathematical knowledge as constructed and disseminated in institutions, with a special focus on investigating the institutional contingency of knowledge as it appears through the phenomenon of didactic transposition (e.g., Bosch & Gascón, 2006; Chevallard, 1992, 1999). The ATD provides a framework that involves both microdidactic, macrodidactic and societal levels (from a subject or question treated in the classroom to the levels of culture and humanity) and how these levels interact. The ATD insists on explicit and independent models of disciplinary knowledge.

Both theoretical frameworks are described as grand theoretical frameworks, that is, well-established and complete frames for research both inside and outside mathematics education (Kieran et al., 2015). The ATD originated within mathematics education but has now expanded to other areas of educational research. CHAT originated in studies of workplace learning and was later brought into educational research, with the case of schools (teachers' workplace). Both frameworks adopt a systemic perspective on teachers, students and schools, with the "activity systems" of CHAT being somewhat similar to institutions in the ATD. There are also important differences. CHAT focuses on subject(s) and their interactions with other nodes of the activity system; the ATD focuses on modelling the (variable and developing) praxis and knowledge involved in a didactical process or praxis; and the institutional constraints affect the process. So, although CHAT may give a much more detailed picture of what the ATD terms the institution, the ATD provides more detailed models of the activities or processes it engages in. This may also be because of the deliberate focus on "the didactic" in the ATD, which is, by nature, evolving, while the focus on "systems" within CHAT reflects an interest in those features behind what appears to be ever changing (like professional activity).

It is probably not meaningful to integrate or combine studies in ATD and CHAT, given that they offer elaborate and independent traditions of research. In the discussion, I show how the results from studies carried out in each contribute valuable and complementary perspectives on the general *problematique* of the present thesis, as outlined in Section 1.

3 Research background: Inquiry in statistics education

3.1 Mathematics and statistics

In most middle schools, statistics as a school subject is located in the discipline of mathematics, being on equal terms with domains like algebra and geometry. However, mathematics and statistics differ in several ways, as do mathematics education and statistical education.

As a science, mathematics is about logic, operations, generalisations, abstractions, deductive reasoning and proving results based on axioms and definitions (Gattuso & Ottaciani, 2011). School mathematics should be authentic to professional mathematical behaviour (Burton, 2004) and include creative ideas, question-posing, mistakes, models, explanations, patterns and an attitude to keep going when it is difficult. A mathematical problem may begin within a given context, but the context is just often a driver or motivator for students to study. One example is realistic mathematics education (RME; Freudenthal, 2006), which is based on the idea that a context can serve as a source for initiating the development of mathematical procedures and concepts, “which then gradually become more formal and general and less context specific” (Van den Heuvel-Panhuizen et al., 2014, p. 521). In mathematics, a pure mathematical focus can be of interest, and mathematical structures are often the aim in and of themselves.

Among statisticians, it is almost a truism that statistics is a mathematical science, but that statistics is not a part of the discipline of mathematics (Moore & Cobb, 2000). Even though definitions of statistics may differ, the individual nature of statistics as a discipline is established as a “general intellectual method that applies wherever data, variation and chance appear. It is a fundamental method because data, variation and chance are omnipresent in modern life. It is an independent discipline with its own core ideas rather than, for example, a branch of mathematics” (Moore, 1998, p. 1254), and statistics is further described as a “methodological discipline. It [statistics] exists not for itself but rather to offer to other fields of study a coherent set of ideas and tools for dealing with data” (Cobb & Moore, 1997, p. 801). In this sense, statistics is about inductive reasoning, uncertain conclusions and insisting on an interpretation in a context; for example, the collection of data, elaboration of data and interpretation of results depend on a context and will not necessarily involve much formal mathematics. A given statistical question and related data can be analysed in many ways and lead to different answers.

In this sense, “Statistics is not just about the methodology in a particular application domain; it also focuses on how to go from the particular to the general and back to the particular again” (Fienberg, 2014, p. 6) in a constant interplay between pattern and context (Cobb & Moore, 1997).

3.1.1 Aims of learning in statistical education

The statistics education literature emphasises three statistical learning aims: statistical literacy, statistical reasoning and statistical thinking. These statistical aims are often contrasted with the more traditional aims of statical teaching, which involves an intensive focus on statistical skills, small data sets, simple statistical descriptors and graphs, procedures and computations that do not prepare students to interpret or communicate data critically (Ben-Zvi & Garfield, 2004). However, there is no common and formal agreement on the definitions of and distinctions between statistical literacy, statistical reasoning and statistical thinking (Garfield et al., 2003).

3.1.1.1 Statistical literacy

Statistical literacy is broadly described as “being able to organize data, construct and display tables, and work with different representations of data. Statistical literacy also includes an understanding of concepts, vocabulary and symbols, and includes an understanding of probability as a measure of uncertainty” (Ben-Zvi & Garfield, 2004, p. 7) and the “ability to access, understand, interpret, critically evaluate, and if relevant express opinions, regarding statistical messages, data-related arguments, or issues involving uncertainty and risk” (Gal, 2019, p. 2).

In an earlier characterisation, statistical literacy is described as being a data consumer and a data producer (Gal, 2002). In a more recent characterisation, statistical literacy is classified as reading and writing the word and world (Wailand, 2017). The statistical literacy perspectives of consumers/producers and reading/writing relate to a double perspective on statistical literacy, involving seeing “statistics as a lens, enable a new view of the world” (Weiland, 2017, p. 37), but also seeing statistics as the possibility to change the “world” by collecting and analysing data and creating new contexts. The notion of statistical literacy is further discussed in Paper IV.

3.1.1.2 Statistical reasoning

One goal of mathematics education is *mathematical reasoning* (Ball & Bass, 2003; Niss & Jensen, 2002). Mathematical and *statistical reasoning* may appear to be similar, but the natures of mathematical reasoning and statistical reasoning are somewhat different.

Because mathematics can be described as the study of logic and patterns, mathematical reasoning involves reasoning about logic and patterns (del Mas, 2004). Mathematical reasoning is defined as establishing a relationship between elements, a condition, and a consequence (Brousseau & Gibel, 2005); and in a broader perspective, it includes investigative patterns and making conjectures, which involves generalisation and arguments (Stylianides, 2008).

Statistical reasoning, however, is characterised as a specific form of reasoning. Statistical reasoning involves developing interpretations based on data and representations of the data. Reasoning in statistics is defined as "understanding and being able to explain statistical processes and being able to fully interpret statistical results" (Ben-Zvi & Garfield, 2004, p. 7), and "a person who can explain why a conclusion is justified demonstrates statistical reasoning" (del Mas, 2004, p. 85). Statistical reasoning is not considered as related to deductive proofs but as a process of studying a statistical problem and engaging in the process of reasoning about *how* and *why*. Statistical reasoning is also exploring and testing a statistical model and investigating if the model is a reasonable fit to a specific context. The concept of statistical reasoning is further discussed in Paper IV.

3.1.1.3 Statistical thinking

It is generally considered that a traditional statistics teaching approach, with a dominant focus on skills and computation, typically lead to situations where "people learn methods, but not how to apply them [the methods] or how to interpret the results" (Mallows, 1998, p. 2). Statistical thinking is described as "noticing, understanding, using, quantifying, explaining and evaluating variation; thinking about the data 'literature'; capturing relevant data and measurements; summarising and representing the data; and taking account of uncertainty and data variability in decision making" (Pfannkuch & Wild, 2004, p. 40). Pfannkuch and Wild describe different elements of statistical thinking and the challenge teachers face in incorporating statistical thinking into classroom practices. The model in Figure 5 shows how they conceptualise and categorise statistical thinking.

(b) DIMENSION 2 : TYPES OF THINKING

GENERAL TYPES

- **Strategic**
  - — planning, anticipating problems
  - — awareness of practical constraints
- **Seeking Explanations**
- **Modelling**
  - — construction followed by use
- **Applying Techniques**
  - — following precedents
  - — recognition and use of archetypes
  - — use of problem solving tools

TYPES FUNDAMENTAL TO *STATISTICAL* THINKING (Foundations)

- **Recognition of need for data**
- **Transnumeration**
  (Changing representations to engender understanding)
  - — capturing "measures" from real system
  - — changing data representations
  - — communicating messages in data
- **Consideration of variation**
  - — noticing and acknowledging
  - — measuring and modelling for the purposes of prediction, explanation, or control
  - — explaining and dealing with
  - — investigative strategies
- **Reasoning with statistical models**
  - — aggregate-based reasoning
- **Integrating the statistical and contextual**
  - — information, knowledge, conceptions

Figure 5: Types of statistical thinking (Wild & Pfannkuch, 1999, p. 226)

Statistical thinking can be difficult to separate from statistical reasoning, but although statistical reasoning includes explaining statistical processes, statistical thinking also involves an understanding of how and why statistical investigations are conducted: "Statistical thinking involves an understanding of the nature of sampling, how we make inferences from samples to populations, and why designed experiments are needed to establish causation" (Ben-Zvi & Garfield, 2004, p. 7). del Mas (2004) describes how it is possible to distinguish the two statistical concepts through the nature of the statistical task: "For example, a person who knows when and how to apply statistical knowledge and procedures demonstrates statistical thinking. By contrast, a person who can explain why a conclusion is justified demonstrates statistical reasoning" (p. 85).

3.1.2 Research focus on statistical literacy, reasoning and thinking

Statistics education research emphasises the importance of statistical literacy, statistical reasoning and statistical thinking.

It is a common understanding that today’s citizens need to be statistically literate in order to cope with everyday life (further discussed in Section 1). Quantitative information and statistics are increasingly present everywhere: in news, on social media, in politics and everyday life—and visualised in charts, graphs, tables or represented in mean, average, correlation and change (Ben-Zvi & Garfield, 2004; Gal, 2002).

Several statistics education researchers (e.g., Gal, 2003; Rumsey, 2002) argue that statistical literacy must be a part of the curriculum in statistics at all levels. The research on statistical literacy points in different directions, for example, documenting a lack of statistical literacy and the inability to apply statistics in everyday life (e.g. Gal, 2002; Verhoeven, 2006), or investigating the development of it in terms of more or less elaborate literacy levels (e.g., Watson & Callingham, 2003). Other studies consider the development of tool to assess statistical literacy (e.g., Watson, 1997), often in relation to representations of data or statistical claims in media—which are sometimes incorrect (e.g., Spiegelhalter & Riesch, 2008).

Studies on statistical reasoning consider both reasoning with concepts and aggregates (e.g., Maxara & Biehler, 2007, Noll & Shaughnessy, 2012); reasoning about different visual representations (Pfannkuch, 2005); and the role of context in reasoning (e.g., Makar & Ben-Zvi, 2011). In particular, it is a recent, frequent hypotheses that certain technological tools have enhanced statistical reasoning at the school level (Biehler et al., 2013), for instance with informal reasoning using dynamic visualisation tools (e.g. Konold & Kazak, 2008; Martins et al., 2015), comparing groups with dynamic visualisation tools (e.g., Frischemeier & Biehler, 2015; Watson & Donne, 2009), or enabling digital simulation and informal inference (e.g., Lee et al., 2015).

The shift from developing skills to statistical thinking has likewise entailed a focus on statistical inquiry models and how students not only learn methods, but how the students analyse data, apply statistical descriptors and interpret results (e.g., MacGillivray & Pereira-Mendoza, 2011; Pfannkuch & Rubick, 2002; Watson, 1997).

Even though there has been an ever-increasing activity in statistics educational research, there is still a “growing demand for statistical skills, knowledge and competencies at the same time that the field of statistics and data science is broadening” (Wild et al., 2018, p. 32). Indeed, “there are many important research questions to be addressed including determining how to best structure, teach and assess the evolving statistical curricula” (Wild et al., 2018, p. 32).

3.2 Statistical inquiry and its relation to study and research paths

The role of inquiry has its roots in science education and, in recent decades, has migrated into the field of mathematics education. Inquiry contributes “to a shift in epistemology from seeing knowledge given as faith to knowledge based on thinking, reflection, experimentation and science” (Artigue & Blomhøj, 2013, p. 798), aiming for students to play an active role in the process of learning. The different theoretical frameworks conceptualise inquiry and illustrate the importance of, for example, the autonomy of students, authenticity of questions, experimental dimension of mathematics and the critical and democratic dimensions of inquiry. The process of inquiry is not only to develop answers, but also to provide reasons for *why* answers are true (Artigue & Blomhøj, 2013).

Research in statistics education generally supports the usefulness of establishing situations where students can explore and inquire statistical issues and problems in a process of posing questions, collecting data, planning how to analyse, visualise, represent and analyse data and develop conclusions. Across a span of more than 40 years, researchers have developed models of these inquiry processes, for example, explorative data analysis (Turkey, 1977), the investigative cycle outlined in the PPDAC model of problem, plan, data, analysis and conclusion (Wild & Pfannkuch, 1999) and the Guideline for Assessment and Introduction in Statistics Education (Bargagliotti et al., 2020; Franklin et al., 2007). A shared idea of the models is that students should go through phases that emulate the experienced statistical researchers’ work, viewed as an ongoing alteration between data and the context throughout the statistical investigation. The Guidelines for Assessment and Instruction in Statistics Education describe the inquiry process as the foundation and core of statistical knowledge: “Spurred by the overabundance of data available in today’s world, the statistical problem-solving process not only remains important, it becomes even more critical for drawing conclusions from data” (Bargagliotti et al., 2016, p. 7).

The investigative cycle (Wild & Pfannkuch, 1999) involves five steps. First, posing a problem starts the inquiry process; it can be a vague idea that needs to be developed or reformulated into a question that can be a driver of the investigation, such as “Do children spend too much time on media … or ?” (Paper I) or “find the correlation between jumping squats and push-ups” (Paper III). The process of planning the design includes reflections on what kind of data are suitable and how to collect the data. Collecting data further includes storing and organising data. The analysis can take various forms: from a focus on specific statistical techniques and statistical representations to a more informal analysis, for example, by using a dynamic visualisation tool. The last part of the investigative cycle is to make sense of data and analysis in view of finding an answer (conclusion) to the problem. Even though the investigative cycle is often illustrated as a cycle with one-way arrows (Wild & Pfannkuch, 1999, p. 226), there are also processes of back and forth between the steps involving tentative analysis and conclusions and more analysis (Konold & Pollatsek, 2002); in addition, the investigative process can include investigative cycles. It is an assumption that students must experience all steps of the inquiry cycle, but not necessarily at the same time. The investigative model is a “nonhierarchical, nonlinear, dynamic way of thinking that encompasses an investigative cycle, an interrogative cycle, types of thinking, and dispositions, all of which are brought to bear in the solving of a statistically based problem” (Pfannkuch & Wild, 2004, p. 41).

There are several major research projects that explored the statistical inquiry process (e.g., Fielding-Wells, 2010; MacGillivray & Pereira-Mendoza, 2011), some of which focused on a single level of the investigative cycle, such as posing problems (e.g., Arnold & Pfannkuch, 2019). Watson et al. (2018) recommend further research on student autonomy: “Student-driven statistical inquiry is worthy of further investigation” (p. 113).

In my research project, I approach statistical inquiry by considering different inquiry models like: the investigative cycle of problem, planning, data, analysis and conclusion (Wild &

Pfannkuch, 1999). The design in Paper I is based on the investigative cycle, and in Paper III, there are elements from the investigative cycle. In Paper IV, the design and analysis are based on SRP (Chevallard, 2015).

3.2.1 Study and research paths – An inquiry approach

The theoretical notion of SRP was introduced in Section 2; we now outline how it has been used and studied in recent research. The aim of an SRP is to engage students in meaningful and challenging real inquiry situations where they learn to pose problems and solve (real) problems, even learning to evaluate and question the knowledge elaborated and further being able to address the same or related problems (Bosch, 2018; Bosch & Winsløw, 2016; Chevallard, 2015). Contrary to the investigative cycle, there are no predetermined phases to direct the inquiry process of the students, and an SRP is modelled not by generic steps but by a tree-shaped graph which outlines the potential or actual semantic connections between questions, starting with the generating question (Winsløw et al., 2013).

For more than 15 years, initial experimentations with SRP have focused on designing, realising and analysing SRP at nearly all educational levels. Publications on these cases mostly indicate successful learning, for instance in terms of the development of advanced praxeologies in teaching is based on an SRP.

In a Danish context, there have been experiments with SRP in upper secondary schools, such as studying painkillers in mathematics/biology and characterising exponential functions in mathematics (Jessen, 2014, 2017). Also, in the setting of a major university college, preservice teachers in primary and lower secondary school were engaged in study and research paths on illness and health issues, in an experiment which seeked to integrate topics from mathematics and biology (Rasmussen, 2016). Researchers have pointed to institutional constraints like examinations and monodisciplinary courses where the disciplines appear one by one, but also to those episodes where the SRP motivate the students and enables them to make two disciplines interact. However, SRP have not become a part of classroom practices apart from these few research initiatives. In a Danish context and prior to our project, there have been (as far as we know) been no experiments with SRP in primary or lower secondary school.

Internationally, especially in Spain, there have been more experiments with SRP (Barquero et al., 2022). Most of these experiments have been carried out at postsecondary level, like the study of population dynamics in mathematics (Barquero et al., 2013); sales prediction and organisation of sharing bikes in mathematics (Serrano et al., 2010); and general elasticity and how to make a machine (Florensa et al., 2016). As constraints and conditions in universitary institutions, the authors also point to the challenge of defining a curriculum by questions and not by a dense set of contents to study. The notion of SRP-TE (SRP in Teacher Education, see e.g. Barquero et al., 2015; 2019) refers to a process where participants are first engaged in an SRP as students, then analyse it as didacticians, and finally design and experiment one or more SRP as didacticians and teachers – all under the supervision of university teachers (Barquero et al., 2018). Resent research on SRP for business students at the university level has been carried out in the context of a statistics course (e.g. Markulin et al., 2021, 2022); the authors show how various course elements nourish SRP and how the SRP supports the evolution of the content of the course all adapted to the students' needs. The researchers also discuss the challenges related to planning SRP interventions, and the need to communicate the constraints and conditions to the research communities.

Even though SRP has been a focus of research for more than 15 years, we are still at the beginning. We still need research on how SRP can unfold without external support from researchers, and more cases of how SRP can function in the specific institutional context of primary and lower secondary school. It is also of interest to examine how different disciplines contribute to shape a SRP, for example, how a statistical focus differs from a mathematical focus in the process of questioning.

In Paper I, we study two different inquiry approaches: one in Japan and one in Denmark. In Paper IV, we study the design and realisation of SRP in statistics in a grade 5 class.

3.3 Paradidactic infrastructure and teacher professional development

In Section 2, we provided a brief overview of the ATD (Chevallard, 1999, 2006) and the complex interaction of teachers' statistical praxeologies, and the corresponding didactic and paradidactic praxeologies. Paradidactic praxeologies are further described in Paper II.
The development of teachers' professional knowledge depends to a large extent on paradidactic infrastructure (cf. Miyakawa & Winsløw, 2019; Winsløw, 2012) that governs and supports teachers' professional activity outside the classroom, for example, curriculum materials, textbooks, time given to prepare teaching, workspace and resources available.

Miyakawa and Winsløw (2013, 2019) have contributed to research into the conditions for teachers' work outside the classroom related to mathematical praxeologies and didactic praxeologies. They exhibit a model of the Japanese paradidactic infrastructure, incorporating institutional frameworks, settings for teachers' study and research, and media used by the teachers. They point out lesson study as an important element of the paradidactic infrastructure in Japan and the development of teachers' shared mathematical and didactic praxeologies. However, lesson study is indeed one of the elements supporting teachers' professional development in Japan, and its function there cannot be understood if it is viewed in isolation from other elements of the paradidactic infrastructure. Miyakawa and Winsløw stress that "Indeed, paradidactic infrastructure is crucial to *any* initiative to develop the teaching profession" (2019, p. 302). Jessen et al. (2019) compare the paradidactic infrastructures in Denmark and Japan to identify the ecologies which the two countries offer for realising viable SRP, in particular through curriculum materials, textbooks and accessibility of ICT. Their analysis shows more parallels than differences.

Also, in other theoretical directions, outside the ATD community, there is a research interest in the "ecology of human development" (Bronfenbrenner, 1999), "whole school change" (Thompson, 2007) and "teacher development across classrooms and schools" (Cobb et al., 2003). Fullan et al. (1990, 2005) present a framework to support the sustainability of teachers' professional development, bringing together teachers, administrators and researchers. They conclude that systemic and cultural changes in schools and classrooms are intimately intertwined, and progress cannot be sustained by individual teachers but must be rooted in and borne by institutions.

3.3.1 Teacher professional development in statistics

Over the past three decades, it has been a common understanding that the quality of teaching depends on the subject-related knowledge of teachers (Rowland & Ruthven, 2011). Since Shullman's (1986, 1987) introduction of the concept of pedagogical content knowledge, several researchers have developed models of teacher knowledge (Ball et al., 2008; Turner & Rowland, 2011). However, even though the area of mathematical knowledge for teaching (MKT) is well researched, there is surprisingly little research on statistical knowledge for teaching (SKT; Batenero et al., 2011; Groth, 2007).

Several researchers have attempted a specialisation of the MKT model (Ball et al., 2008) into a statistics version, naturally named SKT (Burgess, 2011; González, 2021; Groth, 2017; Groth & Berger, 2013). Here, SKT is viewed as consisting of two main components: subject matter knowledge and pedagogical content knowledge. Within subject matter knowledge, three parts are distinguished: First, common statical knowledge, which is the statistical knowledge both needed by teachers and other professions, for example, reading graphs and calculating descriptive statistics. Second, specialised statistical knowledge is then specific to teaching situations, for example, representing statistical "ideas in ways that are comprehensible to students" (Groth & Berger, 2013, p. 249), which is not necessarily needed in other professions. The third is horizon statistical knowledge, which is to know the broader set of mathematical or statistical ideas to which the statistical content connects, e.g. "understanding some epistemological obstacles related to the historical development of probability" (Godino et. al., 2011, p. 273). Pedagogical content knowledge also consists of three conceptions: knowledge of statistical content and teaching, for example, the qualities that different inquiry models can offer, as frameworks for students' statistical investigations; knowledge of the curriculum, for example, how statistics is also present in other subjects; and knowledge of statistical content and students, for example, the way that students approach different statistical tasks or specific statistical difficulties. The SKT thus represents what teachers know about statistics and its teaching. Gonzáles (2014) further develops the conceptual model of the SKT to include beliefs about teaching statistics and teachers' motivation and self-regulation. In the ATD, the teachers' statistical knowledge is modelled in terms of statistical praxeologies and corresponding didactic praxeologies (see Section 2.2).

Eichler and Zepata-Cardona (2016) describe how only a few studies regarding the SKT relate to the actual practice of teaching. This is critical because the knowledge needed to teach is more complex than declarative knowledge. Eichler and Zepata-Cardona state that "it is simplistic to focus exclusively on teachers' content knowledge" (2016, p. 7) and stress that it is important to know how teachers use their statistical content knowledge and pedagogical content knowledge in classroom teaching. da Ponte and Noll (2018) further add that professional developmental research must emphasise the integration of classroom practices and statistical content knowledge while engaging teachers in inquiring about their own classrooms (e.g., Groth, 2017; da Ponte, 2011) and fostering an inquiry position about the teachers' own teaching (Jaworski, 2003). The paradidactic infrastructure—lesson study—represents such an inquiry method and is one approach for teachers to engage in a structured inquiry approach to students' statistical learning, both outside (planning and reflecting) and inside (teaching and observing) classrooms (e.g., Roback et al., 2006).

In general, we agree that more research on teachers and statistical teaching is needed. "Many teachers try to teach statistics like other mathematical topics, focusing on only the results, procedures, graphs, etc., rather than on statistical thinking and reasoning processes. What is it that teachers need to know?" (Bakker et al., 2018, p. 56). We especially need to "investigate teachers in their professional context, in classrooms" (Eichler & Zepata-Cardona, 2016, p. 12). This is one of the leading ideas of this thesis.

4 Research questions and methodology

4.1 Research questions

Some of the leitmotivs and motivations behind the present dissertation were outlined in the introduction, including some which are bound to personal experience and the national context of the author. After the presentation of the theoretical frameworks and background, we can now provide a precise formulation of the objectives of this PhD project in terms of three overall research questions (RQ), which concern statistical, didactic and paradidactic praxeologies. We have formulated the research questions within ATD, but we can reconstruct the research question of paper I as part of RQ3: the set-up for collaboration between teachers and researchers considered (and modelled as two activity systems sharing a common object) is indeed the paradidactic infrastructure whose effect on teachers' learning (development of didactic and paradidactic praxeologies) and is studied in that paper.

RQ1: How can inquiry approaches such as investigative cycles and study and research paths support students' development of statistical praxeologies in school settings?

RQ2: What conditions and constraints for implementing these inquiry approaches in primary and lower secondary school mathematics—and specifically statistics—can be observed in different school contexts?

RQ3: How can paradidactical infrastructure support teachers to design, implement and evaluate teaching that involves statistical inquiry by students?

In section 5, we elaborate further on how the different four papers address the research questions and what answers they contribute. Briefly, RQ1 was addressed in Papers III and IV, RQ2 was addressed in Papers I, III and IV, and RQ3 was addressed in Papers I, II, III and IV. We now outline the methodologies employed in the papers in relation to the three questions.

4.2 Methodology

The papers in this dissertation draw on several closely related projects: two large-scale Danish research and development projects, where only some of the interventions had a research focus on educational statistics; a study of paradidactic infrastructure in Denmark and sustainability of lesson study; a comparative study of how statistical inquiry was implemented in two lessons taught in Denmark and Japan; and a design research project about how SRP can contribute to middle school students' development of statistical knowledge. The context of all research projects was Danish primary and lower secondary school, except in Paper III, where the context also included Japanese lower secondary school.

We combined data from these different projects to address the research questions stated above. However, due to the variety of the projects, both in their nature and context, different methodologies were applied in the papers and thus contribute to the answers to the questions. We now outline how.

4.2.1 Methodology for RQ1: How can inquiry approaches support students' development of statistical praxeologies in school settings?

The methods for answering RQ1 originate from two research projects (and drawn on in Papers III and IV): A descriptive case study (Yin, 2013), and a design research project (Bakker, 2018; Barquero & Bosch, 2015; Cobb et al., 2015). Both projects focused on statistical inquiry and students' development of statistical praxeologies. The data were qualitative in nature: the primary data to answer RQ1 consisted of audio and video recordings of the classroom lessons. The two research projects differed in the sense that the descriptive case study had a research focus on analysing teacher-driven statistical classroom practices of teachers and students in relation to inquiry based statistics teaching, while the intention with the design research project was to question, implement and analyse a didactic design based on the theoretical principles of SRP. In the design study, we drew on teachers' experience and observations to improve the SRP design and implementation. Therefore, to some extent we implemented the cyclic method from lesson study, fully aware that design research aims at developing educational knowledge beyond a given context, while a lesson study aims at improving the shared knowledge of a smaller group of teachers in systematic way, and can be characterised as professional teacher development (Miyakawa & Winsløw, 2009).

The analysis of the data in the two papers was threefold: first, we analysed the lessons semantically and categorised teacher's and students' contribution as questions and answers (further described in Paper III), second, we constructed Q&A diagrams displaying questions and answers by the teacher and students, and third, we modelled the developed answers in terms of statistical praxeologies.

4.2.2 Methodology for RQ2: What conditions and constrains for implementing inquiry approaches in statistics can be observed in different school contexts?

The methods for answering RQ2 originate from three research projects (Papers I, III and IV): an explanatory and critical case study (Flyvbjerg, 2006; Yin, 2014), a descriptive case study (Yin, 2013), and a design research project (Bakker, 2018). All three research projects focused on statistical inquiry. In Paper I, the research focus was on teachers' learning and implementation of new statistical inquiry approaches. In Paper III, the main research focus was on students' development of praxeologies and the hypothesis about how different paradidactic elements can support or hinder statistical inquiry. In Paper IV, the main research aim was to develop new educational knowledge from the design research while involving the teachers in the project (also mentioned in Section 4.2.1) to improve the quality of the design, but also to settle old conceptions of the teacher as the "problem of implementation".

The data gathered in the three projects were qualitative in nature. The data to answer RQ2 consisted of audio and video recordings of classroom lessons, field notes, audio recorded teacher interviews, contextual information about the teacher, for example, if the teacher is a mathematics consultant, contextual information about the school, for example, if the school is attached to a University of Education, video recordings of planning statistical inquiry, teachers' reflection of students learning (only Paper IV), and paradidactic elements like textbook, curriculum and teachers' possibility to participate in professional development activities.

The analysis of the conditions and constraints were different in the thee papers. In Paper I, the analysis consisted of two phases: first, we used the expansive learning cycle to model the design research processes, identifying the contradictions; second, we modelled teachers' and researchers' activity in activity systems. In Paper III, we built the analysis of the conditions and constraints upon the statistical praxeological analysis (described in Section 4.2.1) and differences between the two lived inquiry lessons. In the analysis, we identified paradidactic elements and discussed how these paradidactic elements conditioned the development of statistical inquiry. In Paper IV, we again built the analysis of conditions and constraints on the statistical praxeological analysis, focusing on how the teacher conditioned the dialectic of study and research.

4.2.3 Methodology s for RQ3: How can paradidactic infrastructure support teachers to design, implement and evaluate teaching which involves statistical inquiry?

The methods for answering RQ3 originated from all four research projects (Papers I, II, III and IV): three case studies and one design research project. But it was only in Paper II that the main research focus was on identifying and modelling paradidactic infrastructure, the development and establishment of a new paradidactic infrastructure, lesson study and the effects of current paradidactic infrastructure on teachers' professional development.

The data gathered to answer RQ3 serve to identify and describe a large variety of paradidactic conditions, which can be divided into different institutional levels, for example, national, regional, municipality and local school level, and the elements can be further separated into institutional frameworks, settings for teachers' study and research, and media used or produced by teachers. In Paper II, we have roughly modelled the paradidactic infrastructure in Denmark at the primary and lower secondary school levels.

The tools used to analyse these infrastructural elements are partly different in the four papers. In Paper I, we used the tools of expansive learning cycle and activity systems to model teachers' and researchers' common implementation of statistical inquiry; activity systems, in particular, offer a specific framing to categorize some of the infrastructure. In Paper II, the analysis was mainly based on two lesson study cases, selected to exemplify different aspects of how other paradidactic infrastructures influence the lesson study. In Paper III, paradidactic elements were used to support hypotheses about differences between two statistical inquiry lessons. In Paper IV, some paradidactic elements were implemented in the design of the design research to support teachers, while other paradidactic elements were identified and analysed after implementation.

5 Main results

Each of the three research questions will be addressed separately to show how the four papers answer them.

5.1 How statistical praxeologies are realised through different inquiry approaches

In Papers III and IV, we studied three approaches to implement statistical inquiry and their results in terms of students' development of statistical praxeologies. The analysis processes in the two research projects were similar and started with a semantic modelling of questions and answers in a Q&A diagram; we then used a praxeological analysis to interpret the answers that were produced.

The three statistical inquiry approaches differed in the design of the inquiry approach, institutional perspective, length of inquiry approach, and grade of the participating students. The three statistical inquiry approaches also shared some similarities: the students were all actively engaged in statistical inquiry practice, the students had to relate to, and elaborate, nonroutine problems, they made conjectures and questioned or validated their answers and conclusions.

In Paper III, we analysed two statistical inquiry lessons: one in Denmark and one in Japan. The length of both lessons was a single lesson. The lessons were planned by the mathematics teacher and took its starting point in a question, Q_0.

The Danish case was a case of 'experimental activity with many questions' (Paper III). In the lesson, the teacher posed the following question Q_0: "Find the correlation between jumping squats and push-ups". During the inquiry lesson, Danish eighth-grade students collected their own data by jumping and doing push-ups outside the classroom; the students posed critical questions about the quality and validity of the data, but the students' questions were not pursued. The statistical techniques (instrumented) were all provided by the teacher, and the students ended up with only a few non-justified answers.

The aim of the Danish lesson was not to develop an answer to Q_0. Jumping squats and push-ups were the context of the inquiry and led to two generic type of tasks—T_1 and T_2—about conducting an experiment and determining how the resulting sets of data were correlated. The techniques used for T_1 included the production and collection of two-dimensional data; the students questioned the quality of data, which could be seen as an invitation to study the technology. The lesson instead ended with students trying to apply a given technique to model the data. The students' hypotheses were not discussed. The techniques to find a correlation between two sets of data—T_2—were based on instrumented techniques: techniques of making a table, a scatterplot, a trendline and a function. These instrumented techniques were not further elaborated on. In the discussion about correlation, the students questioned if there was a correlation. The teacher ended the lesson by referring to his subjective judgement, that is, him trusting there would be a meaningful correlation.

The Japanese case was a case of 'structured problem-solving' (Paper III). In the lesson, the teacher posed question Q_0 "Which bus is better?" During the inquiry lesson, the Japanese seventh-grade students investigated given data; the data were especially designed to force the students to question and change various established statistical techniques; the students spent most of the lesson presenting and discussion techniques, technology and theory; the students also developed new techniques, technology and theory. The students' development of statistical praxeologies was firmly directed by the teachers' questions, and less by students' own questions.

As in the Danish case, the aim of the Japanese lesson was not to answer Q_0 *per se* but rather, it was for students to meet statistical needs and, as a response, develop new statistical praxeologies. In the Japanese case, the main type of task was to compare two data sets, including developing new techniques and evaluating potentials in tables and graphs. An example was how the students' techniques made it possible for them to draw out the 'right' information in a table and compare frequencies, how the students discussed whether or not the notion of frequency was a suitable descriptor to use, when it would be appropriate to compare a data set with different numbers of data (technology), and finally how the students and the teacher concluded that (absolute) frequency can only be used as a descriptor to compare corresponding groups within two data sets when data sets have the same size (theory). Another example of students' inquiry can be seen in the process in which the students and teachers engaged in developing new techniques, technology and theory about relative frequency.

The categorisations of the cases as teacher experimental activity with many questions, and no-less teacher directed structured problem-solving, were developed as an outcome of our own analysis. However, when we used the investigative model of problem, plan, data, analysis and conclusion (Wild & Pfannkuch, 1999) to discuss and elaborate further on the two inquiry approaches, we could observe that several different elements in the statistical approach were realised, while others were more absent.

In its design, the 'experimental activity with many questions' approach has a main focus on *problem*, *plan* and *data*—with a clear emphasis on first-hand experience with collecting data; the techniques of the *analysis* were given minor attention and were presented by the teacher, and the *conclusion* was more about subjective judgment than about statistical reasoning based on technology and theory. This 'experimental activity with many questions' approach had a linear structure, and even when students questioned the data, they did not discuss technological or theoretical issues, like on how to collect data. For example, it was not explicitly addressed how much data were needed.

The 'structured problem-solving' approach has strong links with Japanese didactical tradition that goes under this name, and where teachers structure the lesson around one procedural and (towards the end) conceptual problem, and allow students to develop and elaborate on their own procedures in the structured design (Stigler & Hieber, 1999, p. 27). The 'structured problem-solving' approach had a main focus on *data*, *analysis* and *conclusion*, especially the alternation between the three phases—back and forth several times—which is indeed observed in how the students and teacher questioned the techniques and developed new theoretical ideas.

In Paper IV, we analysed the design and implementation of an eight-lesson statistical SRP in grade five. The generating question, Q_0—that is, 'Are we physically active?'—was authentic in the sense that the praxeologies developed were far from limited to statistical praxis and logos.

Q_0 also needed to be studied and researched in a broader context, for example, praxeologies related to broad ideas like well-being and fake news. In Paper IV, we visualised (with Q&A diagram) the students' hypotheses, answers found in media, derived questions, data, inquiry, visit of statistical work, and further developed answers. The students' study and research paths developed in different directions; some students researched how many steps' while others examined averages of high intensity activities or compared how much time classes spent on media. In the SRP, the students autonomously inquired into data using TinkerPlots, drawing especially on the (instrumented) techniques of separating and stacking data in different ways, along with using 'black-box processes' like pressing the buttons Δ, N or % to find the mean, frequency or relative frequency, respectively. To bridge praxis and logos, the techniques were studied by the students and teacher together, and the dialogue was scaffolded by the teacher's questions such as 'What is?' 'How do we?' and 'Why?' (technology and theory). We observed how statistical literacy and statistical reasoning supported each other—as special instance of the crucial dialectic of study and research. In Paper IV, we further (re)defined statistical literacy and statistical reasoning in terms of praxeological elements. Statistical literacy was defined as the ability to engage with statistical types of tasks, techniques and technology, and statistical reasoning was seen as the discursive activity based on statistical technology and theory.

The ideas behind the notion of SRP have been explicitly described (Section 2.2.2.). Even though an SRP is neither cyclic nor linear, but leads instead to develop many different realised paths, it is valuable to discuss the approach in terms of the investigative cycle. In the SRP, we observed all the elements of *problem*, *plan*, *data*, *analysis* and *conclusion*, but the students' study and research did not follow the progression of the investigative cycle. There was no predefined structure to follow, in fact the study and research paths were unpredictable and at times chaotic, when the students moved between research and study activities of problem, data, analysis and problem.

In Paper I, the investigative cycle (data-detective) was explicitly used to design the intervention 'Do children spend too much time with media?' However, in Paper I, the research focus was not on the students or their development of statistical praxeologies, so the statistical inquiry approach and students' development of statistical praxeologies have not been analysed there.

5.2 How the conditions and constraints for implementing statistical inquiry can be observed in different school contexts

The aim of the overall research project was not just to analyse the different inquiry approaches and the resulting statistical praxeologies, but also to study the ecology of statistical inquiry: the conditions that facilitated statistical inquiry and constraints that hindered the development of statistical inquiry.

Indeed, in Papers I, III and IV, we studied the conditions and constraints for implementing statistical inquiry. The research focus in the three articles differed: in Paper I, we studied teachers' professional learning; in Paper III, we studied how paradidactic infrastructure conditioned and constrained statistical inquiry; and in Paper IV, we studied the conditions and constraints for implementing an SRP. The ecological analysis was based on the institutional levels of didactic codetermination (Chevallard, 2002, 2019), and thus based on the fact that what happens in the classroom is connected with conditions and constraints coming from levels beyond the classroom situation, either general or specific. We focused, specifically, on the levels of society, school and discipline.

One of the main conditions for implementing statistical inquiry approaches that we analysed was the curriculum, along with how the curriculum concretely affected the evolution of inquiry at classroom level in different ways (Paper III). We observed how the ecology of statistical inquiry, here in a Danish context, was affected by different curriculum recommendations: statistical aims like *exploring connections of data found in everyday life*; generic educational aims like *implementing physical activity*; and educational trends like *teaching outside the classroom*. In the Japanese context, however, the ecology of the statistical inquiry was mainly affected by the subject specific recommendations, such as *helping students understand the meaning and necessity of representative statistical descriptors*. In the Japanese case, we further observed how the Japanese textbooks supported the curriculum recommendations, particularly through the selection of a "problem of the day" that aligns with the curriculum recommendation and also with the textbook teacher guide, which offers many concrete ideas of how to scaffold students' learning. Similar resources were not available in the Danish case. Indeed, the textbook did not offer problems related to exploring connections between sets of data. In general, it is not unusual for Danish mathematics teachers to develop their own tasks. Danish research on teachers' practice, and also curriculum recommendations (EVA, 2012, 2015), have shown how Danish teachers designed activities with the aim of motivating students; the activities were not necessarily based on concrete curriculum recommendations or ideas developed in textbook material.

Our implementation and study of SRP in primary schools seem to support the optimistic conclusion that curriculum recommendations at different levels, for example, generic educational purposes such as that the student should *gain confidence in own possibilities and … take a stand and act*, and also more specific statistical recommendations, such as students being able to *explore and present own statistical inquiry,* can indeed justify and support a SRP approach (Paper IV).

In the different research projects, we also observed conditions and constraints related to other resources. First, we uncovered how the educational design of statistical inquiry, as published on websites, did not offer sufficient support for teachers to manage a complex cycle of statistical inquiry (Paper I), even though the design included detailed descriptions of lessons, media to study, inquiry tasks, ideas of how to integrate digital tools, teacher questions to pose, and developmental support from a researcher in the process of implementing statistical inquiry. Second, we saw how sophisticated techniques, named *bansho* in Japanese, (teachers' Blackboard organisation), supported students' statistical writing and their development of statistical praxeologies (Paper III, the Japanese case). Third, in the SRP context, we found how digital tools and instrumented techniques, on the one hand, can act as a springboard to further statistical elaboration and validation, (Paper IV) and, on the other hand, may also appear as a 'black-box' technique, where students' questions may not really be answered (Paper III).

The statistical inquiry approaches required teachers to guide the students in various ways during the didactical processes. In our research, we observed how teacher's and students' didactic contracts developed when changing from a relatively routine based approach in statistics education to statistical inquiry. In the SRP (Paper IV), we observed how the students posed derived questions, followed different paths and developed different answers, as the teacher guided the students in the inquiry milieu and scaffolded common milieus of visiting statistical work by asking how and why questions. Students' development of logos, that is (according to our definition), statistical reasoning, only appeared when the students and teacher visit statistical work together.

Another research result related to the teachers' praxis was about *teacher confidence*. The participating teachers in the two national developmental and research projects (Paper I) were so confident with their regular teaching practices that they did not recognise the limitations or contradictions between their regular teaching practices and the inquiry intervention. The confident teachers were difficult to convince of the innovative aspect of statistical inquiry, and the teachers' learning did not correspond to the aims of the intentions of the development project. International research on teaching and learning (OECD, 2016) has shown that Danish teachers have a high level of autonomy, and make far-reaching decisions about materials, curriculum, discipline procedures and how to assess students' learning. Even though this may be seen as having principal benefits, it can also turn out be a constraint when seeking to implement a detailed new inquiry approaches, in which teachers must follow more or less teaching designs. Also, national Danish research points at how some Danish teachers oppose very detailed educational aims and activities designed by others, and instead prefer very broad educational aims, so that they are free to choose how to teach (EVA, 2015, p. 25). This can also explain why the teachers did not develop the practices and reflections which were intended in the national initiatives, based on implementation of detailed inquiry designs.

5.3 How paradidactic infrastructure supports teachers to design, implement and evaluate teachers which involves statistical inquiry

In Papers I, II, III and IV, we studied how *paradidactic infrastructures* supported the implementation of inquiry approaches and teachers' professional development. In all papers, the paradidactic infrastructure was crucially related to the institutional system in which the teachers worked.

Papers I and IV contribute similar but complementary insights, in part due to the different theoretical frameworks. The statistical inquiry projects studied in the papers can be broadly characterised as design research projects, but the form and realisation of design research differed remarkably. In the demonstration school project and the KiDM project (Paper I), the participating teachers and researchers did not jointly question the school institution's statistical practices, nor did the teachers and researchers develop the design of the interventions together. The design research process of SRP (Paper IV) differed from these projects by inviting the teachers to participate in several of the design research phases. The habitual didactical design format habitually associate with SRP was merged with elements of lesson study, in a design where the teachers' didactical experience was purposefully drawn on in the didactical design process, such in the steps analysis of the existing statistical practices, proposing new design ideas, and reflections on classroom implementation and students' learning, which were organised in a cyclic process of three iterations.

Teachers' different ways of participating can affect their perspectives on the outcomes of the two projects. In fact, in the demonstration school and KiDM projects, the developed statistical inquiry design did not respond to needs perceived by the participating teachers. First, the didactic design—a website—did not scaffold the teachers in assuming their new roles as directors of teaching inquiry. Second, the teachers did not recognise the differences between their regular statistical practices and new principles of the statistical inquiry, which was (in Paper I) described as missing quaternary contradictions, hence resulting in an absence of substantial and lasting changes of the teachers' practices related to the teaching of statistics. In contrast, the SRP study was realised with an intern sustained study and research dialectics—in line with the theoretical principles of SRP (Bosch & Winsløw, 2016). It is, however, interesting *if* or *how* the intervention influences the teachers' regular statistical practices and whether the participating teachers continue to implement SRP in primary school.

One main conclusion and recommendation (Paper I) was that (design research) interventions need to establish the interactions between teachers and researchers, which would include reflections and negotiations related to the objective of the design research—that is, students' development of statistical literacy and statistical reasoning. The result is in line with concepts such as co-learning (Jaworski, 2004; Wagnar, 1997) and earlier research recommendations (see Section 3.3). Another result was the need to make contradictions more explicit in design research projects, for example, by inviting teachers to participate in several of the design research phases and not exclusively as the implementor of the given design. The recommendations of Paper I were integrated into the design of the SRP study.

In the research project studying two teacher-driven statistical practices in two different institutions (Paper III), we had little preliminary information about teachers' planning and their way to conceptualise statistical inquiry approaches. Our hypothesis, however, was that paradidactic infrastructure, including shared beliefs about inquiry-oriented statistical teaching in the two institutions, could be reflected in the two inquiry lessons: the open, activity-based Danish inquiry lesson versus the Japanese structured problem-solving lesson.

In the project on sustainability of lesson study (Paper II), our main focus was on paradidactic infrastructure. We developed and studied a first model of paradidactic infrastructure in Denmark (PID), and studied the specific paradidactic elements that affected the cases of lesson study in a Danish context that were presented. The study did not examine, specifically, the effects of lesson study on teachers—how teachers and students developed didactic knowledge—but focused on the institutional conditions for realising lesson study, with reference to well-established practices in Japan. There, lesson study has indeed been a stable element of the paradidactic infrastructure for several decades. We examined the extent to which a different paradidactic infrastructure supports or opposes this new element.

The results suggest that implementing, developing and sustaining lesson study depend on engaging with municipal agents and school management in supporting and providing favourable conditions for conducting lesson studies, for example, prioritising it in scheduling classes and teachers' agenda, acknowledging that lesson study cannot be reduced to only realising certain surface actions for a short period of time, and recognising lesson study as a form of teachers' professional development that demands a sustained investment of expertise, time and effort. The first result is in line with the earlier research results presented in Section 3.3. The second research result indicates teachers' need for external lesson study resources, for example, detailed lesson plans and teacher guides that support how to plan and implement mathematical problems and mathematical inquiry. There are plenty practical guides in Denmark that explain the structure of lesson study (e.g., Kaas et al., 2017; Morgensen, 2015), but the dialectic between lesson study and resources are absent in the Danish context. In Japan, however, the experiences from carrying out lesson study informs commercially produced textbooks and are communicated in teacher journals (e.g., Watanabe, 2019), and textbooks and teacher guides include detailed descriptions and ideas to structure specific lessons, even designs of a blackboard (Tokyo Shoseki Mathematics Textbook Editorial Committee, 2016a; 2016b). The third research result concerns the critical need of the 'knowledgeable other'. In the two lesson study cases presented in Paper II, the knowledgeable others played a large role in supporting the teachers' professional development. However, because few in Denmark have relevant and long experience with lesson study, there is a lack of teachers, teacher educators and researchers to take on the role as knowledgeable other. In short, the development of teachers' professional knowledge in the context of lesson study requires an infrastructure going beyond the professional practices itself, which consists of shared conditions (institutional levels and media) that govern professional development in a long-term commitment of both teachers and institutions.

To sum up the study of paradidactic infrastructure, our results have shown that teachers' involvement in design research processes can indeed lead to genuine statistical inquiry, where students not only inquire into a problem, but also studied statistical works and validated more or less given answers found there. The paradidactic infrastructure affects the implementation of a statistical inquiry in several ways and taking these influences explicitly into consideration is crucial when trying to implement new developmental practices for teachers' professional development.

6 Conclusion and perspective for further research

In the present PhD project, we have investigated students' development of statistical praxeologies, ecologies of statistical inquiries and paradidactic infrastructure, here based on the following research questions:

RQ1: How can inquiry approaches such as investigative cycles and study and research paths support students' development of statistical praxeologies in school settings?

RQ2: What conditions and constraints for implementing these inquiry approaches in primary and lower secondary school mathematics—and specifically statistics—can be observed in different school contexts?

RQ3: How can paradidactical infrastructure support teachers to design, implement and evaluate teaching that involves statistical inquiry by students?

In our four studies, we observed how inquiry approaches led students to develop more or less automous statistical praxeologies. In the SRP study, we saw how the inquiry approach encouraged students to use and develop statistical logos, which is necessary when choosing and combining statistical techniques rather than simply applying given ones. The SRP approach contrasts in this way with common forms of teaching the subject in many countries, where techniques are practiced one by one, and are not further questioned or elaborated on. Our hypothesis was that a focus on questions, appearing "real" in the students' life worlds, *before* thinking about data collection and analysis, has a potential to create emphasis on statistical inference and statistical reasoning, also has a lot of potential in primary school.

Digital tools—especially the way they are used—can be seen as acting as both conditions and constraints. In Paper IV, we saw TinkerPlots used as crucial milieus to generate first hypotheses, to question and explore data and answers found in media and, hence, as a condition for the questioning and media-milieu dialectics. Tools like TinkerPlots indeed has several affordances for students to establish inquiry milieus, in which they can interact and analyse data. However, despite their special importance in relation to statistics, digital tools as such have not been our main research focus. The presented SRP seems to respond to some of the quandary related to how we can understand and improve statistical practices. Here, involving students in the dialectics of study and research in SRP can show how students use digital tools while not losing touch with statistical reasoning. Still, more research with a main focus on the developed digital inquiring milieus is needed. In more directed 'structured problem-solving', we have also observed statistical reasoning which was realised without the use of digital tools. One constraint in relation to digital tools, which is perhaps surprising or counterintuitive, is that the universal availability and recommendation of digital tools (as in Denmark) may render such black-box-free experiences of inquiry less likely or easy to realise, while at the same time connecting to elementary mathematics and especially numeracy.

A favourable condition (shared in all contexts) is an impetus from official documents to investigate 'big' questions, especially in cross-disciplinary activities where statistics can supply crucial tools to represent phenomena and extract precise interpretations of relevance to the questions. The constraints are more varied in different school contexts and according to the specific proposal or design for inquiry-oriented teaching, but certainly teachers' own experience with substantial statistical inquiry would seem to be a general challenge.

The role of the teacher is naturally crucial, and the 'usual' forms of this role (like the teacher mainly supervises student work with exercises) may oppose the role the teacher is assumed to take in relation to formats like SRP. We have observed a variety of assumptions about the nature of inquiry, which can exist among both teachers and researchers, and we have also seen the difficulty of achieving truly shared views. The term *implementation* reflects a specific view of the interaction between researchers and teachers, which is questioned or modified in some paradidactic infrastructures, such as lesson study and SRP based on such forms of collaboration (as explored in Paper IV). It would be interesting to investigate further to what extent this could lead to more consistent assumptions about the goals than what we found in Paper I.

We can imagine a dilemma if larger research and developmental projects fail to implement successful inquiry design, while small and intensively supported projects appear to succeed. The latter can be used in research but not to lead to major changes in a school system. The dilemma between large scale research and developmental projects on the one hand, and smaller and intensively supported projects on the other, also obviously relates to the issue of a sustainable paradidactic infrastructure, as explored in Paper I. We consider that this underlines the need to better understand macro-didactic phenomena, such as paradidactic infrastructure and levels of didactic codetermination, and to find ways to develop the paradidactic infrastructure, rather than focusing exclusively on teaching approaches and the current state of teachers' knowledge, as much recent research has tended to do.

7 References

Arnold, P., & Pfannkuch, M. (2019). Posing comparative statistical investigative questions. In G. Burrill & D. Ben-Zvi (Eds.), *Topics and trends in current statistics education research: International perspectives*, *ICME-13 Monographs* (pp. 173-195). Springer. https://doi.org/10.1007/978-3-030-03472-6_8

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

Bakker, A. (2018). *What is design research in education?* Routledge.

Ball, D. L., & Bass, H. (2003). Making mathematics reasonable in school. In J. Kilpatrick, W. Gary Martin & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 27-44). National Council of Teachers of Mathematics (NCTM).

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education, 59*(5), 389-407.

Bargagliotti, A., Franklin, C., Arnold, P., Gould, R., Johnson, S., Perez, L. & Spangler, D. A. (2020). *Pre-K-12 guidelines for assessment and instruction in statistics education II (GAISE II). A framework for statistics and data science education.* National Council of Teachers of Mathematics (NCTM).

Barquero, B., & Bosch, M. (2015). Didactic engineering as a research methodology: From fundamental situations to study and research paths. In A. Watson & M. Ohtani (Eds.), *Task design in mathematics education: An ICMI study 22* (pp. 249-272). Springer. https://doi.org/10.1007/978-3-319-09629-2_8

Barquero, B., Bosch, M., Florensa, I., & Ruiz-Munzón, N. (2022). Study and research paths in the frontier between paradigms. *International Journal of Mathematical Education in Science and Technology*, *53*(5), 1213-1229. http://doi:10.1080/0020739X.2021.1988166

Barquero, B., Bosch, M., & Gascón, J. (2013). The ecological dimension in the teaching of mathematical modelling at the university. *Recherches en Didactique des Mathématiques*, *33*(3), 307-338.

Barquero, B., Bosch, M., & Romo, A. (2015). A study and research path on mathematical modelling for teacher education. In K. Krainer & N. Vondrová (Eds.), *Proceedings of the Ninth Congress of the European Society for Research in Mathematics Education (CERME 9)* (pp. 809-815). Prague, Czech Republic: Charles University in Prague, Faculty of Education and ERME.

Barquero, B., Bosch, M., & Romo, A. (2018). Mathematical modelling in teacher education: Dealing with institutional constraints. *ZDM - International Journal on Mathematics Education*, *50*(1), 31-43. https://doi.org/10.1007/s11858-017-0907-z

Barquero, B., Florensa, I., & Ruiz-Olarría, A. (2019). The education of school and university teachers within the paradigm of questioning the world. In M. Bosch, Y. Chevallard, F. García & J. Monaghan (Eds.), *Working with the anthropological theory of the didactic in mathematics education. A comprehensive casebook* (pp. 189-212). Routledge.

Batanero, C., Burrill, G., & Reading, C. (2011). *Teaching statistics in school mathematics-challenges for teaching and teacher education: A joint ICMI/IASE study: the 18th ICMI study*. Springer.

Ben-Zvi, D., & Garfield, J. B. (2004). Statistical Literacy, Reasoning, and Thinking: Goals, Definitions, and Challenges. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 3-16). Kluwer Academic Publishers. https://doi.org/10.1007/1-4020-2278-6_1

Biehler, R., Ben-Zvi, D., Bakker, A., & Makar, K. (2013). Technology for enhancing statistical reasoning at the school level. In M. A. Clements, A. J. Bishop, C. Keitel, J. Kilpatrick & F. K. S. Leung (Eds.), *Third international handbook of mathematics education* (pp. 643-689) Springer.

Blomhøj, M. & Jensen, T. H. (2003). Developing mathematical modelling competence: Conceptual clarification and educational planning. *Teaching Mathematics and its applications*, *22*(3), 123-139.

Blum, W. & Fermi, R. (2009). Mathematical modelling: Can it be taught and learnt? *Journal of Mathematical Modelling and application, 8*(1), 45-58.

Bosch, M. (2018). Study and research paths: A model for inquiry. In B. Sirakov, P. N. De Souza, & M. Viana (Eds.), *Proceedings of the International Congress of Mathematicians 2018, Vol. 3*, (pp. 4001-4022). World Scientific. https://doi.org/10.1142/9789813272880_0210

Bosch, M., & Gascón, J. (2006). Twenty-five years of the didactic transposition. *ICMI bulletin*, *58*, 51-65.

Bosch, M., & Gascón, J. (2014). Introduction to the anthropological theory of the didactic (ATD). In A. Bikner-Ahsbahs & S. Prediger (Eds.), *Networking of theories as a research practice in mathematics education* (pp. 67-83). Springer. https://doi.org/10.1007/978-3-319-05389-9_5

Bosch, M., & Ruiz-Munzón, N. (2022). The curriculum problem and the paradigm of questioning the world. In Y. Chevallard, B. Barquero, M. Bosch, I. Florensa, J. Gascón, P. Nicolás, & N. Ruiz Munzón (Eds.), *Advances in the anthropological theory of the didactic* (pp. 173-177). Birkhäuser. https://doi.org/10.1007/978-3-030-76791-4

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*, *35*(2), 357-399.

Bronfenbrenner, U. (1999). Environments in developmental perspective: Theoretical and operational models. In S. L. Friedman & T. D. Wachs (Eds.), *Measuring environment across the life span: Emerging methods and concepts* (pp. 3-28). American Psychological Association.

Brousseau, G. (1997). *The theory of didactical situations in mathematics*. (N. Balacheff, M. Cooper, R. Sutherland, & V. Warfield (Eds. and Trans.). Kluwer. https://doi.org/10.1007/0-306-47211-2

Brousseau, G., & Gibel, P. (2005). Didactical Handling of Students' Reasoning Processes in Problem Solving Situations. In C. Laborde, M.-J. Perrin-Glorian, & A. Sierpinska (Eds.), *Beyond the apparent banality of the mathematics classroom* (pp. 13-58). Springer.

Burgess, T. A. (2011). Teacher knowledge of and for statistical investigations. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics – Challenges for teaching and teacher education: A joint ICMI/IASE Study: The 18th ICMI Study* (pp. 259-270). Springer.

Burton, L. (2004). *Mathematicians as enquirers: learning about learning mathematics.* Springer Science + Business Media.

Chen, J. C., Lin, F. L., & Yang, K. L. (2018). A novice mathematics teacher educator–researcher's evolution of tools designed for in-service mathematics teachers' professional development. *Journal of Mathematics Teacher Education*, *21*(5), 517-539.

Chevallard, Y. (1992). Fundamental concepts in didactics: Perspectives provided by an anthropological approach. In R. Douady & A. Mercier (Eds.), *Research in didactique of mathematics. Selected papers* (pp. 131-167). La Pensée Sauvage.

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*, *19*(2), 221-266.

Chevallard, Y. (2002). Organiser l'étude 3. Ecologie et regulations. In J.-L. Dorier, M. Artaud, M. Artigue, R. Berthelot & R. Floris (Eds.), *Actes de la 11e École d'Été de Didactique des Mathématiques* (pp. 41-56). La Pensée Sauvage.

Chevallard, Y. (2006). Steps towards a new epistemology in mathematics education. In M. Bosch (Ed.), *Proceedings of the Fourth Congress of the European Society for Research in Mathematics Education (CERME 4)* (pp. 22-30). Sant Feliu de Guíxols, Spain: FUNDEMI IQS – Universitat Ramon Llull and ERME.

Chevallard, Y. (2007). Passé et présent de la Théorie Anthropologique du Didactique. In L. Ruiz Higueras, A. Estepa & F. J. García (Eds.), *Sociedad, escuela y matemáticas. Aportaciones de la Teoría Antropológica de lo Didáctico* (pp. 705-746). Publicaciones de la Universidad de Jaén.

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

Chevallard, Y. (2019). Introducing the anthropological theory of the didactic: An attempt at a principled approach. *Hiroshima Journal of Mathematics Education*, *12*, 71-114.

Chevallard, Y. (2022) Challenges and advances in teacher education within the ATD. In Y. Chevallard, B. Barquero, M. Bosch, I. Florensa, J. Gascón, P. Nicolás, & N. Ruiz Munzón (Eds.), *Advances in the anthropological theory of the didactic* (pp. 81-89). Birkhäuser. https://doi.org/10.1007/978-3-030-76791-4

Cobb, G. W., & Moore, D. S. (1997). Mathematics, statistics, and teaching. *The American mathematical monthly*, *104*(9), 801-823.

Cobb, P., Jackson, K., & Dunlap, C. (2015). Design research: An analysis and critique. In L. English & D. Kirshner (Eds.), *Handbook of international research in mathematics education* (pp. 493-515). Routledge.

Cobb, P., & McClain, K. (2004). Principles of instructional design for supporting the development of students' statistical reasoning. In J. Garfield & D. Ben-Zvi (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 375-395). Springer.

Cobb, P., McClain, K., de Silva Lamberg, T., & Dean, C. (2003). Situating teachers' instructional practices in the institutional setting of the school and district. *Educational Researcher*, *32*(6), 13-24.

Corica, A. R., & Otero, M. R. (2015). The mathematics teacher's profession: the perspective of future. *European Journal of Science and Mathematics Education, 3*(2), 145-158.

da Ponte, J. P. (2011). Preparing teachers to meet the challenges of statistics education. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics: Challenges for teaching and teacher education. A Joint ICMI/IASE Study* (pp. 299-309). Springer.

da Ponte, J. P., & Noll, J. (2018). Building capacity in statistics teacher education. In D. Ben-Zvi, K. Makar, & J. B. Garfiels (Eds.) *International handbook of research in statistics education* (pp. 433-455). Springer, Cham.

del Mas, R. C. (2004). A comparison of mathematical and statistical reasoning. In D. Ben-Zvi & J. B. Garfoeld (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 79-95). Kluwer.

Demonstrationsskoler.dk. (2015a). *Unge og medier - statistik.* [Young people and the media – statistics] Retrieved December 13, 2022, from http://auuc.demonstrationsskoler.dk/materialer/innovation/forloeb3/om

Demonstrationsskoler.dk. (2015b). *Unge og medier - statistik.* [Young people and the media – statistics] Retrieved December 13, 2022, from http://auuc.demonstrationsskoler.dk/materialer/innovation/forloeb3/faser/hvordan-bruges-statistik-i-hverdagen

Dvir, M., & Ben-Zvi, D. (2018). The role of model comparison in young learners' reasoning with statistical models and modeling. *ZDM - International Journal on Mathematics Education*, *50*, 1183-1196. https://doi.org/10.1007/s11858-018-0987-4

Eichler, A., & Zapata-Cardona, L. (2016). *Empirical research in statistics education. ICME-13 Topical Surveys.* Springer Open. https://doi.org/10.1007/978-3-319-38968-4

Engeström, Y. (1987*). Learning by expanding: An activity-theoretical approach to developmental research*. Orienta-Konsultit.

Engeström, Y (1999). Activity theory and individual and social transformation, In: Y. Engeström, R. Miettinen, & R.–L. Punamäki (Eds). *Perspectives on activity theory* (pp. 19-38). Cambridge University Press.

Engeström, Y. (2001). Expansive learning at work: Toward an activity theoretical reconceptualization. *Journal of Education and Work*, *14*(1), 133-156.

Engeström, Y., & Sannino, A. (2010). Studies of expansive learning: Foundations, findings and future challenges. *Educational Research Review, 5*(1), 1-24.

EVA (2012). *Fælles Mål. En undersøgelse af lærere brug of Fælles Mål.* Danmarks Evalueringsinstitut.

EVA (2015). *Undersøgelse af læreres og forældres forståelse af forenklede Fælles Mål*. Danmarks Evalueringsinstitut.

Fielding-Wells, J. (2010). Linking problems, conclusions and evidence: Primary students' early experiences of planning statistical investigations. In C. Reading (Ed.), *Proceedings of the 8th international conference on teaching statistics (ICOTS8).* Voorburg, The Netherlands: International Statistical Institute.

Fienberg, S. E. (2014). What is statistics? *Annual review of statistics and its application*, *1*, 1-9.

Fitzallen, N., & Watson, J. (2010) Developing statistical reasoning facilitated by Tinkerplots. In K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9).* Voorburg, the Netherlands: International Statistical Institute.

Florensa, I., Bosch, M., Gascón, J., & Mata, M. (2016). SRP design in an elasticity course: The role of mathematic modelling. In E. Nardi, C. Winsløw, & T. Hausberger (Eds.), *Proceedings of the first Conference of the International Network for didactic research in university mathematics* (pp. 191-200). University of Montpellier and INDRUM.

Flyvbjerg, B. (2006). Five misunderstandings about case-study research. *Qualitative Inquiry*, *12*(2), 219-245.

Foot, K., & Groleau, C. (2011). Contradictions, transitions, and materiality in organizing processes: An activity theory perspective. *First Monday*, *16*(6), 1-18.

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). *Guidelines for assessment and instruction in statistics education (GAISE) report: A pre-K–12 curriculum framework.* National Council of Teachers of Mathematics (NCTM).

Frischemeier, D., & Biehler, R. (2015). Preservice teachers' statistical reasoning when comparing groups facilitated by software. In K. Krainer & N. Vondrová (Eds.), *Proceedings of the Ninth Congress of the European Society for Research in Mathematics Education (CERME 9)* (pp. 643-650). Prague, Czech Republic: Charles University in Prague, Faculty of Education and ERME.

Freudenthal, H. (2006). *Revisiting mathematics education: China lectures*. Kluwer Academic Publisher.

Fullan, M., Bennett, B., & Rolheiser-Bennett, C. (1990). Linking classroom and school improvement. *Educational leadership*, *47*(8), 13-19.

Fullan, M. (2005). The meaning of educational change: A quarter of a century of learning. In A. Lieberman (Ed.), *The roots of educational change: International handbook of educational change* (pp. 202-216). Springer Science & Business Media.

Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International Statistical Review*, *70*(1), 1-25. https://doi.org/10.1111/j.1751-5823.2002.tb00336.x

Gal, I. (2003). Teaching for Statistical Literacy and Services of Statistics Agencies, *The American Statistician*, 57(2), 80-84.

Gal, I. (2019). Understanding statistical literacy: About knowledge of contexts and models. In J. M. Contreras, M. M. Gea, M. M. López-Martín & E. Molina-Portillo (Eds.), *Actas del Tercer Congreso. Internacional Virtual de Educación Estadística* (pp. 1-15). Available online: www.ugr.es/local/fqm126/civeest.html

Garfield, J., delMas, R. & Chance, B. (2003). The Web-based ARTIST: Assessment Resource Tools for Improving Statistical Thinking. Paper presented at *the Annual Meeting of the American Educational Research Association* (pp. 1-22). Available online: https://www.causeweb.org/cause/archive/artist/articles/AERA_2003.pdf

Gattuso, L., & Ottaviani, M. G. (2011). Complementing mathematical thinking and statistical thinking in school mathematics. In C. Batanero, G. Burrill & C. Reading (Eds.), *Teaching Statistics in School Mathematics-Challenges for Teaching and Teacher Education: A Joint ICMI/IASE Study: The 18th ICMI Study*, (pp. 121-132). Springer.

Godino, J. D., Ortiz, J. J., Roa, R. & Wilhelmi, M. R. (2011). Models for Statistical Pedagogical Knowledge. In C. Batanero, G. Burrill & C. Reading (Eds.), *Teaching Statistics in School* 271 *Mathematics-Challenges for Teaching and Teacher Education: A Joint*

ICMI/IASE Study (pp. 271-282). Springer Science + Business Media. https://doi.org/10.1007/978-94-007-1131-0_27

Gonzáles, O. (2014). *Examining Venezuelan secondary school mathematics teachers' statistical knowledge for teaching: Focusing on the instruction of variability-related concepts.* [Doctoral dissertation; Hiroshima University]. Available online: https://core.ac.uk/download/pdf/222952772.pdf

Groth, R. E. (2007). Research commentary: toward a conceptualization of statistical knowledge for teaching. *Journal for research in Mathematics Education*, *38*(5), 427-437.

Groth, R. W. (2017). Developing statistical knowledge for teaching during design-based research. *Statistics Education Research Journal, 16*(2), 376-396.

Groth R. E. & Berger, J. A. (2013). Mapping the structure of knowledge for teaching nominal categorical data analysis. *Educational Studies of Mathematics, 83*, 247-265. https://doi.org/10.1007/s10649-012-9452-4

Hegedus, S., & Tall, D. (2015). Foundations for the future: The potential of multimodal technologies for learning mathematics. In L. D. English & D. Kirshner (Eds.), *The third handbook of international research in mathematics education* (pp. 543-562). Routledge.

Infa.dk, (1985). *Informatik i skolens fa*g [Informatics in school subjects]. Retrieved January 15, 2023 from http://www.infa.dk

Jaworski, B. (2003). Inquiry as a pervasive pedagogic process in mathematics education development. In M. A. Mariotti (Ed.), *European Research in Mathematics Education III: Proceedings of the Third Conference of the European Society for Research in Mathematics Education (CERME 3)* (pp. 1-10). Bellaria, Italy: University of Pisa and ERME.

Jaworski, B. (2004). Grappling with complexity: Co-learning in inquiry communities in mathematics teaching development. In M. Høines & A. Fuglestad (Eds.), *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education, 1* (pp. 17-36). PME

Jessen, B. E. (2014). How can study and research paths contribute to the teaching of mathematics in an interdisciplinary setting? *Annales de Dedactique et de Science Cognitives*, *19*, 199–244.

Jessen, B. E. (2017). How to generate autonomous questioning in secondary mathematics teaching? *Recherches en Didactique des Mathematiques, 37*(2-3), 217-246.

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: The cases of Denmark and Japan. In M. Bosch, Y. Chevallard, F. García, & J. Monaghan (Eds.), *Working with the anthropological theory of the didactic in mathematics education* (pp. 118-138). Routledge. https://doi.org/10.4324/9780429198168

Kaas, T., Kristiansen, H., Møller, H., Skott, C. K. & Østergren-Olsen, D. (2017). *Lektionsstudiebogen*. [The lesson study book]. Hans Reitzels Forlag.

Kaptelinin, V., Kuutti, K., & Bannon, L. (1995). Activity theory: Basic concepts and applications. In B. Blumenthal, J. Gornostaev, & C. Unger (Eds.), *Human-Computer Interactions* (pp. 189-201). Springer.

KiDM.dk. (2017). *Mål.* [Aims of 3rd intervention]. Retrieved December 13, 2021, from http://kidm.dk/matematik/laerer/matematik-laerer/oversigt/de-tre-indsatser/indsats-3/binde-knuder/maalsaetning/

Kieran, C., Doorman, M., & Ohtani, M. (2015). Frameworks and principles for task design. In A. Watson & M. Ohtani (Eds.), *Task design in mathematics education: An ICMI study 22* (pp. 19-81). Springer Nature.

Konold, C., & Kazak, S. (2008). Reconnecting data and chance. *Technology innovations in statistics education*, *2*(1), 1-37.

Konold, C., & Pollatsek, A. (2002). Data analysis as the search for signals in noisy processes. *Journal for research in mathematics education*, *33*(4), 259-289.

Lampert, M. (1990). When the problem is not the question and the solution is not the answer: Mathematical knowing and teaching. *American educational research journal*, *27*(1), 29-63.

Lee, H., Tran, D., & Nickel, J. (2015) Simulation approaches for informal inference: Models to develop understanding. In K. Krainer & N. Vondrová (Eds.), *Proceedings of the Ninth Congress of the European Society for Research in Mathematics Education (CERME 9)* (pp. 707-717). Prague, Czech Republic: Charles University in Prague, Faculty of Education and ERME.

MacGillivray, H., & Pereira-Mendoza, L. (2011). Teaching statistical thinking through investigative projects. In C. Batanero, G. Burrill & C. Reading (Eds.), *Teaching Statistics in School Mathematics-Challenges for Teaching and Teacher Education: A Joint ICMI/IASE Study: The 18th ICMI Study* (pp. 109-120). Springer.

Makar, K., & Ben-Zvi, D. (2011). The role of context in developing reasoning about informal statistical inference. *Mathematical Thinking and Learning*, *13*(1-2), 1-4.

Mallows, C. (1998). The zeroth problem. *The American Statistician*, *52*(1), 1-9.

Markulin, K., Bosch, M., & Florensa, I. (2021). Project-based learning in statistics: A critical analysis. *Caminhos da Educação Matemática em Revista*, *11*(1), 200-220.

Markulin, K., Bosch, M., Florensa, I., & Montañola, C. (2022). The evolution of a study and research path in statistics. *Épijournal de Didactique et Epistémologie des Mathématiques pour l'Enseignement Supérieur*, *1*, 1-16.

Martins, M., Monteiro, C., & Carvalho, C. (2015). How teacher understand sampling when using Tinkerplots. In K. Krainer & N. Vondrová (Eds.), *Proceedings of the Ninth Congress of the European Society for Research in Mathematics Education (CERME 9)* (pp. 715-721). Prague, Czech Republic: Charles University in Prague, Faculty of Education and ERME.

Maxara, C., & Bielher, R. (2007) Constructing stochastic simulations with a computer tool: Students' competencies and difficulties. In D. Pitta-Pantazi & C. Philippou (Eds.), *Proceedings of the Fifth Congress of the European Society for Research in Mathematics Education (CERME 5)* (pp. 762-771). Larnaca, Cyprus: University of Cyprus and ERME.

Miyakawa, T., & Winsløw, C. (2009). Didactical designs for students' proportional reasoning: an "open approach" lesson and a "fundamental situation". *Educational Studies in Mathematics*, *72*, 199-218.

Miyakawa, T., & Winsløw, C. (2013). Developing mathematics teacher knowledge: The paradidactic infrastructure of "open lesson" in Japan. *Journal of Mathematics Teacher Education*, *16*, 185-209.

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

Morgensen, A. (2015). *Lektionsstudier i skolen. Kollegial sparring gennem fælles studier*. [Lesson studies. Sharing knowledge]. Dafolo.

Moore, D. S. (1998). Statistics among the liberal arts. *Journal of the American Statistical Association*, *93*(444), 1253-1259.

Moore, D. S., & Cobb, G. W. (2000). Statistics and mathematics: Tension and cooperation. *The American Mathematical Monthly*, *107*(7), 615-630.

Niss, M. and Jensen, T. H. (2002). *Kompetencer og matematiklæring - ideer og inspiration til udvikling af matematikundervisningen i Danmark.* [Competences and mathematics - ideas and inspiration for the development of mathematics education in Denmark.] Undervisningsministeriet.

Noll, J., & Kirin, D. (2017). TinkerPlots TM model construction approaches for comparing two groups: Student perspectives. *Statistics Education Research Journal*, *16*(2), 213-234.

Noll, J., & Shaughnessy, J. M. (2012). Aspects of students' reasoning about variation in empirical sampling distributions. *Journal for Research in Mathematics Education*, *43*(5), 509-556.

OECD (2016), *Supporting Teacher Professionalism: Insights from TALIS 2013*, TALIS, OECD Publishing, https://doi.org/10.1787/9789264248601-en.

Pfannkuch, M. (2005). Thinking tools and variation. *Statistics Education Research Journal*, *4*(1), 83-91.

Pfannkuch, M., & Rubick, A. (2002). An exploration of students' statistical thinking with given data. *Statistics Education Research Journal*, *1*(2), 4-21.

Pfannkuch, M., & Wild, C. (2004). Towards an understanding of statistical thinking. In D. Ben-Zvi & J. Garfield (Eds.). *The challenge of developing statistical literacy, reasoning and thinking*, (pp. 17-46). Springer.

Rasmussen, K. (2016). The direction and autonomy of interdisciplinary study and research paths in teacher education. *REDIMAT*, *5*(2), 158-179. http://dx.doi.org/10.4471/redimat.2016.1753

Rasmussen, L. (1987). *Stokastik – et propædeutisk forskningsprogram 1 & 2.* [propaedeutic stochastical research programme 1&2]. Danmarks Lærerhøjskole. Matematisk Institut.

Roback, P., Chance, B., Legler, J., & Moore, T. (2006). Applying Japanese lesson study principles to an upper-level undergraduate statistics course. *Journal of Statistics Education, 14*(2), 1-20.

Rowland, T., & Ruthven, K. (2011). Introduction: Mathematical knowledge in teaching. In T. Rowland & K. Ruthven (Eds.) *Mathematical knowledge in teaching* (pp. 1-5). Springer.

Rumsey, D. (2002). Statistical Literacy as a Goal for Introductory Statistics Courses, *Journal of Statistics Education 10*(3), 1-12.

Serrano, L., Bosch, M., & Gascón, J. (2010). Fitting models to data: The mathematising step in the modelling process. In V. Durand-Guerrier, S. Soury-Lavergne, & F. Arzarello (Eds.), *Proceedings of the Sixth Congress of the European Society for Research in Mathematics Education (CERME 6)* (pp. 2186-2195). Lyon, France: Institut National de Recherche Pédagogique and ERME.

Shulman, L. S. (1986). Those who understand: knowledge grown in teaching. *Educational Researcher*, *15*(2), 4-14.

Shulman, L. S. (1987). Knowledge and teaching: Foundation of the new reform. *Harvard Educational Review*, *57*(1), 1-22.

Skott, C. K. (2018). Konsulenters fortællinger om en intervention. [Consultants' stories about an intervention]. In J. Bundsgaard, M. Georgsen, S. Graf, T. I. Hansen, & C. K. Skott (Eds.), *Skoleudvikling med it. Forskning i tre demonstrationsskoleforsøg I* [School development with ICT. Research in three demonstration school projects I] (pp. 167-185). Aarhus Universitetsforlag.

Skott, C. K., & Østergaard, C. H. (2016). How does an ICT-competent mathematics teacher benefit from an ICT-integrative project? In C. Csíkos, A. Rausch, & J. Szitányi (Eds.), *Proceedings of the 40th conference of the international group for the psychology of mathematics education, 4* (pp. 211-218). PME.

Skott, C. K., & Østergaard, C. H. (2018). En matematiklærers professionelle udvikling. [Professional development of a mathematics teacher]. In J. Bundsgaard, M. Georgsen, S. T. Graf, T. I. Hansen & C. K. Skott (Eds.), *Innovativ undervisning med it. Forskning i tre demonstrationsskoleforsøg II* [Innovative teaching with ICT. Research in three demonstration school projects II] (pp. 157-178). Aarhus Universitetsforlag.

Spiegelhalter, D. J. and Riesch, H. (2008). Bacon sandwiches and middle-class drinkers: the risk of communication risk, *Significance 5(1),* 30-33. https://doi.org/10.1111/j.1740-9713.2008.00277.x

Stigler, J. & Hiebert, J. (1999). *The teaching gap*. New York: The Free Press.

Stylianides, G. J. (2008). An analytic framework of reasoning-and-proving. *For the learning of mathematics*, *28*(1), 9-16.

Thomson, P. (2007). *Whole school change. A literature review.* Arts Council England.

Tokyo Shoseki Mathematics Textbook Editorial Committee (2016a). *Sansu Math, Textbook, grade 4*. Koyo Publishing Inc.

Tokyo Shoseki Mathematics Textbook Editorial Committee (2016b). *Sansu Math, Teacher's guide 4A*. Koyo Publishing Inc.

Tukey, J. W. (1977). *Exploratory data analysis vol 2.* Addison-Wesley Publishing Company.

Turner, F. & Rowland, T. (2011). The knowledge quartet as an organizing framework for developing and deeoening teachers' mathematical knowledge. In T. Rowland & K. Ruthven (Eds.), *Mathematical knowledge in teaching* (pp. 195-212). Springer.

Van den Heuvel-Panhuizen, M. & Drijvers, P. (2014). Reference Realistic Mathematics Education. In S. Lerman (Ed.), *Encyclopedia of Mathematics Education* (pp. 521-525). Springer Science+Business Media. http://doi.org/10.1007/978-94-007-4978-8

Verhoeven, P. (2006). Statistics Education in the Netherlands and Flanders: An Outline of Introductory Courses at Universities and Colleges, In *Proceedings of the Seventh International Conference on Teaching Statistics (ICOTS7)* (pp.1-5). Available at: https://www.stat.auckland.ac.nz/~iase/publications/17/3A4_VERH.pdf

Vygotskij, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Wagner, J (1997). The Unavoidable Intervention of Educational Research: A Framework for Reconsidering Research-Practitioner Cooperation. *Educational Researcher, 26*(7), 13-22.

Wake, G., Swan, M., & Foster, C. (2016). Professional learning through the collaborative design of problem-solving lessons. *Journal of Mathematics Teacher Education*, *19*, 243-260.

Watanabe, T. (2019). Lesson Study and Textbook Revisions: What Can We Learn from the Japanese Case?. In R. Huang, A. Takahashi & J. P. da Ponte (Eds.) *Theory and Practice of Lesson Study in Mathematics: An International Perspective* (pp. 183-200). Springer.

Watson, J. M. (1997). Assessing Statistical Thinking Using the Media, In I. Gal & J. B. Garfield (Eds.) *The Assessment Challenge in Statistics Education* (pp. 107-121). IOS Press and The International Statistical Institute.

Watson, J., & Callingham, R. (2003). Statistical literacy: A complex hierarchical construct. *Statistics Education Research Journal*, *2*(2), 3-46.

Watson, J., & Donne, J. (2009). TinkerPlots as a research tool to explore student understanding. *Technology Innovations in Statistics Education*, *3*(1). 1-35.

Wei, G. (2017). Dynamics of teacher professional learning: From cultural-historical activity theoretical perspective. *Culture, Biography and Lifelong Learning*, *3*(1), 35-50.

Weiland, T. (2017). Problematizing statistical literacy: An intersection of critical and statistical literacies. *Educational Studies in Mathematics*, *96*(1), 33-47.

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, *67*(3), 223-248. https://doi.org/10.1111/j.1751-5823.1999.tb00442.x

Wild, C. J., 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*, (pp. 5-36). Springer.

Williams, J., & Wake, G. (2007). Black boxes in workplace mathematics. *Educational Studies in Mathematics*, *64*(3), 317-343.

Winsløw, C. (2012). 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. Mathematics curriculum materials and teacher development* (pp. 291-304). Springer.

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. https://doi.org/10.1007/s10649-012-9453-3

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

Yin, R. K. (2014). *Case Study Research: Design and methods* (5th ed.). Thousand Oaks: SAGE Publications.

8 Research papers

8.1 Paper I: Contradicting activity systems – Learning from large-scale interventions that fail to change mathematics teaching practice as intended

Contradicting activity systems – Learning from large-scale interventions that fail to change mathematics teaching practice as intended

Dorte Moeskær Larsen[1], Camilla Hellsten Østergaard[2] & Klaus Rasmussen[3]

[1] University College Lillebaelt, Odense, Denmark, dmla@ucl.dk

[2]Universit of Copenhagen & University College Copenhagen, Copenhagen, Denmark, choe@kp.dk

[3] University College Copenhagen, Copenhagen, Denmark, qlausras@gmail.com

Abstract

The implementation of large-scale intervention and development projects is often problematic, and the impacts of such projects usually fall somewhat short of what was expected. Additionally, the rationalities of intervention projects are not carried over into classroom teaching as directly as expected. This problem is generally known, but comprehensive explanations continue to elude the research community at large. Using the theory of cultural-historical activity theory (CHAT), we propose that the heart of the problem lies in the expansive learning process that teachers undergo. This process is driven by unrecognised contradictions in terms of cultural and historical origin, which are fundamentally different from the processes governing the projects. We analyse two cases taken from two large Danish professional development projects. In each case, we focus on a teacher as part of two activity systems ('the project' and 'the classroom') and how the contradictions within and between these shape learning through epistemic actions. The results indicate the importance of making these contradictions apparent and accessible to everyone in the activity systems. Because of these various contradictions, the agency conferred upon teachers leads to unintended outcomes.

Keywords: mathematic education, professional development, implementation, statistical reasoning, cultural-historical activity theory

Introduction

Researchers and educational reformers have developed countless interventions to improve classroom teaching and, consequently, students' learning of mathematics. Interventions often begin as pilot studies with few participants, develop into interventions with hundreds of participants, and end as implementations or educational reforms involving parts of or whole educational systems. Interventions have been studied to determine what happens when they are implemented in classrooms, that is, whether the interventions worked; for whom they worked; and where, when, and why they worked (Century & Cassata, 2016).

Maaß and Artigue (2013) defined 'implementation' as setting a planned intervention or innovation in motion so that new research results can lead to the development of new interventions, which are further disseminated. Century and Cassata (2016) describe implementation research as the study of the several efforts on the part of education researchers to bring about successful and lasting change. Several researchers have expressed the idea that widespread implementation projects often have disappointing results in terms of their impact on classroom teaching practices (Century & Cassata, 2016). Skott (2009, 2020) argues that the field of intervention research is often overly optimistic regarding the potential impact of interventions. Other researchers (e.g., Cai et. al., 2017; Jankvist et al., 2020) argue for the potential of addressing mathematics education from an implementation point of view, but to do so, we must broaden the existing implementation research frameworks to better understand the various aspects of implementation.

The intention of this paper is not to attribute the challenges of implementation to a lack of knowledge or willingness on part of the teachers, nor is it to suggest that implementation is flawed. Rather, we follow Terhart (2013) in accepting that the perceived success or failure of educational implementations is intrinsically difficult to estimate because it is a matter of interpretation and *post-hoc* reconstruction. Our intention is to analyse the implementation processes from various perspectives and thus obtain a broader perspective that includes teachers, students, researchers, and communities. We thus analyse a 'classroom perspective' and a 'project perspective' to determine how the interaction between these shapes professional development. To do this, we draw on cultural-historical activity theory (CHAT) (Engeström, 2001; Engeström & Sannino, 2010), which provides us with the theoretical tools we need to deal with the multitude of factors involved in attempting to change educational practices via interventions. Using this framework, we adopt the position that these factors are cultural and historical, as well as describable within activity systems. Other frameworks would propose the factors to be, e.g., institutional or epistemological in nature (Bosch & Gascón, 2014). We have chosen CHAT for several reasons, firstly its comprehensiveness in describing two or more loci of human activity, i.e., activity systems, that interact. Secondly, the CHAT framework is a unit of analysis that focuses on how implementations must be embedded and contextualised in the meaningful life activity of the participants.

In the following, we describe the elements of CHAT, which enables us to state our research question with theoretical precision. Then, we provide a methodology section, in which the overall contexts of the two Danish projects are described. From each project, we present one case linked to a teacher. Further characteristics of the intervention projects are presented at appropriate places throughout the analysis of the cases.

Theoretical considerations

The professional development that emerges from participation in educational intervention projects can be seen as an expansion of existing knowledge in essentially unknown directions. The notion of *expansive learning* focuses on communities as learners and on the transformation and creation of knowledge. The theory of expansive learning focuses on the transformation and creation of culture, and knowledge is constructed as a collective activity (Engeström & Sannino, 2010). In expansive learning, learners are integral parts of a wider network in which the construction of essentially new knowledge takes place. The network of relationships is not merely between individuals, who are called *subjects*, but also encompasses *artefact, rules, object, community,* and the *division of labour*. Taken together, these make up an *activity system*.

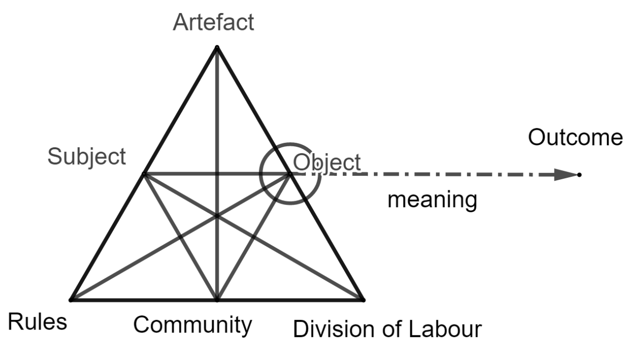


Figure 1 General model of an activity system.

Activity systems are often depicted by a triangular illustration featuring *nodes* (Figure 1) (see Cole & Engeström, 1993; Engeström & Sannino, 2010). Activity systems are networks of the above-mentioned sociocultural nodes, with each having complex structures and interactions that shape the entirety of the system. In fact, CHAT is based on Vygotsky's (1978) and Leontiev's (1978) ideas of object-oriented action, in which learning is not simply an activity located at the level of the private individual human mind. Rather, learning takes place within a system in which one or more subjects act upon an object of activity, e.g., the professional development of teachers, with the purpose of generating an outcome, e.g., an increase in teachers' capability for competent pedagogical action. Expansive learning thus manifests primarily as changes in the object of the collective activity. The *subject* refers to individuals or a group of people engaged in an activity whose position and point of view are adopted as the perspective for the analysis (this is the primary distinction between the subject and community nodes). The *artefacts* of the activity system mediate every human action and experience (Vygotskij, 1978) and can be internal, external, psychological, or material. Artefacts can be embedded in *material*, e.g., curriculums, textbooks, and portfolios, or remain primarily in their *mental* form, e.g., jargon, methods, learning theories. When a subject uses an artefact, that artefact affects the way the subject carries out an action (Kaptelinin et al., 1995). The subject likewise subtly shapes and enhances the artefact to make it more effective and useful, which again changes the way the subject completes a task. The *community* refers to a wider group of individuals or subgroups who share interests or cultures that interact within an activity system. The community members negotiate the *division of labour* and the *rules* for participation within the community. The rules can be implicit or explicit regulations, norms, conventions, and standards that constrain actions within the activity system. The division of labour can be divided into a "horizontal" division of tasks and a "vertical" division of power and status (Engeström & Sannino, 2010).

Within an activity system, contradictions build up over time and may lead to a transformation of the system (Engeström, 2001; Engeström et al., 1999). Contradictions can accumulate within and between activity systems (Engeström, 2011). Contradictions can be brought about consciously, leading to a change in one node or multiple nodes in the activity system, which eventually transforms the entire system. Contradictions are also the foundation for innovative attempts to change activities: 'Contradictions are a sign of richness in the activity system' (Foot, 2014, p. 337). Contradictions become the driving force behind expansive learning when they are addressed in a way that identifies and uses a (new) object as a motive for learning. This motive compels subjects to transform collective activity because it makes personal sense and holds the promise of increased agency in classroom processes.

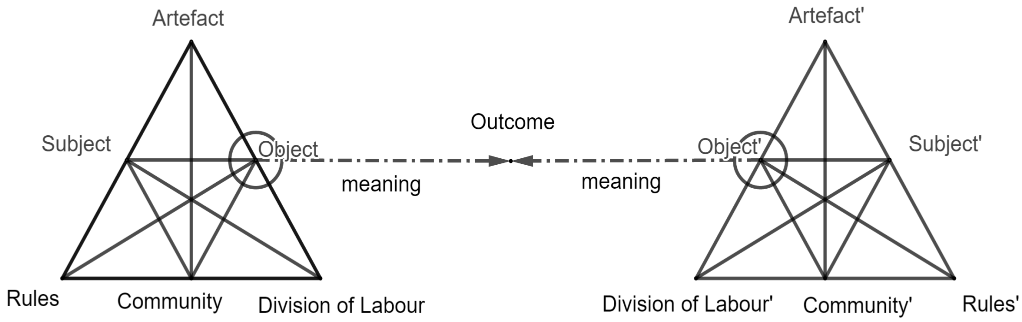


Figure 2 Model of interacting activity systems. The connecting arrows are traditionally drawn between objects, while the "node of contact" can be any or several.

In the process of learning, the subject begins to question the recent practices and logic of their activity and participation in the activity system, and this gradually expands into a collective movement (Cole & Engeström, 1993; Engeström, 2008). A learning process often begins, for subjects in an activity system, when one activity system 'comes into contact' with another (e.g., when there is a shared objective or a new rule that applies to both, or when an artefact from another activity system is introduced) (Figure 2). When designing interventions, it is important that the interventions respond to and build on emerging contradictions in the activity systems (Engeström, 2011).

Expansive learning involves a cyclic process with seven epistemic actions, in which

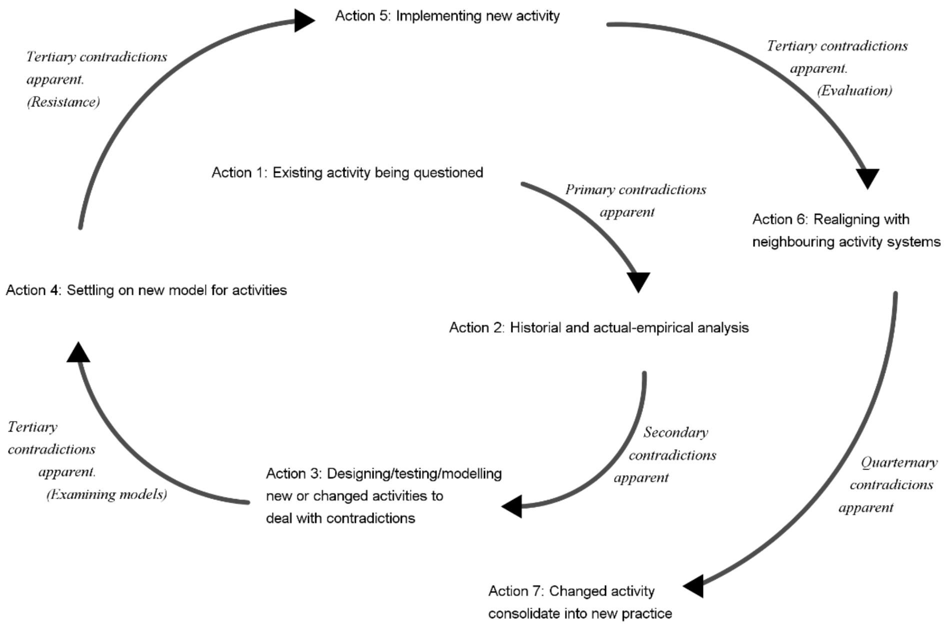


Figure 3 Ideal-typical sequence of actions and associated contradictions in an expansive learning cycle

contradictions 'propel' learning or shape what is learned (Figure 3).

There are four types of contradictions: primary contradictions appear within any node of an activity system (e.g., new rule–old rule). Secondary contradictions occur between two or more nodes (e.g., new object–old artefact). Tertiary contradictions can be observed between a newly established mode of activity and the remnants of the previous modes of activity. Quaternary contradictions can be seen between a newly reorganised activity and its neighbouring activity system(s) (Foot & Groleau, 2011).

The process of expansive learning is a process of successively addressing contradictions, e.g., an established teaching practice is deemed less desirable when considering a curricular change. The cycle is not a universal formula for stages, and it may be difficult or even impossible to 'find a concrete collective learning process which would cleanly follow the ideal-typical model' (Engeström & Sannino, 2010, p. 7). The logic of the expansive cycle is that a new cycle begins when an existing pattern of activity is questioned. The cycle ends when a new pattern of activity has been consolidated. When using the expansive cycle model in analyses, the model has been used to identify the various actions and thus clarify the process itself (Foot & Groleau, 2011). Engeström (1999, p. 33) argues that 'it seems promising to analyse these cycles in terms of the stepwise formation and resolution of internal contradictions in activity systems. However, only a few studies in mathematics education have hitherto examined the *contradictions* within or between activity systems: Chen, Lin, and Yang (2018) investigated the contradictions between a novice mathematics teacher educator and mathematics teachers in an in-service course, while Williams and Wake (2007) considered contradictions between identifying mathematics in the systems of college and the workplace.

The expansive learning process is divided into seven actions (Figure 3). The first action is to question and criticise aspects of existing practice. The second action involves using historical analysis to explain how the present state of affairs came about, as well as an actual empirical analysis, which is intended to 'explain the situation by constructing a picture of its inner systemic relations' (Engeström & Sannino, 2010, p. 7). The third action is to model a solution by, for example, proposing or designing a new or revised activity that offers a solution to the contradictions. The fourth action involves examining and elaborating on the model to understand its potential and limitations. The fifth action is to implement the new model. The sixth action is to reflect on and evaluate the process, including critical discussions and rejections of the changed activity. The seventh action is to consolidate the changed activity into a new stable form of practice.

Research question in theoretical terms

This article explores two professional development projects and how a teacher in each project expanded their learning as a result. Using the CHAT framework, we investigate the interaction between the activity systems of the teachers in the classroom and the activity systems of the professional development projects. The two professional development projects share an objective, that is, to enhance teaching and, ultimately, the learning of students. Our generalised research question regarding teacher participation in large implementation projects thus becomes the following: how is the professional learning of mathematics teachers driven by contradictions and mediated by relationships in and between nodes of the different activity systems?

Methodology

We use CHAT as our theoretical framework (Eisenhart, 1991) to establish and analyse the two cases. This has been done before (e.g., in Stouraitis et al., 2015), but in this paper, the cases stretch over longer periods. It is therefore important not to use a static model. We want to focus on the trajectory of the activity systems over time, which makes the activity theory a particularly useful theoretical lens for characterizing activity. By focusing on the contradictions that drive learning in and shaping the outcomes of an activity system, we can learn about the evolution (or transformation) of a system.

The framework was not used as a tool to design the interventions themselves, nor was CHAT used in the implementation of the interventions. The CHAT framework was utilised during the collection of data regarding the implementations, the formulation of the research question, and the analysis. Opfer and Pedder (2011) argue that teacher learning must be examined using a complex theoretical framework, such as CHAT, in order to understand the network of variables that influence the process of learning new knowledge. By understanding the complex dynamics involved in teachers seeking and sharing knowledge within a professional developmental project, researchers and administrators can support teachers in learning how to navigate this complexity so as to develop knowledge that will improve their practice (Engeström & Sannino, 2010).

The two projects were evaluated quantitatively with respect to student achievement, but this did not address or explain what influenced implementation – that is, *how* and *why* the projects developed. To understand this, we require details. We must study the interface between project and classroom, as well as the overall context of the project – we require an in-depth case study. Our methodological design is closely connected to the CHAT framework, meaning that the data are not stripped of their context. Teachers' backgrounds are important to consider when analysing interviews, and the specific context of classroom observations is important to the data collection.

The two projects that provided the context for our cases were 'Information and Communications Technology (ICT) in the Innovative School' (the Demonstration School Project) and 'Quality in Danish Language and Mathematics Teaching' (KiDM). The projects were neither linear nor formative interventions (Engeström & Sannino, 2010) but, rather, a mix of the two. In a linear intervention, e.g., a Randomized Controlled Trial (RCT), the goals are known beforehand by the researcher, difficulties are often interpreted as indicating weaknesses in the design, and the aim is to control all variables and achieve a predefined solution. In contrast, a formative intervention begins with participants (subjects) recognising a problem, which is then dealt with using an intervention that is shaped and negotiated by subjects and a researcher. The intention is for subjects to gain agency and ownership of the process.

The cases explored in this article revolve around two teachers: Ea in the Demonstration School Project and Maya in KiDM. These two projects were selected because of certain similarities and differences: they are the largest national implementation projects in mathematics education in Denmark in the last 15 years. They both build upon the understanding that professional development must be practice oriented and teachers must be involved in reflection and meaning making. The KiDM project was designed based on the results from the Demonstrations School Project, and KiDM adds several new initiatives, such as questions to pose so as to help teachers during the implementation process. The two projects constitute two explanatory cases (Yin, 2014) explaining how and why the projects develop as they do. The teachers within the cases are selected on the basis of their eagerness to participate in the projects and being experienced and innovative teachers:

Ea, in the Demonstrations School Project, is a mathematics teacher with 14 years of experience. Ea questioned how to teach mathematics/statistics and was eager to discuss the teaching of mathematics. Ea was an early adopter (Sahin, 2006) of ICT in the mathematics classroom. She was appointed by the school management to support her colleagues in integrating ICT and has previously participated in other ICT teaching development projects.

Maya also has 14 years of experience and is the mathematics supervisor[1] at her school, a position she has also held at other schools in the municipality. Maya is the author of a book about assessment in mathematics.

Both Ea and Maya expressed, before their participation in the projects, that they were very interested in developing their teaching practice and therefore volunteered for the projects and even made themselves available for this research.

The cases of Ea in the Demonstration School Project and Maya in KiDM are *critical cases* and have importance for the general issue of professional development and implementation (Flyvbjerg, 2006). As critical cases, they are of the 'most likely' type (Flyvbjerg, 2006, p. 231) in the sense that, if Ea and Maya were unable to establish the intended mathematical classroom practices, less competent teachers would also be unable to establish them.

[1] Diplomas (60 ECTS) in mathematics supervision have been available since 2009, and many schools have a local supervisor in mathematics who initiates changes and supervises fellow teachers.

Data collection and methods

Ea participated in the Demonstration School Project for two years and implemented the intervention two times thereafter. The Demonstration School Project provided us with 31 video observations of Ea's classroom lessons: sixteen observations during the Demonstration School Project [Observation Ea intervention I 1–12; Observation Ea intervention II 13–16] and 15 observations from Ea's classroom between implementations and after the last implementation [Observation Ea teaching 1–15]. We also conducted four semi-structured interviews with Ea [Interview Ea 1–4], all of which were audio recorded. Interviews 1 and 2 were collected after the first implementation, and Interviews 3 and 4 were collected after the second implementation. The data further included the project website (in Danish) (Demonstrationsskoler.dk, 2015a; Demonstrationsskoler.dk, 2015b), a written description of the intervention design (Skott, 2014), and an interview with one of the project researchers. We did not have access to the raw interview data but, rather, referred to an analysis and interpretation of the interview (Skott, 2018).

Maya participated in KiDM, providing us with four video observations from her classroom [Observation Maya 1–4] and one semi-structured interview with Maya after the implementation [Interview Maya]. The data also included six evaluation sheets from teacher meetings at her school [Evaluation sheets Maya 1–6] and a transcript from a group evaluation meeting [Group evaluation Maya]. Additional follow-up data were obtained two years after the intervention [Correspondence Maya; Local framework Maya]. Further data included reports developed by the KiDM project (Hansen et al., 2020; Hansen & Bundsgaard, 2016). References to data sources will, in the analysis section below, be denoted by the use of the square brackets in this format: [Name, Date, Recording start in minutes].

All data have been transcribed, and the first part of our analysis consisted of coding, with the intention of closely representing data by performing both word-by-word coding and incident-to-incident coding (Charmaz, 2006). This first part of the analysis produced 27 categories (e.g., 'teamwork', 'students learning', 'inspiration', 'mathematical communication', and 'about the project'). These categories were then transformed into a theoretical coding by locating, elaborating, and discussing *nodes* (Figure 1) by moving between the data, CHAT theory, and interpretation. Data were used in the document analysis (Bowen, 2009), specifically the categories derived from the data were used to identify parts of the documents that addressed the same issues. Sorting the data into categories linked to nodes enabled us to tentatively point out contradictions when reviewing incidents in several of the data sources.

As a second part of our analysis, we used the expansive learning cycle to model the interventions. We identified the actions and the subjects participating in the actions. We then matched the contradictions to the expansive learning cycle model by fitting them to the actions, particularly Actions 1, 2, 5, and 6.

In the third and final part of the analysis, we modelled the activity systems by identifying subjects, describing mediating artefacts, and elaborating the object. The process is visualised by using the triangle models (Figures 5 and 7), which we use to explain how the contradictions emerge in or between the nodes of the involved activity systems.

Results – Ea in the Demonstration School Project

Analysis to identify actions

Ea participated in the Demonstration School Project, together with two other mathematics teachers from her school. The school had 900 students and was an association of three smaller schools in a rural area.

Action 1: In the first action, the researchers leading the overall project were the active participants. The researchers questioned the existing practice of students' learning, seeking to identify how teachers can use ICTs to initiate, negotiate, and establish two new statistical practices:

> *One where the students are to be critical towards the use of statistics and one where the students are to investigate tendencies and patterns in datasets. In that sense, students need to examine, understand, and critically evaluate data and statistical results that permeate their daily lives.* (Demonstrationsskoler.dk, 2015a)

Ea was also involved in the first action of questioning, criticising, or rejecting aspects of the existing practice. Ea did not meet the researchers, but she teaches statistics and questions her existing practice, as well as the aim of problem solving and statistical skills.

Action 2: In the second action, the researchers reviewed research on statistical reasoning and the use of digital resources (Skott & Østergaard, 2016, 2018). Research indicates that students best develop statistical knowledge through an explorative process in which they combine an overall data perspective with a modelling perspective to address statistical questions that seem authentic to the students (Ben-Zvi & Makar, 2016; Shaughnessy, 2007). Historical and empirical analysis revealed how most of the teaching of statistics takes place in schools: students are presented with statistical descriptors and procedures, which they subsequently practice using datasets that seldom have meaningful contexts (Cobb & McClain, 2004; Petrosino et al., 2003). Teaching statistics cannot be understood without the use of digital artefacts (Pepin et al., 2013), including opportunities to create dynamic visualisations of mathematical contexts (Laborde, 2001).

Action 3: In the third action, the researchers designed a sequence of lessons to demonstrate the desired new practices. This sequence of lessons was called 'Youngsters and ICTs' and aimed at the sixth grade. The objective was to encourage students to reason about data using digital support, adopting a critical stance towards statistics use, as well as to investigate tendencies and patterns in datasets (Demonstrationsskoler.dk, 2015a). The sequence, consisting of 15 lessons, was published on a website, with detailed descriptions of how to organise lessons, resources to use, and activity formats (Demonstrationsskoler.dk, 2015b). The sequence of lessons framed and prescribed how teachers should engage students in statistical investigations: formulate statistical problems; generate, analyse, and reason about data; interpret results; and disseminate them in and out of the school.

The proposed teaching practice had to integrate the use of spreadsheets and MiniTools to support students' data analysis and reasoning processes. A central tenet of the desired teaching practice was to include and capitalise on students' mathematical contributions in classroom discussions. 'Youngsters and ICTs' was, therefore, designed to engage students in a real-life context, using a newspaper article as a starting point. During the sequence, students were meant to work with the statistical content and, ultimately, communicate their new knowledge to people outside the classroom.

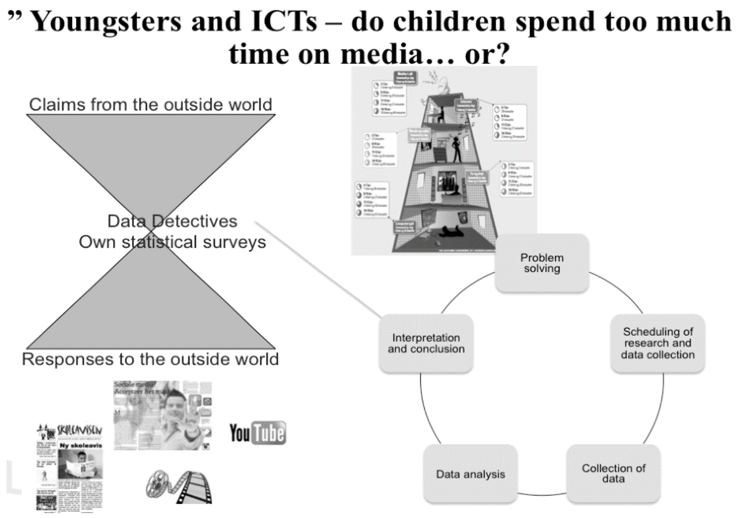


Figure 4 Illustration of elements involved in the desired new practice exemplified in the “Youngsters and ITCs” lesson sequence

Students could be described as ‘data detectives’ cycling through a process of problem formulation, planning, data collection, data analysis, interpretation, and conclusion (Figure 4). During the design of ‘Youngsters and ICTs’, participating teachers were regularly asked by mail to comment on the content and design. The teachers had time reserved to participate in these activities, but none of the participating teachers made any comments. This is likely because commenting on professional development courses is not a regular practice.

Action 4: In the fourth action, a demonstration lesson is carried out. The teachers were presented with the “Youngsters and ICTs” design and website, including the lesson plan for each lesson, questions to pose in support of students’ learning, and handout materials for the students.

Action 5: In the fifth action, the participating teachers implemented the sequence of lessons in their own classrooms. During the implementation, a researcher visited all schools, observed one lesson together with the teachers involved at the school, and discussed the content and student learning process.

To elaborate on action five, we provide an example of a discussion that took place in Ea’s classroom. Ea’s class had been working with the part of the lesson sequence that dealt with the question ‘Do children spend too much time with media?’ In the process, the students had developed a digital survey, collected data, and analysed the data using Excel and were now ready to present their interpretations and conclusions to their classmates. In the classroom episode cited below, a group of students presents their results, and we see how Ea facilitates the presentation session and the associated discussion [Observation Ea Implementation I 15, 28 March 2014, 34:00]:

Student Lærke: [The presentation begins] Exploring grade 6.

Student Michelle: What is . . . [the presentation reads, "What is the mean on weekends and weekdays?" There is a picture of three iPads. The student reads aloud.] Oh, is the mean on weekends and weekdays? There is a very big difference between the days we have studied. If we look at iPads, there is a huge difference. Saturday, it's 21 hours, Sunday, it's 19.5 hours, Monday, it's 28 hours, and Tuesday, it is 83.5 hours. And the mean is . . .

Student Lærke: [Four circle diagrams and several pictures of TV and iPads are represented. The presentation reads, "[Do] children spend too much time on media?"] Okay. Do we spend too much time on media? Oh, the mean of all days — we spend 9 hours and 16 minutes using medias. And, maybe, that's a bit too much per person per day. Oh, and on a weekday, Monday and Tuesday, on these days, every person spends 9 hours and 20 minutes. And the weekend, Saturday and Sunday, and it's the two days that have been put together, so it gives the two days is 9 hours and 11 minutes. And, all the days, it was so 9 hours and 16 minutes. [Applause.]

Ea: Okay. You [the students] can pose questions . . .

Student Peter: There were too many words. Therefore, I did not listen. Because it was just numbers, numbers, numbers, Monday, Tuesday. It became frustrating.

Ea: You did not understand the purpose of mentioning the numbers?

Student Peter: They kept saying Monday, Tuesday, and Wednesday. There was a maximum value, and then, there was a minimum value.

Student Sophie: There [were] a lot of pictures [on the Keynote, a presentation tool].

Ea: That is good. Yes.

Ea: Anyone? The presentation is really nice. It's a nice front page [on the KeyNote], etc. You should begin your presentation by saying what the actual goal was: [Do] young people spend too much time on media? Do their parents think they spend too much time on media? It does not appear until late in your presentation. Can you see it? Michelle repeats a lot of numbers that I do not associate with anything. Can you follow me? Yes. So, there is a very high mean per day, you say. I think the first thing you say is when you've added the numbers together. Then, you have taken the mean. Does it not appear clearly?

Student Lærke: No.

Ea: Somehow, there is no reason for all the talking if you do not explain anything with the numbers. You just tell me the numbers, but what are you telling with the numbers? Can you see what I mean? It's great that you've calculated all that. It's good. Nevertheless, there is no need to calculate if we do not need the calculations for anything. So, this is where you have to say: okay, we think it's really a lot of time people spend on media because we can see that the highest values, on the weekends, are 8 and 9 hours . . . But since the mean is only . . . we really think it's okay. Then, use the numbers to describe what you want to tell me.

Student Michelle: Yes.

In total, five groups present their calculations and interpretations regarding 'Do children spend too much time on media?'. The other students and Ea provide responses.
While Action 5 was clearly part of the expansive learning cycle, it did not, in itself, suggest what learning took place or why the lesson unfolded as it did. We, therefore, proceed to Action 6.

Action 6: In the sixth action, which involved reflecting on and evaluating the process, teachers and researchers participated in an after-implementation workshop, in which the discussion of the lesson sequence and students' learning opportunities took place. Ea was explicit about how she organised her mathematics lessons. She admitted that she did not like to orchestrate mathematical discussions. Traditionally, Ea provided short classroom directions, followed by individual or small group work with the textbook or iPad, partly because *'I fail to tell them anything in 20 minutes . . . I think I often supervise more than I teach . . . because I do not think* *they listen when I stand at the blackboard*' [Ea Interview 2, 27 October 2014, 10:00]. Thus, Ea rarely assembled the entire class for joint mathematical discussions.

In the above classroom example, Ea follows the recommendations built into the lesson sequence and attempts to establish a more dialogical teaching approach, in which the students, in groups, tell about and show their work to the rest of the class. After each presentation, the class applauds, and they discuss the presentation quality, i.e., the way the students communicated and represented their analyses, but not the quality of the statistical content.

Ea considered 'Youngsters and ICTs' to be insufficiently innovative: '*I do not think there has been enough ICT . . . but mathematically, it's another way of teaching than I have taught in the past*' [Interview Ea 2, 27 October 2014, 31:35]. Ea considered mathematical skills to be a prerequisite for problem-solving: '*I am a bit old-fashioned. I think it is most important they have skills . . . I can pose open problems, but if they do not have the skills, then they cannot solve them'* [Interview Ea 2, 27 October 2014, 59:20].

The researcher, who participated in the after-implementation workshop, was explicit about the challenges the teachers met in the intervention. In her view, the teachers experienced the lesson sequence as being significantly different from their usual teaching, partly because statistics, in many cases, is taught with a focus on skills: '*They [the teachers] are always thinking*: "*We spend too much time talking here, we need more skills . . . that's what the kids want*"' (Skott, 2018, p. 177). This is partly because the design of the intervention did not include the teachers: '*It's not because we did not want to include the teachers [in the design-phase], but the most important thing has been that is it based on research results*' (Skott, 2018, p. 179).

The researcher also elaborated regarding their view of the intervention: '*I have always had reservations about whether this [the intervention] could succeed. Do I believe in this?*' (Skott, 2018, p. 177). The researcher did not directly challenge the intervention's theoretical approach at any point but, rather, reiterated that it had been difficult to establish a culture of cooperation. The researcher would have liked the intervention to contain more practical and collaborative elements, stating that it was primarily through the practical elements of the intervention that teachers developed their teaching.

Action 7: After the intervention period concluded, Ea and a colleague decided to make 'Youngsters and ITCs' part of their sixth-grade curriculum: '*We want to develop the design . . . We can do better, right?*' [Interview Ea 4, 1 June 2015, 34:00]. Ea and her colleagues continued to work with ‘Youngsters and ITCs’ over the next 4 years, but this ended when the teachers became a part of another project, KiDM.

Analysis of contradictions in activity systems

To interpret the actions identified above in the expansive learning cycle, we now focus on contradictions in two involved activity systems: that of Ea and the students in her class and that of the researchers involved in the Demonstration School Project, particularly the ‘Youngsters and ITCs’ lesson sequence. In our analysis, we do not present all contradictions identifiable in or among the different nodes in a single activity system or among the nodes in the two activity systems. We focus on the four types of contradictions, pointing out specific nodes involved in the complex process of learning. While we highlight a limited number of nodes, it is important to remember that the entire network of activity system nodes is in the background of every contradiction.

The primary contradiction consists of a need for new statistics teaching practices. This contradiction is articulated in the activity system of the Demonstration School Project, meaning that there is a dissatisfaction with the current way of teaching and learning statistics (the object) in the activity system of the classroom in general. The primary contradiction, mainly located in the object node, becomes a driving force for learning in both activity systems because they share the same desired outcome of enhancing students’ statistical reasoning. The need for new statistics teaching practices is not restricted to a need for new statistics content and ways of

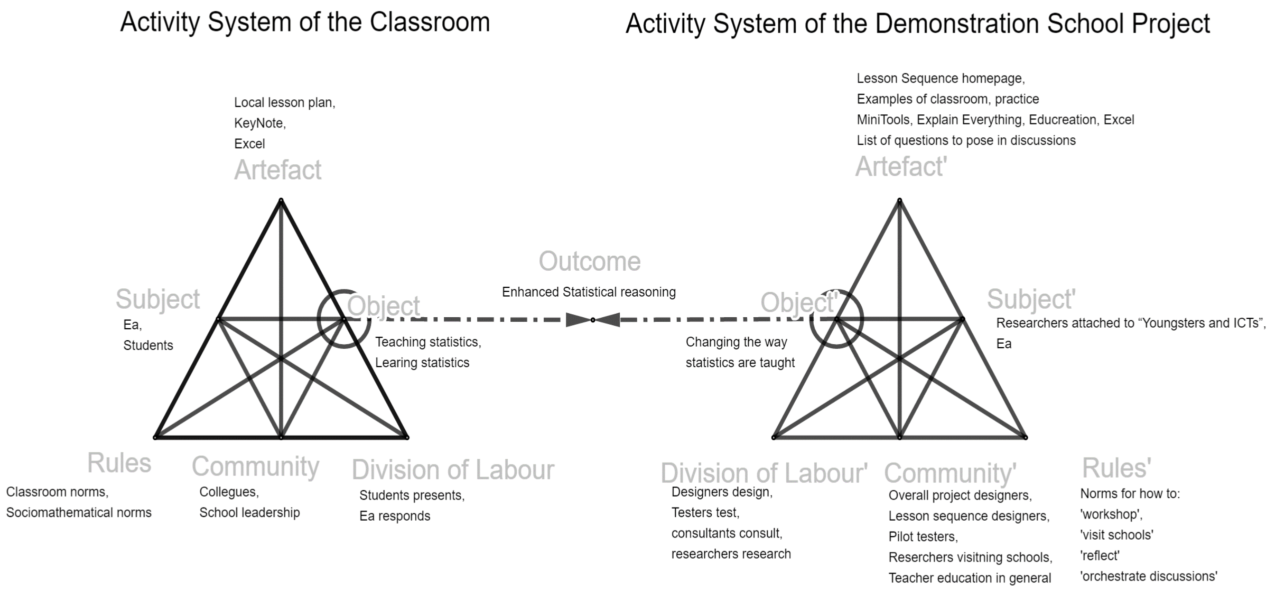


Figure 5 The activity systems in ‘Youngsters and ITCs’

teaching statistics but also includes a need to learn how to utilise new tools (artefacts) and how to make changes to teaching practice in classrooms.

Secondary contradictions in the classroom activity system support the primary contradiction: new artefacts, like dynamic data representation software, contradict the old ways of analysing and presenting data. Teaching must also change in order to benefit from new software (c.f. Action 2 above). Students' role as data detectives (new objective) contradicts the typical classroom division of labour, in which the teacher is responsible for posing questions about data (c.f. Action 3 above). Secondary contradictions are also apparent in the activity system of the project: if teachers are to engage in new ways of teaching statistics (object), designers must make this easy to do within the economic and practical constraints of the project. Therefore, a website (Demonstrationsskoler.dk, 2015b) with the lesson sequence (artefact) was introduced. Teachers were invited to comment, but none did, indicating a contradiction in the division of labour (designers design, teachers teach) (c.f. Action 3 above).

Tertiary contradictions are apparent in Actions 4 to 6, in which the new model for teaching statistics was settled upon; implemented at numerous participating schools; and, consequently, shaped so as to align with other existing activity. The process of settling on the final form of the lesson sequence for wider implementation (Action 4) was propelled by a need to dispel contradictions between the nodes of the two activity systems. New artefacts, such as the project lesson plan, contradicted local lesson plans, which were mostly very brief if they existed at all. The list of questions for teachers to pose to students was made to ameliorate the contradiction between the whole-class discussions envisioned in the project (*rules*) and the prevailing verbal interactions in the classroom, which usually followed a pattern of "question-response-evaluation" (*rules*). This discussion-contradiction continued into Action 5, where it gave rise to resistance and evaluation, which were unfavourable regarding the project.

Ea later stated in an interview (c.f. Action 6) that she did not feel comfortable in her verbal communication about mathematics and rarely attempted whole-class discussions. She saw her role as teacher as more of a supervisor for each individual student. This illustrates a contradiction between Ea's wish to follow the intention of the lesson sequence to orchestrate whole-class discussions (*subject* and *rules*) and the existing division of labour in the classroom. This contradiction was touched upon in several other interviews with Ea: '*There are more joint discussions [in the intervention], and I think it is good for some of the students*'; '*I think there was too much talking'* (Interview Ea 2). Ea attempted to change the classroom practice, but what was intended as an opportunity for discussion devolved into comments mostly regarding presentation form. Little was said about how to communicate statistics and much less about reasoning with statistics. Consequently, Ea was not fully convinced that she should change the usual division of labour and rules of the classroom activity system.

The aim of integrating ICT into 'Youngsters and ITCs' was to enable students to use technology to reason about data, be critical about the use of statistics, and investigate tendencies and patterns in datasets. Looking at the presentations recounted in Action 5, we do not see the use of such investigative competences. Instead, we see how the students used KeyNote and other non-dynamic data charts (all done in Excel and not with the recommended MiniTools) to present their results to Ea and their classmates. In short, we see a contradiction between the ways ICT artefacts were used to investigate data in the project and communicate data in the classroom. This also reveals a quaternary contradiction in how Ea was regarded by colleagues and leadership as an ICT-competent teacher: for Ea to learn a new didactic use of ICT, she had to abandon this sense of regard, which was tied to her self-image, temporarily. Ea had a difficult time turning this contradiction into a motive for learning.

After the intervention, Ea and her teacher colleagues instituted 'Youngsters and ITC' as a stable part of teaching statistics to the sixth grade and began to consolidate some of it into a new practice. In a way, the Demonstration School Project changed the activity of Ea's classroom, not as intended but, rather, by realigning it with other existing activities. Ea felt agency and wanted to develop and improve the design. Our data regarding Ea's participation in 'Youngsters and ITCs' did not reveal quaternary contradictions. Elements from 'Youngsters and ICT' were used for several years, until they were questioned anew when the school activity system encountered the KiDM activity system.

Results – Maya in the KiDM project

Analysis to identify actions

Maya was involved in developing the mathematical culture at her small village school (300 students), and among other things, she appreciated her work in organising group meetings with her fellow math teachers. At these group meetings, Maya was responsible for the agenda, which could, for example, focus on pedagogical aspects of mathematics or the arrangement of a math day (a whole day with focus only on mathematics).

In the following, we analyse the KiDM project to identify five actions in the expansive learning cycle Maya took part in during the KiDM project.

Action 1: In the first action, questions were raised by the Danish Ministry of Education, in collaboration with the Danish School Headmaster Association and the Danish Teachers Association. The question was how to increase teaching quality in school mathematics and Danish language classes. A consortium of two universities and four university colleges was formed to design an intervention project that would test potential answers to the question. The overall project parameters were written by a group of researchers dubbed 'the KiDM steering group'. Notably, they handled the 'internal' part of the question by deciding that increasing the quality of school mathematics would be equivalent to making teaching more inquiry based. The project ran for three years and was structured as a randomised controlled trial. It ultimately involved 87 schools with 267 classes and 4,681 students. Maya was indirectly involved in the first action because she was working in school practice, meaning that she is specifically challenged to make her teaching more open and inquiry based.

Action 2: The second action, historical and actual empirical analysis, was conducted by the KiDM steering group. First, a systematic literature review was carried out regarding inquiry-based mathematics. Then, a qualitative interview study was undertaken at the schools (Dreyøe et al., 2017; Michelsen et al., 2017). Through these interviews, the teachers were directly involved in this action, and although Maya was not part of the interviews, she is indirectly represented through these interviews because of her position as a teacher. The analysis concluded with the formulation of three grand principles of inquiry-based teaching to be implemented in the KiDM intervention.

Action 3: The third action was to design a suitable KiDM model for inquiry-based mathematics. It was developed in a design-based research collaboration process involving ten primary and lower secondary teachers, as well as mathematics supervisors, associate professors from teacher education, and university professors. This action involved an iterative process in which the intervention was tested a number of times (Hansen et al., 2020). In total, eight schools were involved in the process of modelling the new proposed mode of mathematics teaching activity.

As in the Demonstration School Project, the intervention was built around a website containing a detailed teachers' guide and accompanying pages for students, detailing various inquiry-based activities. In addition to the descriptions of tasks, the teachers' guide contained direct guidelines regarding how the activities should be introduced, what hints could be given during lessons, and how the collective class discussions should be designed.

Groups of teachers and supervisors at the participating schools began the intervention by taking part in an all-day meeting that introduced both the inquiry teaching concepts and the practical organisation of the project. While teachers tried out the inquiry teaching activities in their own classrooms, they were to meet and discuss these activities with their local colleagues a total of six times for 1–2 hours.

Action 4: The intervention was tested after the development phase in a pilot run at 14 schools (56 classes). It was adjusted slightly before the full trial started. The pilot schoolteachers commented on the interventions, both at evaluation meetings and through questionnaires.

Action 5: One KiDM teaching activity was called 'How many knots?' and inspired by Burke et al. (2008). In this activity, the students produced their own statistical datasets by tying knots on a string. The students began individually, tying as many knots as they could in one minute. Afterward, every student counted how many knots each had made. Then, in groups of three to four students, they prepared charts in a form of their own choice to provide an overview of the numbers. The following question was then posed: '*If a man suddenly comes through the door and asks how many knots a student in this class can tie in one minute, what can you tell him?*'

One of the aims, as written on the website teachers' guide, was for students to '*collect, arrange and describe numerical data*', '*explain informal experiences about frequency and average*', '*argue for their choice of systematisation*', and '*use reasoning in the inquiry work*' (KiDM.dk, 2017, our translation).

All groups of students were supposed to plan a presentation of their systematisation and their answers to 'the man'. The website teacher guide suggested that each group should develop a short video about their findings and associated reasoning. It contained an example video filmed at one of the development schools (c.f. Action 3). In addition, the guide suggested questions for teachers to pose during whole-class discussions in order to focus specifically on the students' argumentation for their choices and findings.

To explore the full trial implementation in more detail, we provide an example of a discussion that took place in Maya's classroom, specifically a whole-class discussion after group work. Maya has asked the students to make a poster, not a video as recommended, presenting their findings. In the whole-class discussion, the groups took turns presenting their posters in front of the class (Figure 6)

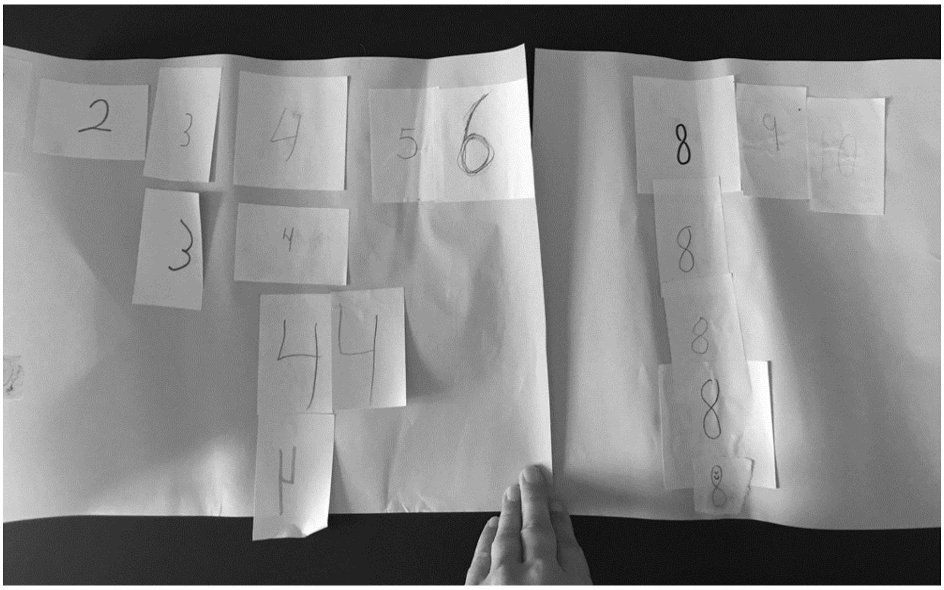


Figure 6 Student poster with Post-it notes representing the number of knots each student tied (frequency grouping)

[Observation Maya 4, 1 May 2018, 0:21 and 11:24].

[Four students are at the blackboard, ready to present their poster]

Maya: Okay, please explain.

Student Peter: Okay, the smallest number is 1.

Student Michael: And the biggest number is 10.

Student Molly: And there are just as many who have tied 8 and 4 knots.

Student Olga: And, together, there are 99 knots altogether.

Maya: Very nice. Give the group a hand!

> *[Applause. The next group goes to the board. The remainder of the groups continue presenting in a similar way. During the last group presentation, the following takes place.]*
>
> *Maya: I thought, in your group, there were two different kinds of mean — the one where you were 18 students and the one where you were 14 students. What did you learn from*

it? Did you have some extra thoughts about average? Why is one number lower than the other?

Maya: I would like to follow your reasoning. How could we get two numbers — shouldn't it be the same?

Student Roberta: No.

Maya: Why not?

Student Roberta: Because [stops] . . . I do not know.

Maya: No, but then there may be some others in the class who can figure it out. Anyone? Why is it like that? Can you explain it?

Student Dagmar: I do not know . . .

Maya: Okay, what do you say Noah?

Student Noah: For example — those knots we've counted . . . If it's like, that the new students . . . probably had a little lower number than the others . . . then we can deduce that it [the mean] will get lower. If each new student had made eight knots, then it might have been the same.

Maya: Maybe . . .

Student Noah: Maybe?

Maya: Does everybody understand what Noah says?

Several Students: Yes.

Maya: That's because the four students who came in late, they tied fewer knots, so the mean was a bit lower when they were included. That is a good observation and a good consideration . . . If the new students had [tied] 10 each, then it [the mean] would have been higher . . . so that's why . . .

Student Vilfred: There is something else, Maya! It's also because you divide it [the total] with more and then it [the mean] will be a smaller number.

Maya: Good enough — interesting . . . [the lesson ends].

Action 6: Maya participated in the intervention as a teacher of a fifth-grade class and in her capacity as the school's mathematics supervisor. Maya felt that the many specific questions provided for student reflection were pertinent and rewarding. Maya also emphasised the use of the specific suggested materials as brilliant in several of the provided tasks, e.g., she lauded the use of Post-it notes during the 'How many knots' teaching activity. However, she also expressed concerns about how well the KiDM ideas about inquiry-oriented teaching could be used with students in situations involving different mathematical content.

Action 7: At Maya's school, they intended to implement elements of the KiDM project in the coming year [Group Evaluation Maya, 20 June 2018, 33:44]:

> *We have to put it in as part of our teacher development next year. We will look at innovative exercises, and each class must do one or two explorative activities each year... Yes, so we all must do two inquiry activities, and we must present it to each other and try to develop it together. So, we must spend time on it. And it may well be there is somebody who chooses to grab some of this [KiDM material].*

In a follow-up correspondence two years later [Correspondence Maya, 22 June 2020], Maya reported that KiDM material was still available at the school and remembered vaguely that the material had been used on several occasions. More notably, KiDM background material explaining the pedagogics and didactics behind inquiry-based teaching had been discussed at meetings among the mathematics teachers. It had inspired the teachers to develop a common framework for teaching that involved mathematical experiments [Local framework Maya, December 2019].

Analysis of contradictions in activity systems

Maya took part in two systems of activity: the classroom and the KiDM project. The object of the first was the teaching and learning of students in the classroom; the second sought to change how statistics is taught (Figure 7).

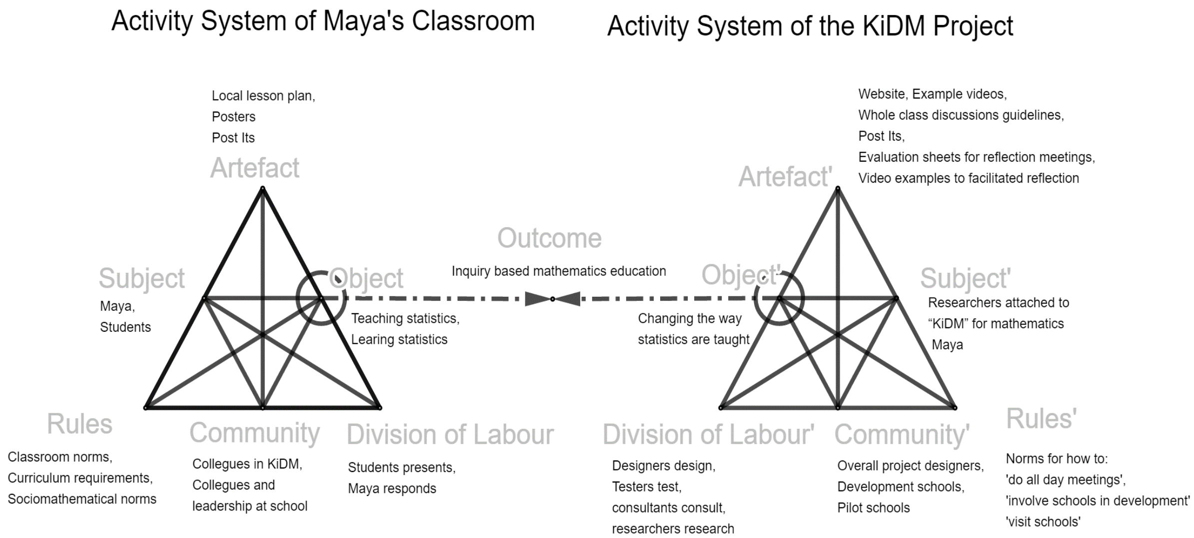


Figure 7 Activity systems in KiDM

The primary contradiction is expressed in the overall project description written by the KiDM steering group (Action 1). They were motivated by a contradiction between what some research sees as quality in teaching, i.e., having an inquiry-based approach (Artigue & Blomhøj, 2013), and how teaching generally takes place, as reported in a baseline investigation incidentally prepared within the Demonstration School Project (Hansen, 2019, p. 45). The baseline showed that most statistics lessons in Danish schools are used for training skills. They do not involve problem solving or open-ended tasks. By participating in KiDM, Maya and colleagues accepted the primary contradiction as a form of motivation to attempt a new method of teaching statistics.

The secondary contradiction is apparent in the historical and empirical analysis of Action 2. There is a contradiction between the existing division of labour and rules regarding the responsibilities in the classroom and the new object of inquiry-based teaching: students must assume ownership of the inquiry process, and the teacher must relinquish control. Inquiry teaching (*object*) reveals a contradiction with the curriculum (*rules*): will students get around to the curriculum requirements? A contradiction is also apparent between the old and new objects: will students realise the common mathematical structures and abstract mathematical concepts, or will their learning only adhere to the specific context of the inquiry (Hansen & Bundsgaard, 2016)? Such contradictions caused the formulation of the design principles, reflected in Action 3, in which a great deal of the KiDM resources and project activities were allocated to a collaborative design process with teachers from three different schools.

Several tertiary and commonly known contradictions led the KiDM project to settle on a regime with many artefacts intended to facilitate high implementation fidelity. One such contradiction, in the "bottom" nodes, stems from the classic double bind in which teachers are obliged to prepare well for lessons but, for economic reasons, rarely have adequate time to do so. Therefore, teacher meetings tend to become bogged down in day-to-day business, leaving little time for reflection and development discussions.

Looking at the discussion that took place in Maya's classroom (Action 5), we notice that Maya modified elements in the 'How many knots' lesson. Short movies were not prepared by the students, revealing a tertiary contradiction between the artefacts introduced by KiDM and the artefacts usually employed in the class: '*In the plan, it's written "can" and not "must"*' [Maya interview, 20 June 2018, 23:55]. While this change in mediating artefacts may superficially seem harmless, it is detrimental to the KiDM objective of having reasoning in whole-class discussions reflect students' systematisation and development of understanding. The intention of the KiDM project was for students to express their explanations and reasoning in the problem-solving process through video presentations. In Maya's classroom, students made many calculations to obtain their results, but this was not reflected in their presentations, meaning the students did not get to explain or argue for their calculations. Consequently, they did not have the opportunity to convince their fellow students of the validity of their claims. There appears to be a tertiary contradiction between the communicative norms in the classroom and those encapsulated in the mediating artefacts of the KiDM project. The KiDM material set the stage for explanations and follow-up questions, but this was not the norm in Maya's classroom. At the end of the presentations, Maya did ask a follow-up question regarding one aspect of the presentations: she asked why the average changed when more students participated in the dataset. Noah attempted to explain his reasoning process, but this did not lead to a discussion that other students could join. Rather, the discussion quickly closed again, with arguments never being fully verified or rejected. Apparently, there is a contradiction in the existing division of labour, in which Maya is supposed to perform the valuation, and the objective of inquiry-based mathematics, in which students take on this obligation.

In evaluating her participation in KiDM, Maya revealed a tertiary contradiction between her and the school community, which mirrors the case of Ea. Maya was an experienced mathematics supervisor and thought highly of her own teaching competences: '*I participated in some projects and gave some suggestions to the changes in mathematics teaching, and then, I was headhunted. I was asked if I wanted to start as a mathematics consultant in the municipality*' [Maya Interview, 20 June 2018, 04:25, 04:40 and 03:03]. In addition, she expressed that she was participating in professional development because '*They were impressed with my work. I was a first-mover...*' and '*It wasn't new [knowledge], but I think it is interesting to grow*.' Maya's interpretation of these statements gave her confidence to modify the elements prescribed by the KiDM project, although it made her miss key features of inquiry-based teaching (the new object). She readily used Post-it, but while she praised the questions provided for reflection, she did not actively use them in class (Actions 5 and 6).

While Maya and colleagues consolidated their inquiry-based teaching into a newly changed practice (Action 7), we did not find any data suggesting manifestations of quaternary contradictions.

Discussion and concluding remarks

Through the analysis of the two cases, we have now clarified how the professional learning of mathematics teachers was driven by contradictions and mediated by relationships in and between nodes of the different activity systems. Notably, in both cases, the primary contradiction was mainly articulated in one activity system and only experienced in practice in the other activity system. The interventions, therefore, did not respond directly to a need felt by the subjects of the classroom activity systems. This fact may be meaningful to decision-makers within the school community, but it does not necessarily affect teachers, who are ultimately the gatekeepers of changes in classroom practice. While we see a less linear intervention design in KiDM than in the Demonstration School Project, we are still far from letting researchers and teachers question school teaching practice together. The question is how this can be possible given that researchers and teachers occupy different positions in developmental processes and research.

The outcome of expanded knowledge is not clearly envisaged by subjects in each activity system, which should come as no surprise given the 'learning as expansion' metaphor, which, by definition, deals with learning in situations in which the outcome of learning is not known *a priori*. It is possible to view the behaviour of Ea and Maya as reasonable ways to resolve or reduce the contradictions that surfaced in their teaching. By making incremental changes to their practice, they contributed to an expanding learning circle showing how teaching should be conducted, integrating innovative ICT use or inquiry. Ea and Maya were both self-assured, which made it personally and socially challenging to turn contradictions into motives for learning. They showed scepticism, a tertiary contradiction revealed by signs of resistance. Assuming that a critical stance is a virtue, especially in the classroom activity system community node, this impedes close adherence to the interventions. Critical reflection is also valued in the project activity system community for cultural-historical reasons. Danish teacher education, to which many of the project researchers are attached, has a long tradition of valuing teachers who are able to form their own informed opinions independently. This is consistent with Engeström's (2011) focus on agency as important in developing a viable interventionist methodology. This study, however, shows that agency is a double-edged sword for the makers of educational interventions: agency is required for teachers to engage in the project, but agency is also what enables teachers to interpret the project according to contradictions in 'their own' activity system, which are not necessarily the same as those in the project activity system. Consequently, the establishment of agency without paying attention to contradictions entails expansive learning in directions not wholly aligned with the project activity system.

One surprising fact was that no quaternary contradictions were obvious in our data. If we assume that Ea and Maya did indeed substantially change their practices, we would expect evidence that they had to justify their new practice to parents, colleagues outside the projects, or others. It may be that quaternary contradictions were not apparent, because Ea and Maya's changed classroom practices fit well with the classroom activity system, but in light of all the other contradictions, we suspect that the elements of the project activity systems were assimilated, to use a Piagetian notion, into existing practices. This assimilation was so seamless that neighbouring activity systems were not affected. If we take the bleaker outlook that Ea and Maya changed very little of their practice, this would also explain the absence of quaternary contradictions, something that is supported by their own 'confession' of having seen little that was new in the intervention. Seeing Ea and Maya as critical cases, it is difficult to imagine less experienced teachers change their classroom practices in accordance with the projects' intended aims.

Finally, through our analysis of the interactions between the activity systems of the teachers in the classroom and two professional development projects, we have explored different emerging contradictions that were mediated by different nodes in the CHAT model.

To make a contradiction a driving force in these intervention projects, such contradictions must be the key point in the negotiations between the activity systems. Teachers and researchers must make the contradictions apparent to everyone in the activity systems. We specifically point to the absence of quaternary contradictions as problematic for the intentions of these kinds of large-scale interventions. We, therefore, propose an even greater focus on how the various contradictions might be clarified during a professional development process and, in particular, the quaternary contradictions.

In our study, we have used the expansive learning cycle to identify and analyse the progression of the projects through actions. The model helps to create a bridge between the individual classrooms and the more general aims of the projects. The analysis maps a dialectic between large-scale intervention studies and more locally determined solutions that can lead to change. A limitation or concern regarding the use of the CHAT framework, however, is that the map can become overly comprehensive. This can be a strength because the framework matches the complexity of the projects like few other frameworks do, or it can a weakness because it does not readily point to any single problematic issue. The Demonstration School Project and KiDM had ambitions to be both largescale intervention projects with control variables and also interventions aimed at involving the participants in the development process. We question whether it is possible to satisfy both desires at the same time.

To clarify this study's contributions, we will attempt to categorize our results based on the five reasons to perform implementation research described by Century and Cassata (2016), which were recently used in Koichu, Aguilar and Misfeldt (2021). These reasons are as follows: *'(i) inform innovation design and development; (ii) understand whether and to what extent the innovation achieves desired outcomes for the target population; (iii) understand relationships between influential factors, innovation enactment, and outcomes; (iv) improve innovation design, use, and support in practice settings; (v) develop theory*' (Koichu et al., 2021, p. 977). The findings in this study can be assigned to three of these categories (i, ii, and iii) and therefore contribute broadly to the field of research. Firstly, the result shows that, when designing interventions, it is important to focus on the agency conferred by all the contradictions, including the quaternary contradiction. This is likely a political or project-management challenge, but we find that this study helps by providing important information that can be used when planning future projects (reason i). The analysis in this study also attempts to determine whether the interventions produce the expected results (reason ii), and points out that, in terms of involving teachers, agency is complex because critical teachers do not always follow an intervention plan. This study helps to show that we must be aware of this duality in future projects. Finally, the results also provide us with a better understanding of the relationship between the involved activity systems and the fact that many nodes in activity systems shape the outcome (reason iii) of attempts to change practice.

References

Artigue, M., & Blomhøj, M. (2013). Conceptualizing inquiry-based education in mathematics. *ZDM - International Journal on Mathematics Education*, *45*(6), 797–810.

Ben-Zvi, D., & Makar, K. (2016). International perspectives on the teaching and learning of statistics. In D. Ben-Zvi, & K. Makar (Eds.), *The teaching and learning of statistics* (pp. 1–10). Springer International Publishing.

Bosch, M., & Gascón, J. (2014). Introduction to the Anthropological Theory of the Didactic (ATD). In A. Bikner-Ahsbahs & S. Prediger (Eds.), *Networking of Theories as a Research Practice in Mathematics Education* (pp. 67–83). Springer International Publishing.

Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, *9*(2), 27–40.

Burke, M. J., Luebeck, J., Martin, T. S., McCrone, S. M., Piccolino, A. V., & Riley, K. J. (2008). *Navigating through reasoning and proof in Grades 9-12*: National Council of Teachers of Mathematics.

Cai, J., Morris, A., Hohensee, C., Hwang, S., Robison, V., & Hiebert, J. (2017). Making classroom implementation an integral part of research. *Journal for Research in Mathematics Education*, *48*(4), 342–347.

Century, J., & Cassata, A. (2016). Implementation research: Finding common ground on what, how, why, where, and who. *Review of Research in Education*, *40*(1), 169–215.

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. SAGE Publications

Chen, J. C., Lin, F. L., & Yang, K. L. (2018). A novice mathematics teacher educator–researcher's evolution of tools designed for in-service mathematics teachers' professional development. *Journal of Mathematics Teacher Education*, *21*(5), 517–539.

Cobb, P. (2007). Putting philosophy to work—Coping with multiple perspectives. In F. K.

Lester (Ed.), *Second handbook on research of mathematics teaching and learning* (pp. 3–38). NCTM, Information Age Publishing Inc.

Cobb, P., & McClain, K. (2004). Principles of instructional design for supporting the development of students' statistical reasoning. In J. Garfield & D. Ben-Zvi (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 375–395). Springer.

Cole, M., & Engeström, Y. (1993). A cultural-historical approach to distributed cognition. In G. Salomon (Ed.), *Distributed cognitions: Psychological and educational considerations* (pp. 1–46). Cambridge University Press.

Demonstrationsskoler.dk. (2015a). *Unge og medier - statistik*. Retrieved December 13, 2021, from http://auuc.demonstrationsskoler.dk/materialer/innovation/forloeb3/om

Demonstrationsskoler.dk. (2015b). *Unge og medier - statistik*. Retrieved December 13, 2021, from http://auuc.demonstrationsskoler.dk/materialer/innovation/forloeb3/faser/hvordan-bruges-statistik-i-hverdagen

Dreyøe, J., Michelsen, C., Hjelmborg, M. D., Larsen, D. M., Lindhart, B. K., & Misfeldt, M. (2017). *Hvad vi ved om undersøgelsesorienteret undervisning i matematik - Forundersøgelse i projekt Kvalitet i Dansk og Matematik - Delrapport 2*. Undervisningsministeriet (Danish Ministry of Education).

Eisenhart, M. (1991). Conceptual frameworks for research circa 1991: Ideas from a cultural anthropologist; implications for mathematics education rese. In R. G. Underhill (Ed.), *Proceedings of the thirteenth annual meeting - North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 202–219).

Engeström, Y (1999). Activity theory and individual and social transformation, In: Y. Engeström, R. Miettinen, and R.–L. Punamäki (eds). *Perspectives on activity theory*. Cambridge University Press, (pp. 19–38).

Engeström, Y. (2001). Expansive learning at work: Toward an activity theoretical reconceptualization. *Journal of Education and Work*, *14*(1), 133–156.

Engeström, Y. (2008) Crossing boundaries in teacher teams. In Y Engeström (Ed.) *From teams to knots: Activity-Theoretical studies of collaboration and learning at work.* (pp. 86-107). Cambridge University Press

Engeström, Y, Miettinen, R & Punamaki, R. L. (1999). *Perspectives on Activity*

Theory.Cambridge University Press.

Engeström, Y., & Sannino, A. (2010). Studies of expansive learning: Foundations, findings and future challenges. *Educational Research Review*, *5*(1), 1–24.

Engeström, Y. (2011) From design experiments to formative interventions. *Theory and Psychology*, *21*(5), 598–628.

Flyvbjerg, B. (2006). Five misunderstandings about case-study research. *Qualitative Inquiry*, *12*(2), 219–245.

Foot, K. (2014). Cultural-historical activity theory: Exploring a theory to inform practice and research. *Journal of Human Behaviour in the Social Environment*, *24*(3), 329–347.

Foot, K., & Groleau, C. (2011). Contradictions, transitions, and materiality in organizing processes: An activity theory perspective. *First Monday*, *16*(6),1–18.

Hansen, T., Elf, N., Misfeldt, M., Gissel, S., & Lindhart, B. (2020). *Kvalitet i Dansk og Matematik*. Retrieved December 13, 2021, from https://static-curis.ku.dk/portal/files/236902924/Slutrapport_Kvalitet_i_dansk_og_matematik.pdf

Hansen, T. I., & Bundsgaard, J. (2016). *Effektmåling af demonstrationsskoleforsøg. Afrapportering af kvantitative undersøgelser på tværs af de tre demonstrationsskoleprojekter i AUUC-konsortiet*. Retrieved December 13, 2021, from https://www.forskningsdatabasen.dk/en/catalog/2290382366

Jankvist, U., Aguilar, M. S., Dreyøe, J., & Misfeldt, M. (2019, February). Adapting implementation research frameworks for mathematics education. *In Eleventh Congress of the European Society for Research in Mathematics Education* (No. 08). Freudenthal Group; Freudenthal Institute; ERME.

Kaptelinin, V., Kuutti, K., & Bannon, L. (1995). Activity theory: Basic concepts and applications. In B. Blumenthal, J. Gornostaev, & C. Unger (Eds.), *Human-Computer Interactions* (pp. 189–201). Springer.

KiDM.dk. (2017). *Aims of 3rd intervention*. Retrieved December 13, 2021, from http://kidm.dk/matematik/laerer/matematik-laerer/oversigt/de-tre-indsatser/indsats-3/binde-knuder/maalsaetning/

Koichu, B., Aguilar, M. S., & Misfeldt, M. (2021). Implementation-related research in mathematics education: the search for identity. *ZDM–Mathematics Education*, *53*(5), 975-989.

Laborde, C. (2001). The use of new technologies as a vehicle for restructuring teachers' mathematics. In L. Fou-Lin & T. J. Cooney (Eds.), *Making sense of mathematics teacher education* (pp. 87–109). Springer.

Leontiev, A. (1978). *Activity, personality, and consciousness*. Prentice-Hall.

Maaß, K., & Artigue, M. (2013). Implementation of inquiry-based learning in day-to-day teaching: A synthesis. ZDM - Mathematics Education, 45(6), 779–795. https:// doi. org/ 10. 1007/ s11858-013-0528-0

Michelsen, C., Dreyøe, J., Hjelmborg, M., & Larsen, D. (2017). *Forskningsbaseret viden om undersøgende matematikundervisning*. Retrieved December 13, 2021, from https://www.forskningsdatabasen.dk/en/catalog/2393691204

Opfer, V. D., & Pedder, D. (2011). Conceptualizing teacher professional learning. *Review of Educational Research*, *81*(3), 376–407.

Pepin, B., Gueudet, G., & Trouche, L. (2013). Re-sourcing teachers' work and interactions: A collective perspective on resources, their use and transformation. *ZDM - International Journal on Mathematics Education*, *45*(7), 929–943.

Petrosino, A. J., Lehrer, R., & Schauble, L. (2003). Structuring error and experimental variation as distribution in the fourth grade. *Mathematical Thinking and Learning*, *5*(2–3), 131–156.

Sannino, A. (2008). Sustaining a non-dominant activity in school: Only a utopia? *Journal of Educational Change 9*(4), 329–338.

Sahin, I. (2006). Detailed review of Rogers' diffusion of innovations theory and educational technology-related studies based on Rogers' theory. *The Turkish Online Journal of Educational Technology, 5*(2), 14–23.

Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational Researcher*, *27*(2), 4–13.

Shaughnessy, J. M. (2007). Research on statistics' reasoning and learning. In K. Frank & Jr. Lester, (Eds.), *Second handbook on research of mathematics teaching and learning* (pp. 957–1010)., NC: Information Age Publishing.

Skott, C. K. (2014). *Beskrivelse af interventionsdesign for "It i den innovative skole – nye kompetencer, nye organiseringsformer i det 21. århundrede"*. Retrieved December 13, 2021, from http://auuc.demonstrationsskoler.dk/sites/default/files//u103/beskrivelse_af_interventionsdesign_itis_040714.pdf

Skott, C. K. (2018). Konsulenters fortællinger om en intervention. In J. Bundsgaard, M. Georgsen, S. Graf, T. I. Hansen, & C. K. Skott (Eds.), *Skoleudvikling med it. Forskning i tre demonstrationsskoleforsøg I* (pp. 167–185). Aarhus Universitetsforlag.

Skott, C. K., & Østergaard, C. H. (2016). How does an ICT-competent mathematics teacher benefit from an ICT-integrative project? In C. Csíkos, A. Rausch, & J. Szitányi (Eds.), *Proceedings of the 40th conference of the international group for the psychology of mathematics education* (pp. 211–218).

Skott, C. K., & Østergaard, C. H. (2018). En matematiklærers professionelle udvikling. In *Innovativ undervisning med it. Forskning i tre demonstrationsskoleforsøg II* (pp. 157–178), Aarhus Universitetsforlag.

Skott, J. (2009). Theoretical loops in mathematics education: Theory-practice relationships, revisited. *Asian Journal of Educational Research and Synergy 1(2), 74–88.*

Skott, J. (2020). Lakatos, Columbus and the expectations of research impact on practice. *Journal of Mathematics Teacher Education*, *23*, 333–337.

Stouraitis, K., Potari, D., & Skott, J. (2015). Contradictions and shifts in teaching with a new curriculum: The role of mathematics. In *CERME 9-Ninth Congress of the European Society for Research in Mathematics Education* (pp. 3262-3268).

Terhart, E. (2013). Teacher resistance against school reform: Reflecting an inconvenient truth. *School Leadership and Management*, *33*(5), 486–500.

Vygotskij, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Williams, J., & Wake, G. (2007). Black boxes in workplace mathematics. *Educational Studies in Mathematics*, *64*(3), 317–343.

Yin, R. K. (2014). *Case Study Research* (5 ed.). Thousand Oaks: SAGE Publications.

Zazkis, R., & Chernoff, E. J. (2008). What makes a counterexample exemplary? *Educational Studies in Mathematics*, *68*(3), 195–208.

8.2 Paper II: Sustainable Lesson Study and Paradidactic Infrastructure: The case of Denmark

SUSTAINABLE LESSON STUDY AND PARADIDACTIC INFRASTRUCTURE: THE CASE OF DENMARK

Camilla Hellsten Østergaard[1] and Carl Winsløw[2]

University of Copenhagen

[1] choe@kp.dk and [2] winslow@ind.ku.dk

Abstract

Most widely published books and papers on lesson study emphasise that a main attraction of this format is that it enables teachers to build and maintain a shared knowledge base for the teaching profession. This, however, clearly requires that lesson study is established as more than an occasional experiment, which is new to most teachers involved and involves only exceptional teachers. Therefore, research on lesson study increasingly focuses on the question of sustainability of the format as a practice-based form of professional development. We present and expand the recent theoretical idea of paradidactic infrastructure and analyse sustainability from the institutional perspective of the Anthropological Theory of the Didactic and especially based on the idea of paradidactic infrastructure. Our data come from the Danish context where lesson study in mathematics has been experimented with for more than 10 years, first by individual researchers and now, increasingly driven by institutions, such as schools or municipalities.

Keywords: Lesson study, sustainability, paradidactic infrastructure, professional development

Introduction

Lesson study (LS) was essentially introduced to the Western world by a rare “best-seller” in the educational science literature authored by Stigler and Hiebert (1999). In both the USA and Denmark, the dissemination of LS has to some extent, been based on individuals’ first-hand experience of it in Japan and in video recordings that relay some of what LS may contain and produce in terms of mathematics teaching. Slowly, several Danish scholars began to take interest in LS as described by international (especially American) scholars, such as Stigler and Hiebert (1999).

The purpose of this paper is to investigate the sustainability of LS in the Danish context, drawing on theoretical tools that have so far only been used in the context of Japan. Danish efforts to implement LS began relatively late compared to those of other Western countries. Such efforts first happened in the United States, and Lewis (2002) already reported on a number of "barriers facing LS" in this particular country and noted that "research is needed to understand the supporting conditions that have enabled lesson study to succeed at some sites" (p. 19). Newer empirical research has studied the effects of efforts to bring about such conditions, such as providing additional resources like prepackaged LS kits with templates, LS steps and mathematics resources (Akiba & Wilkinson, 2016; Lewis & Perry, 2014), involving leaders of schools or bigger districts in adapting LS in organisational structures and routines (Akiba, Howard & Wilkinson, 2019) or introducing the "knowledgeable other" to ensure the effectiveness of LS (Takahashi, 2013; Takahashi & McDougal, 2016).

The literature also contains case studies of small- and large-scale implementations of LS in dozens of other countries, including Denmark, Norway and Sweden. We note here that most such studies report some kind of immediate "success" of LS. For example, teachers have reported on such success and what they learned from the experience; but in this paper, we will not be concerned so much with the effects of LS on teachers and students as with the institutional conditions that allow or hinder the continued practice of LS.

What are the necessary and sufficient conditions for LS to become a regular and teacher-led activity, particularly in a Danish context? What could these conditions be, in general? The question was reformulated and approached during ICME13 in Hamburg as, *What makes lesson study "viable" in some contexts, where the activity is regularly done by all teachers with no external initiative?* (Qauresma et al., 2018, p. vii). To examine an indeterminate "what" in a research context, a theoretical framework, which specifies the boundaries and kinds of the research object, is required. The contribution of this paper is mainly to introduce and exemplify the recent theoretical construct of *paradidactic infrastructure* as a means to frame studies that investigate the above question more precisely. Within the Anthropological Theory of the Didactic, this construct models the conditions under which the teaching profession operates (in particular, learns) outside of teaching situations (Winsløw, 2012; Miyakawa & Winsløw, 2019).

We present this theoretical framework along with the more precise research questions on the paradidactic infrastructure for LS in Denmark after the background section. Then, we proceed to investigate our questions through two case studies, which illustrate important constraints and potentials for LS in Denmark.

Background: Lesson study in a Danish context

Danish researchers following the international literature became aware of LS starting in the late 1990s, along with everyone else (cf. introduction). It began to be disseminated to Danish teachers and teacher educators some years later via popularisation texts (Winsløw, 2004, 2006, 2009) and workshop presentations. The first implementation of LS in Denmark was in 2009 at the National Knowledge Centre for Mathematics Teaching (NAVIMAT, in existence until 2012). Led by Kaj Østergaard and Erik Bilsted, the elaboration of LS involved seven mathematics teachers, and Bilsted (2010) reported on this elaboration by presenting the lesson plans of seven "super lessons" together with a reprint of the work of Winsløw (2009). Since then, several other professional development and research projects based on LS involving both in-service and pre-service teachers, have been initiated. We focus here on LSs which have been documented in national and international academic journals.

A good amount of attention has been paid to LS as part of teacher education. Nyboe and Rasmussen (2015) described LS as a formal framework for practice elements in teacher education, which allows teaching to become a common and concrete object of concern for pre-service teachers, teachers and teacher educators. Jørgensen, Rostgaard and Mogensen (2016) presented a two-year research project in which LS was carried out by pre-service teachers in collaboration with their internship supervisors (who are experienced teachers). One conclusion was that LS fostered collaboration that aligned the goals of in-service teachers and internship supervisors as the internship became an avenue for knowledge sharing and a rewarding collaboration. These studies are mainly descriptive, but others are more theoretical. Skott and Østergaard (2015) used the *knowledge quartet* to categorise Danish pre-service teachers' mathematical *knowing* for teaching as it developed through LS during their internships. Østergaard (2016) used the the Anthropological Theory of the Didactic to explore the potentials of LS to improve pre-service teachers' learning during internship and concluded that LS constitutes an interface between teacher education and the internship programme.

Moving beyond the setting of teacher education, Skott and Møller (2016, 2017, 2019) used *Patterns of Participation* to conceptualise and understand teachers' *knowing* and participation from a social perspective. One of their aims was to contribute a deeper theoretical understanding of individual teacher learning and meaning-making derived from participation in LS. Another aim was to explore the conflicts that emerged when Danish teachers engaged in LS. The authors concluded that it is necessary to address broader issues of culture, such as teachers' collaborative work and power relations between the teachers in order to adapt LS in a Danish context. They also identified cultural norms as an explanation of implementation problems. Bahn (2018) used the Theory of Didactical Situations in the analysis of experienced teachers' work using open-ended lessons in the context of LS. He found that the open-ended approach was useful to some extent and that LS was effective in developing teachers' knowledge about the connection between teaching and learning. He also found that considerable external support was necessary for teachers' progress in this regard. Finally, in a mostly theoretical paper, Winsløw, Bahn and Rasmussen (2018) argued that the use of theoretical frameworks is crucial for systematic research on LS. They exemplified the use of the Theory of Didactical Situations and the Anthropological Theory of the Didactic using two Danish cases of pre- and in-service teachers' development.

We have not found documentation of Danish LS projects (involving the same teachers) which lasted for more than a few years or of research analysing the sustainability of LS in Denmark. Sustainability is an implicit concern in several of the documented implementation projects and in some of the more theoretical analyses of these projects, such as the identification of cultural issues as important conditions for implementing LS (Skott & Møller, 2019). We now propose a general and explicit framework.

Theoretical Framework and Research Questions

In the Anthropological Theory of the Didactic (Chevallard, 1999, 2006), we model knowledge in terms of praxeologies. Praxis refers to practical knowledge and "know-how" and encompasses tasks and techniques for solving tasks, and logos is explicit knowledge about the praxis and consists of technology, including explanations and justifications of the techniques, and theory, which is an abstract discourse clarifying or justifying the technology. In line with the work of Miyakawa and Winsløw (2019), we use the term *paradidactic* to refer to phenomena related to school teaching but found outside of teaching situations themselves. *Paradidactic praxeologies* refer to the practices and explicit knowledge (Chevallard, 1999), which teachers deploy and develop as they interact with other teachers, prepare or evaluate a lesson, and so on. As the examples show, some of these praxeologies are very similar to concrete *didactic praxeologies*, which occur in the classroom and are aimed at developing *mathematical praxeologies* among the students. Other paradidactic praxeologies, such as meetings or courses uniting all teachers from a given school, typically do not address any specific didactic praxeology and certainly not any mathematical praxeology.

The professional praxeologies of mathematics teachers constitute a very complex composition of mathematical praxeologies, didactical praxeologies and paradidactic praxeologies. LS activity demonstrates that this matter cannot be reduced to the forms of practice and logos, which are visible in mathematics classrooms, especially because in the latter, teachers are usually alone and cannot share their experiences and inferences with colleagues. The construction of professional knowledge requires, more generally, an infrastructure (frameworks, settings, organisations, etc.) that goes beyond the professional practice itself. In some professions, this infrastructure is prestigious and visible, for example, the medical profession with its international scientific journals, conferences, organisations, etc. In the case of mathematics teachers, there is also such an infrastructure - the *paradidactic infrastructure* (cf. Winsløw, 2012; Miyakawa & Winsløw, 2018) - which consists of the shared conditions (institutions, levels, media, etc.) that govern and nurture teachers' work outside the classroom, whether they work in large or small groups or alone (as is often the case for the day-to-day work related to teaching). Certainly, some of these conditions are quite visible and almost trivial, such as the time and workspace which teachers are given to prepare their teaching and the official documents and other resources, such as textbooks, that are available for this work. Other conditions may be less visible and more difficult to describe. LS is but one of the elements of conditions that are shared by Japanese teachers, and it depends on multiple other elements of the Japanese paradidactic infrastructure.

Miyakawa and Winsløw (2019, Table 1) displayed a wider field of the paradidactic infrastructure in Japan. The consideration of such conditions drew our attention to the fact that paradidactic infrastructure is often *national* in nature. Unlike the medical profession, the profession of mathematics teaching does not possess a solid and internationally shared knowledge base, with the possible exception of that which pertains strictly to the mathematical praxeologies that it exists to disseminate. The experiences, results and assumptions of the profession are still mostly shared at a national level or even more locally. The internationalisation of mathematics education research over the past century or so has the potential to produce a wider knowledge base, but unlike what holds for the medical profession, it remains mainly known to a small segment of scholars, despite their efforts to interact with teachers and contribute to the paradidactic field of activity, including pre- and in-service teacher education. The national differences of paradidactic infrastructure also imply that the investigation of the conditions for conducting LS has to be done at a national level.

In particular, we are interested in *sustainable LS*, meaning LS that is conducted regularly over several years and with resources normally available to schools (i.e., not as an initiative or with external funding). This does not entirely exclude external support (even in the form of funds); for example, in Japan, "knowledgeable others" are usually present (Takahashi, 2014), but the activity must not be merely an initiative by external institutions and should not, in principle, be limited in time.

We focus on the paradidactic infrastructure of Danish mathematics teachers in primary and lower secondary school and its implications for the possibility of sustainable LS. The reason for limiting our scope to these levels is that these teachers share a number of important institutional conditions: their initial education takes place at university colleges and is very different from the university programmes that prepare upper secondary teachers. In Denmark, the primary and lower secondary levels (grades 1–9) are mandatory and delivered in the same school institutions unlike higher-level education. The relative homogeneity of the teachers is also reflected in common work conditions, a common union for teachers and the fact that the schools in which they teach are either administered by one of the country's 98 municipalities or state-sponsored private institutions (catering to roughly 20% of the children in grades 1–9).

With these preliminaries and assumptions, we present in Table 1 a rough overview of the main institutions and elements of the paradidactic infrastructure in Denmark (PID) in analogy with the Japanese case (Miyakawa & Winsløw, 2019, Table 1). Our first portrait of the PID is based on the authors' brainstorming sessions and informal conversations with colleagues. It is thus more a rough picture than a final map. We note that the institutional levels involved in Table 1 can be roughly connected to the levels of society and school in Chevallard's (2002, 2019) hierarchy of didactic codetermination.

We now explain, in a little more detail, the elements of the table that are the most important (indicated in boldface in the table). The School Department of the Ministry of Education, currently The Department for Teaching and Quality, is in charge of the national, centralised regulations and management of schools. In particular, it produces the national "Common Goals", which specify the learning content for each subject at a given grade level, and creates (in part by delegation to committees or companies) a number of mandatory tests corresponding to these goals, including the exams at the end of lower secondary school. There is no centralised production or control of teaching materials. Most teachers rely on textbooks published by commercial editors usually authored by 2–4 teachers or teacher educators. Teacher educators are generally employed at University Colleges, where primary and lower secondary teachers are initially educated. These institutions also engage in development projects and in-service education related to school teaching, often financed by the Ministry of Education or by private foundations. Individual researchers at universities (such as the second author) may, exceptionally, be involved in such activities. Primary and lower secondary school teachers do not usually interact with universities except for the minority who (like the first author) pursue a masters or doctoral degree in addition to regular teacher education at the bachelor's level.

Institutional level	Institutional frameworks	Settings for teacher study and research	Media used or produced by teachers
National	**Ministry of Education** - **School department** Ministry of Higher Education and Research - **University colleges** - Universities National Teacher Union - Danish Math Teacher Association (MTA) **- MTA Publishing Co.** Math counsellors' network **Private foundations** Private textbook publishers	EMU (webpage with national curriculum and research) Teaching material centres Masters programmes Annual conference Theme conferences Book fairs	**"Common Goals", tests** Skolekom lists Folkeskolen.dk "Mathematics" Magazine Books and other resources **Commercial textbooks**
Region	Regional sections of MTA	In-service courses	
Municipality	School commission **School administration** **Educational Centre/consultants** Math Resource Centre /Counsellor teachers	In-service courses New teaching materials, meetings and courses on specific themes	
Local school	**School principal** School management Math team(s)	Meetings (for all teachers) Math team meetings	School Intranet

Table 1. Elements of Paradidactic Infrastructure in Denmark.

The public schools are funded and managed by the 98 municipalities. Each school has a board (appointed by the municipality) which appoints a school principal who is responsible for the day-to-day operation of the school, including the hiring of teachers and regulation of their working conditions. The principal appoints other members of the school management, who may take on specific pedagogical responsibilities. At the municipal level, within the school administration, there is also a unit - sometimes called the Educational Centre - which supports teachers through in-service courses and the like, usually staffed by consultants who may be specialised in some school subject(s). The organisation of this service depends on the municipality (size, political choices).

As noted in the previous section, LS in Denmark has so far mainly been initiated by individuals from university colleges and universities, often in relation to short-term projects, such as PhD dissertations. Such projects, and indeed the two cases considered, could offer some insight into conditions which favour or inhibit LS and enable us to hypothesize on sustainability requirements. More precisely, we will use our model of PID (Table 1) and the two cases to investigate the following questions:

RQ1: How do the different elements of PID affect (support, hinder) specific efforts to implement LS?

RQ2: What can be learnt from such specific cases regarding the potential for sustainable LS activity in Danish schools, for instance in terms of necessary changes in PID or in the method of implementing LS?

The two case studies are largely explorative (Streb, 2012) in the sense that we use them to generate hypothetical answers for the above questions, which are supposed to illustrate the meaning and pertinence of the proposed framework and of the concrete model of PID presented above. The hypothetical responses to RQ1 are summed up after each of the cases, and RQ2 is discussed in the discussion and conclusion.

The two cases are not comparable with regards to the length of the studies, numbers of participating teachers, academic support or the mathematical content studied in the classroom. However, both are cases of LS implemented over a long period in a Danish context and are aimed at making LS a viable and sustainable part of teachers' professional development. The cases were selected to exemplify the different aspects of how PID affects the efforts to implement LS.

Both cases draw on documents, pictures and field notes collected by the authors while observing and participating in the two contexts. In the first case, audio recordings of meetings were also used. The data is used to support the generation and illustration of hypotheses and not as "proof" of their general validity.

Case 1: Learning to do Lesson Study is not Sufficient

ABC (pseudonym) School is a large public school near Copenhagen, and it is funded as a so-called resource centre of mathematics education in its municipality. In early 2018, the mathematics consultants of the municipality and the school management decided to establish a "LS environment" at ABC School in collaboration with two mathematics educators from the nearby University College (including the first author). Two explicit aims for this activity were cited: to increase teachers' explicit focus on students' mathematics competences (a key notion in the current curriculum) in the classroom and investigate how teachers’ collaboration can improve the quality of mathematics teaching. Given the role of the school, another aim was for teachers to subsequently help other schools establish a LS environment. LS was totally new to the teachers, who were often teaching alone; at the time of the study, they had not experienced collaborative situations, including observing teaching and students’ learning together.

The project began with an introductory workshop on LS, given by the mathematics educators, during which a teacher from another municipality (the one considered in Case 2) gave an open lesson, which was followed by a common reflection session for the participants (as a first experience of this part of LS). At the end of the workshop, the principal of the school that the presenting teachers came from narrated his experiences with LS. In short, the introductory workshop gave teachers of ABC School a first impression on how LS works. Later in the spring of 2018, two teams of mathematics teachers (grades 3 and 7) started their own LS with the two mathematics educators as facilitators. After a common workshop on "reasoning competence" and "problem solving competence" (cf. Niss and Jensen, 2002), they planned a research lesson and completed three cycles of teaching (observing), reflecting and revising the lesson plan.

In the autumn of 2018, one new LS team joined the first two, and the spring workshop was repeated, but this time, a teacher from one of the first teams gave the open lesson. Again, LSs were done within the team, and one teacher from each of the teams focused on how to facilitate and guide a LS process.

We will now discuss one of the research lessons for grade 3, taught in the autumn of 2018. The presentation is based on lesson plans, field notes from classroom observations, student worksheets (as shown in Figures 2 and 3), audio recordings of reflections sessions, project meetings with teachers and municipality partners and interviews with teachers. Reflections, evaluations and interviews relating to the PID were all transcribed and coded (Table 1).

The research lesson was the last of three, and it was the second time that the team of the three teachers participated in LS. In the research lesson and the following reflection, the three teachers, one of the mathematics educators, an invited "knowledgeable other" from a university college and a journalist from a major national teacher magazine, were present. The lesson plan had been revised twice based on the experiences from the first two trials and reflections. The target knowledge was students' understanding of fractions and their representations.

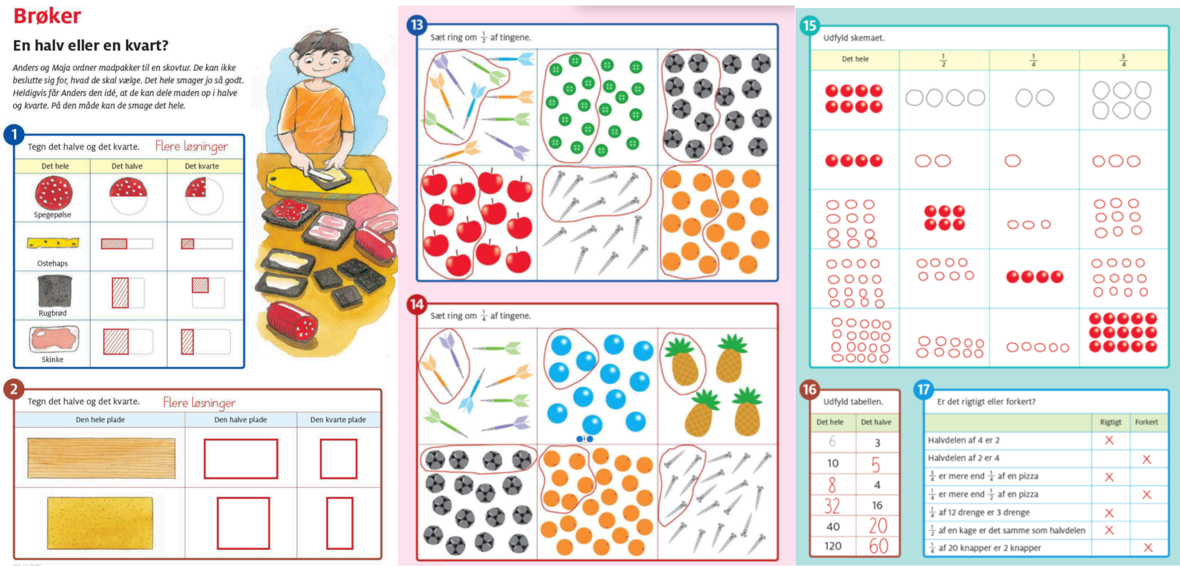


Figure 1: Grade 3 textbook: fraction tasks and examples of techniques (Lindhardt, Jensen & Møller, 2021)

In the textbook used at ABC school, the following type of tasks on fractions were covered: 1) draw the half and the quarter of a figure, 2) put a ring around ½ and ¼ of things, 3) fill out the form: whole, half, quarter, three quarters and 4) Is this right or wrong? (see the different tasks in figure 1). The suggested techniques included 1) making drawings of the correct size, 2) counting the total amount of elements in a group and marking ½ and ¼ of the element, 3) drawing whole-part relations and 4) ticking the correct box. The level of students' activities focused on practical knowledge and know-how techniques. The students were not asked to reflect on their techniques. While planning the research lesson, the teachers expanded the focus from praxis to also include logos (e.g., reasoning about the size of a fraction and the notation of a fraction).

The teachers' experience suggested that it was difficult for students to encounter fractions other than ½ and ¼ in their surroundings and everyday lives, where integers and decimal numbers appear more frequently. The teachers wanted the students to visualise fractions in different ways and formulated the fraction problems: *Why is 1/8 smaller than ¼, while 8 is greater than 4?* The broader goals of the research lesson related to students' inquiry skills and mathematical reasoning, involving work with different kinds of representations. As part of the planning of the lesson, the teachers designed a worksheet to support the students' work with drawings, written explanations, calculations and everyday stories (see Figures 2 and 3). The teachers found the fraction problem simple, but by working on the problem themselves, they came up with several potential ways for the students to reason and use representations to support their reasoning.

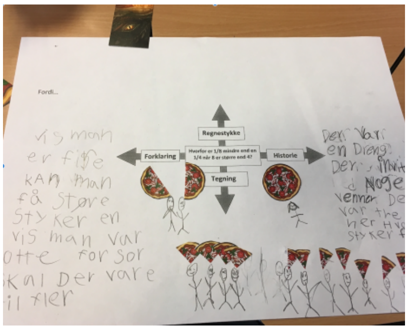


Figure 2: Worksheet with students' reasoning about sharing a pizza: "if there are four [people], you can have bigger slices [of pizza] than if there are eight [people] because you would need more pieces for more [people]"; the explanation is illustrated with drawings of people and pizzas. Everyday story: "A boy invites three friends; how big a slice of pizza does each of the children get?"

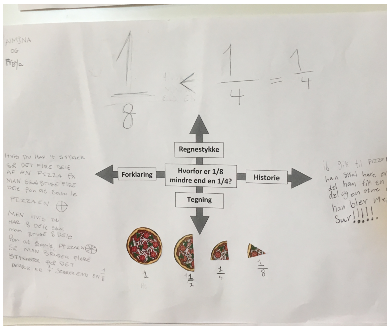


Figure 3: Worksheet with students' reasoning: "if you have four pieces of pizza, you need four pieces to form a [whole] pizza" (illustrated by a circle divided into four pieces), "but if you have eighths you need 8 pieces to form the pizza" (illustrated by a circle divided into eight pieces); "then you need more pieces and because of that, 1/4 is bigger than 1/8". On top: "1/8 < 1/4 = 1/4".

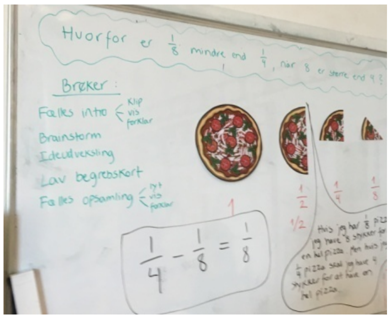


Figure 4: White board with students' reasoning and responses: "If I have 1/8 of a pizza, I need 8 pieces to make a full pizza, but if I have 1/4 of a pizza, I need 4 pieces to make a full pizza".

During the observation of students' work, teachers focused on their use of the different representations, especially the drawings and the explanations, to help students understand the problem. The teachers also noticed that students progressively used more formal mathematical language during the entire class discussion (partly reflected in Figure 4). During the following reflection meeting, the teachers stated what they had learned from observing students' dialogues, especially regarding their reasoning. Many of these points were quite generic and could be relevant beyond the research lesson. The participants also discussed the difficulties of orchestrating a whole class discussion and, in particular, drew substantial connections between different students' solutions.

The "knowledgeable other" stated that the lesson plan was robust and described the lesson as a complex lesson: "a delicious bite, where we get special knowledge about children's thinking on fractions, part and whole, in the process". He further complemented the lesson because the focus was on students' reasoning, and the teacher did not tell them what to do or think: "I categorise this as *reasoning without authority*".

In an oral evaluation of the experience, all teachers from the different teams indicated that they wanted to continue doing LS. They described LS as a certain kind of "noise" that, in a good way, forced them to be curious, experiment and make challenging changes in their teaching. This kind of "noise" was absent when teachers worked alone. In short, the teachers explained that they learned mathematics and teaching mathematics when they observed students' learning and discussed new ways to orchestrate mathematics lessons. However, the teachers also described the LS project as challenging; they felt that it "doubled" their work. Some were confused due to organisational obstacles, such as a substitute teacher not showing up. The teachers stated that they would like more "completed" lesson plans in Danish.

In the evaluation, the consultants from the municipality and the school management discussed how to support the project in a better way and how to construct a new structure around the LS. The consultants and principals had attended a workshop or a research lesson and the subsequent reflection, and they described the potentials of the project. Their focus was not on the teachers' learning to teach mathematics, including learning mathematics, but on their learning of how to do LS. They asked the teachers severally whether they were ready to start facilitating LS processes at some of the other schools in the municipalities and whether it was possible to organise open lessons, etc. In several ways, it seemed like the management and consultants had different goals from those of the teachers and mathematics educators.

In spite of the positive evaluation, the project stalled during the following winter as the municipality and school management became engaged in other projects. They cancelled several LS meetings and explained that it was difficult to find teachers who had time for the project. The mathematics educators' interpretation was that the municipality and principals saw the first two rounds of LS as a completion of the goal of learning how to do LS and had moved on to other goals.

To sum up this case, at the beginning of the project, LS was strongly supported by three institutional levels of PID: the university college and the "knowledgeable other" from the national level, consultants from the municipality level and the school management from the local school level. During the project, the teachers had mixed experiences with LS. On one hand, they were confused about organisational obstacles (e.g., extra math team meetings), which were not scaffolded by the school management and on the other hand, the teachers experienced professional development during LS. The teachers' development was not recognised by the school management or the consultants from the municipality. During the second year, the support for the two last levels disappeared, and the project was no longer viable. What will happen in the future is hard to predict, but the teachers' positive stance on LS may influence the school and municipality in the long run.

Case 2: Learning During Lesson Study Requires External Input

Lyngby-Taarbæk is a municipality north of Copenhagen, with about 55000 inhabitants and nine ordinary public schools. As in other large municipalities, the central school administration employs "consultants" to support the schools' teaching in specific subjects, for instance, in relation to new national initiatives and more generally, cross-school collaboration and teachers' in-service education. The present case concerns an in-service course organised by the consultant for mathematics and science, Dr Jacob Bahn. He took up this position after completing his PhD thesis (2018) on LS as a method for implementing inquiry-based mathematics teaching. The actual experiments for the thesis were carried out at three schools in the Lyngby-Taarbæk municipality, where teachers and pre-service teachers had experimented with LS occasionally in relation to pre-service teacher education and sporadic events. In 2018 and 2019, Bahn organised two in-service courses, which were attended by about 60 of the mathematics teachers in the municipality; some attended both courses. The courses lasted for about 8 weeks, involving five 3-hour sessions and some homework for participants. The idea was to engage all teachers in one LS, focusing on a common theme, which in 2018, was “mathematical inquiry” (with examples mainly from counting problems in relation to simple geometric shapes). Teachers signed up for the courses as “teams” of at least three teachers from each school.

The courses started with classical coursework related to the theme and subsequently, teamwork to design a research lesson with some support from course teachers. Following the first two or three sessions, all teams experimented with their lesson at their own school, with one of the members teaching (usually one of their own classes) and the rest of the team observing and participating in the reflection meeting. In most cases, Bahn also participated as an observer and “knowledgeable other” (*koshi* in Japanese). After the experiments, a course session was devoted to the teams for them to give short PowerPoint presentations of their lessons and observations (often involving pictures and short video clips from the lesson experiment). At the end of this session, the participants chose which lessons to observe together in the final sessions. These sessions were organised as “open lessons”; the selected teams invited all course participants to observe, and the teams discussed their research at their schools.

We now focus on the work carried out by one of the teams selected to present an open lesson in the 2018 course; we call this team “T”. The analysis is based on field notes from the research lesson, the lesson plan and student worksheets (as shown in Figure 6). The description and analysis of the case was verified by the organizer of the course and lesson study in question; only a few adjustments were needed, mainly concerning the context of the in-service course.

Most of the teachers from T had grade 3 classes and initially decided to design their research lesson for this level. They initially discussed the difficulty of having students do “mathematical inquiry” at this level, where they were still working on basic operations in addition to other basic skills, such as reading. They then selected a problem from teaching material that they had initially studied together and that one of them had previously used at this level. It concerned what the participants called “investigating the coordinate system”, which is a new and difficult topic in grade 3. The problem selected was based on a figure from the textbook material, which was also used as the first handout in the lesson (Figure 5). In Figure 5, there was also a question:

“How far is it from the car to the house?” In the lesson plan, the teachers formulated a more elaborate problem:

> Carla’s parents have bought a new house. Carla’s mother says that Carla must find her way home from the car. How many routes can Carla choose (the shortest route)?

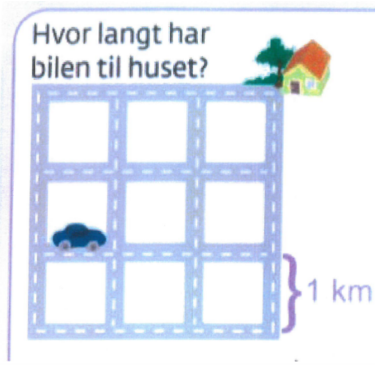


Figure 5: First handout (copy from Teglskov & Kristensen, 2012, p. 50).

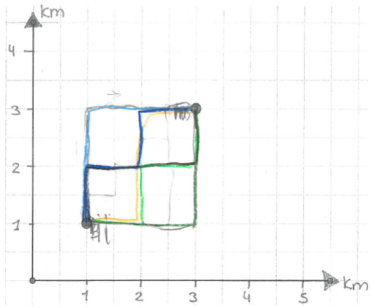


Figure 6: Coordinate system handout with student writing in colour.

Figure 7: Whiteboard writing, in part by a group of students (arrows also present in their personal writing).

In the open lesson observed, the teacher first devolved the question in Figure 5, and after about 10 minutes, the students’ answers were formulated and validated in a short whole-class session. Here, the teachers only talked about the results and methods while pointing at a smartboard showing Figure 5 (and no other visual support). Some students made measurements (with rulers) along the line segment from the car to the house; the need for the car to stay on the road and hence, to measure distances along these was institutionalized by the teacher, based on formulations by students. With this convention, students agreed that the distance was 4 km. Then, the teacher devolved the question from the lesson plan (quoted above) and distributed the handout shown in Figure 6. The students were seated in groups but got individual handouts; some students worked together but still drew on their own handouts. The teacher realised that some students considered all grid lines in Figure 6 to be roads and interrupted the group work to explain how the roads appeared in the coordinate system using a whiteboard as shown in Figure 7. The last five minutes of the lesson was a whole-class episode during which the whiteboard was used again (see Figure 7; the arrows were added during the episode). During this episode, coloured routes were drawn, and proposals that were not 4 km long were rejected but not drawn; the group who invented the arrow notation explained their reasoning while the teacher summarized on the whiteboard (to the right). The fact that there were six routes was not questioned or validated, but based on the handout and observations, only the group that used the arrow notation had found as many as six, while others found three or four. The lesson ended with the teacher reiterating the explanation of the six routes (Figure 7).

During the reflection session, teachers provided many interesting observations. Some pertained to the handouts, which contained some confusing aspects for students (in Figure 4, the car looked 1 km long, the unclear location of the car, etc.; in the second handout, the grid lines could be confused with roads, as seen in Figure 6). Others questioned whether the two problems were pertinent to inquiry as they had unique answers. Many comments pointed out that the teacher did not manage to guide all 29 students or ensure they got feedback on their work. Some questioned the clarity of the tasks, including the rules to follow. For instance, students drew routes that were longer than 4 km, and some even talked about "finding as many routes as possible".

Dr Bahn pointed out that important points of the two tasks were not made clear to the students: most students had measured the "bird fly distance", and the reason given for the rejection of this was just a rule ("cars run on roads"). Similarly, whether the result (six) for the last task was actually correct, was not discussed or explained, although reasoning is an important part of inquiry. The knowledgeable other pointed out the richness of the last task and the need for systematic reasoning to get more than "some examples of routes", as many students did. He focused on the "arrow notation" invented by one student group and explained how it can be developed into a systematic way of keeping track of different routes even for the generalized problem.

The episode illustrated several more general challenges with LS that were also noted by Bahn (2018). Teachers often deviated substantially from the lesson plan in actual lessons, for instance, by changing the way that the main problem was explained to the students (here, in two steps instead of one, as the plan suggests) and by adding explanations in the middle of a period of problem solving. While T had foreseen (and in the first experiment, observed) the tendency of students to measure distances diagonally, they developed no ideas to deal with it other than institutionalizing the "taxicab" distance as correct. Finally, the last counting task, which was rich in potential for inquiry, was not thoroughly analysed by T. In the lesson plan, the solution ("six") was not even mentioned or justified, and we found no similarities to the earlier experiment in terms of possible student solutions. All of these shortcomings are of course important to notice in order to improve the teachers' experience with LS and their mathematical and didactical knowledge.

In summary, in this case, the school and the municipality, along with the (national) university were major factors in enabling and sustaining the LS activity. The course organised by the municipality was an element of PID that could (if pursued over several years) make LS viable as an activity, assuming strong coordination with the school management and input from external individuals educated at the postgraduate level in practice-related didactics and mathematics. We have reason to believe that the long-term effort with LS in the municipality has considerably affected a number of leading teachers but analysing individual teachers' experiences is outside of the scope of the current paper.

Discussion and Conclusion

Both cases demonstrate the decisive role of the municipality and school in PID in terms of initiating, modifying or even cancelling LS activities. Takahashi and McDougal (2017) made a similar observation in the U.S.: to ensure sustainability, "enthusiasm for lesson study" must be clearly established at the school level; purely organisational support is not enough. It is crucial for the decision makers to be informed about the long-term goals and functions of LS in countries, such as Japan, where such formats are a regular part of teacher development. While LS does not have a strict format, almost all of these limited experiences of Danish teachers were attained in collaboration with other teachers who were equally inexperienced with the format and with educators or researchers who usually also had very limited experience (occasionally, including some observations in Japan). It is also worth noticing that Danish teachers are novices in developing lesson plans and do not have the support of the Japanese mathematics education literature on LS. So, as in other countries where the format is new, "learning to do LS" in Denmark is a considerable challenge.

We propose that sustainability should not merely be taken to mean an autonomous LS entity where teachers, after intensive support with professional development and new resources, can continue without external initiatives. The cases we have examined suggest, rather, the importance of external academic support from entities, such as the "knowledgeable other", as well as organisational support from the management at the local school and consultants from the municipality. This vision of sustainability is also consistent with other experiences and positions expressed in the research literature. Clivaz and Takahashi (2018, p. 161) note that “Lesson study cannot thrive in a vacuum; the greater its support, the greater its impact will be”. In our case, we consider that PID must provide more than substantial conditions and knowledge on “learning to do LS” and must provide professional nourishment and possibilities to share knowledge. This suggests an inclusive expansion of PID, more substantial than what is suggested by, for example, Akiba and Wilkinson (2016) in the form of prepackaged LS kits and LS steps.

Yet, LS is not an end in itself; its *raison d'être* in countries like Japan is to develop teachers' professional knowledge over the duration of their careers and at the same time, make this knowledge shared among members of the professional community (cf. Stigler & Hiebert, 1999). Although there have been reports on (usually and broadly speaking) "novice LS experiences" as successful and beneficial for teachers, such limited experiences with LS will at best, show the teachers how the activity works. With no infrastructure to engage in LS afterwards, such endeavours are as useless as a train with no tracks.

The paradidactic infrastructure can, thus, not be reduced to punctual opportunities to engage in a LS. The learning gained from LS depends on many factors, and our limited experiences and most studies on similar endeavours in other countries suggest that more or less seasoned "guides" of the activity are needed in order to create a meaningful, professional experience. In Japan, the "knowledgeable others", who are crucial participants in any LS, often come from the profession itself (Takahashi, 2014). As expert teachers who lead and guide the planning of lessons, they have simply learned the craft from participating in LS with more experienced peers. It may look like the hen and the egg, but even in Japan, LS grew out of the experience of learning from external experts (Isoda, 2007). The two cases show how the inclusion of such experience also make a difference to the outcome in the Danish context. The development of PID to make LS an important and useful format for professional development will thus require a focus not only on the creation of the time and place for teachers to engage in LS but also on the training of a number of "experts" who can contribute to guiding the activity and who can act as knowledgeable others in reflection meetings.

References

Akiba, M., Murata, A., Howard, C., Wilkinson, B., & Fabrega, J. (2019). Race to the top and lesson study implementation in Florida: District policy and leadership for teacher professional development. In *theory and practice of lesson study in mathematics* (pp. 731–754). Springer, Cham.

Akiba, M. & Wilkinson, B. (2016). Adopting an international innovation for teacher professional development: State and district approaches to lesson study in Florida. *Journal of Teacher Education*, *67*(1), 74–93.

Bahn, J. (2018). *Enquiry based mathematics education and lesson study.* [Unpublished doctoral Dissertation, University of Copenhagen]. Retrieved Dec. 19, 2019, from https://www.ind.ku.dk/begivenheder/2018/bahn/

Chevallard, Y. (1999). L'analyse des pratiques enseignantes en théorie anthropologique du didactique. *Recherches en Didactique des Mathématiques*, *19*(2), 221–266.

Chevallard, Y. (2002). Organiser l'étude. 3. Écologie & regulation.'. In J. L. Dorier, M. Artaud, R. Berthelot & R. Floris (Eds.), *Actes de la 11e École d'Été de Didactique des Mathématiques* (pp. 41–56). Grenoble: La Pensée Sauvage.

Chevallard, Y. (2006). Steps towards a new epistemology in mathematics education. In M. Bosch (Ed.), Proceedings of the IV conference of the European society for research in mathematics education (pp. 21–30). Barcelona: Universitat Ramon Llull.

Chevallard, Y. (2019). Introducing the anthropological theory of the didactic: An attempt at a principled approach. *Hiroshima Journal of Mathematics Education*, *12*, 71–114.

Clivaz, S. & Takahashi, A. (2018). Mathematics lesson study around the world: Conclusions and looking ahead. In M. Quaresma, C. Winsløw, S. Clivaz, J. P. da Ponte, A. NíShúilleabháin & A. Takahashi (Eds.), *Mathematics lesson study around the world: Theoretical and methodological issues* (pp. 153–164). New York: Springer.

Isoda, M. (2007). Where did lesson study begin and how far has it come? In M. Isoda, M. Stepens, Y. Ohara and T. Miyakawa (Eds.), *Japanese lesson study in mathematics: Its impact, diversity and potential for educational improvement* (pp. 8–15). New Jersey: World scientific.

Jørgensen, E., Rostgaard, P., & Mogensen, A. (2016). Lektionsstudier i læreruddannelsens praktik - et professionelt løft? [Lesson Study in teacher education practicum a professional enhancement?] In B. O. Hallås & G. Grimsæth (Eds.), *Lesson study iennordiskkontekst* (pp. 123–141). Oslo: Gyldendal.

Lewis, C. (2002). Does lesson study have a future in the United States? *Nagoya Journal of Education and Human Development*, *2002*(1), 1–23.

Lewis, C. & Perry, R. (2014). Lesson study with mathematical resources: A sustainable model for locally-led teacher professional learning. *Mathematics Teacher Education and Development*, *16*(1). Retrieved from https://files.eric.ed.gov/fulltext/EJ1046670.pdf 1

Lindhardt, B., Jensen, H. N. & Møller, M. T. (2021). *Kontext +3B, Lærervejledning/Web*. Alinea. Retrieved from https://kontextplus.alinea.dk/course/QsBd-kontext-3b

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.

Niss, M. and Jensen, T. H. (2002). *Kompetencer og matematiklæring - ideer og inspiration til udvikling af matematikundervisningen i Danmark.* Copenhagen: Undervisningsministeriet.

Nyboe, H. & Rasmussen, K. (2015). Lektionsstudier som praksissamarbejde. *UngePædagoger*, *2015*(2), 31–42.

Østergaard, C. H. & Skott, C. K. (2016). Establishing a mathematical lesson study culture in Danish teacher education. Presentation at WALS, 2016.

Østergaard, K. (2016). *Teori-praksis-problematikken i matematiklæreruddannelsen - belyst gennem lektionsstudier.* IMFUFA tekst nr. 501/ 2016. Roskilde University

Quaresma, M., Winsløw, C., Clivaz, S., da Ponte, J., NíShúilleabháin, A., & Takahashi, A. (Eds.) (2018) . *Mathematics lesson study around the world: Theoretical and methodological issues.* Cham: Springer.

Quaresma, M., Winsløw, C., Clivaz, S., da Ponte, J. P., NíShúilleabháin, A., & Takahashi, A. (2018). *Mathematics lesson study around the world.* Cham, Switzerland: Springer, p. vii

Skott, C. K. & Møller, H. (2016). Et deltagelsesperspektiv på lektionsstudier som metode til lærerudvikling i Danmark. In B. O. Hallås & G. Grimsæth (Eds.), *Lesson study i en nordisk kontekst* (pp. 230–249). Oslo: Gyldendal.

Skott, C. K. & Møller, H. (2017). The individual teacher in lesson study collaboration. *International Journal for Lesson and Learning Studies*, *6*(3), 216–232.

Skott. C. K. & Møller, H. (2019). Adaptations of lesson study in a Danish context: Displacements of teachers' work and power and relations. *Teaching and Teacher Education*, *87*(2020). https://doi.org/10.1016/j.tate.2019.102945.

Skott C. K. & Østergaard, C. H. (2015). Lektionsstudier i dansk læreruddannelse. *Nordic Studies in Mathematics Education*, *20*(1986), 101–122.

Stigler, J. & Hiebert, J. (1999). *The teaching gap*. New York: The Free Press.

Streb, C. (2012). Explorative case study. In A. Mills, G. Durepos & E. Wiebe (Eds.), *Encyclopedia of Case Study Research* (pp. 372–373). Thousand Oaks: Sage Publ. Inc.

Takahashi, A. (2014). The role of the knowledgeable other in lesson study: Examining the final comments of experienced lesson study practitioners. *Mathematics Teacher Education and Development*, *16*(1), 4–21.

Takahashi, A. & McDougal, T. (2016). Collaborative lesson research: Maximizing the impact of lesson study. *ZDM*, *48*(4), 513–526.

Takahashi, A. & McDougal, T. (2017). Collaborative lesson research (CLR). In M. Quaresma, C. Winsløw, S. Clivaz, J. P. da Ponte, A. NíShúilleabháin & A. Takahashi (Eds.), Mathematics lesson study around the world: Theoretical and methodological issues (pp. 143–152). New York, NY: Springer.

Teglskov, R. & Kristensen, B. (2012). *Multi3A, Elevbog.* Copenhagen: Gyldendal Uddannelse.

Winsløw, C. (2004). Quadratics in Japanese. *Nordic Studies in Mathematics Education*, *9*(1), 51–74.

Winsløw, C. (2006). *Didaktiske elementer. En introduktion til matematikkens og naturfagenes didaktik.* Copenhagen: Biofolia.

Winsløw, C. (2009). Et mysterium om tal - og japanske lektionsstudier. *MONA*, *2009*(1), 31–43.

Winsløw, C. (2012). 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. Mathematics curriculum materials and teacher development* (pp. 291–304). New York: Springer.

Winsløw C., Bahn J., & Rasmussen K. (2018). Theorizing lesson study: Two related frameworks and two Danish case studies. In M. Quaresma et al. (Eds.), *Mathematics lesson study around the world: Theoretical and methodological issues* (pp. 123–142). Cham: Springer.

8.3 Paper III: 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 statistical classrooms

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 statistical classrooms

Camilla Hellsten Østergaard

University of Copenhagen and University College Copenhagen

camilla.h.oestergaard@ind.ku.dk and choe@kp.dk

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 scrips but also analyse 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_k) 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 analyse 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 analyse the cases from an institutional perspective in order to understand how cultural factors shape statistical inquiry. We address the following questions:

RQ1: What are the dynamics between questions and answers in the Danish and Japanese lessons?

RQ2: What statistical praxeologies are developed as answers?

RQ3: What are the principal similarities and differences between the lessons?

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 centre 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 analysed 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 centre in a large municipality in Denmark. As a resource centre, 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…".

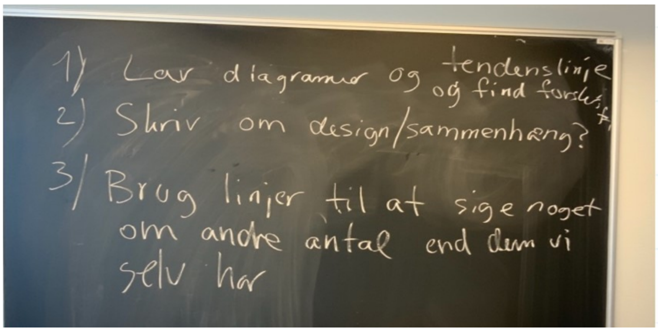


Figure 1. Teacher’s writing on the blackboard, translated into English: i) draw diagrams and trendlines and find the functional equation; ii) write about design/correlation? and iii) use lines to say something about quantities other than those we have.

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, others in GeoGebra. A few students succeed in constructing a scatterplot and trendline (Figure 2). After finding a trendline and a linear regression function, the

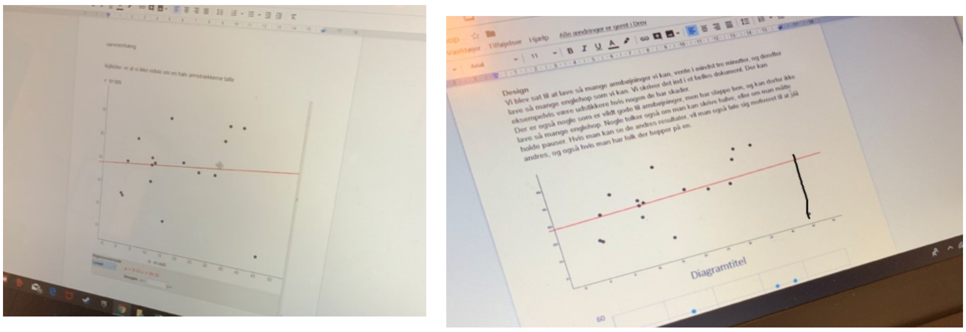


Figure 2. Two students’ scatterplots and trendlines.

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…”

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).

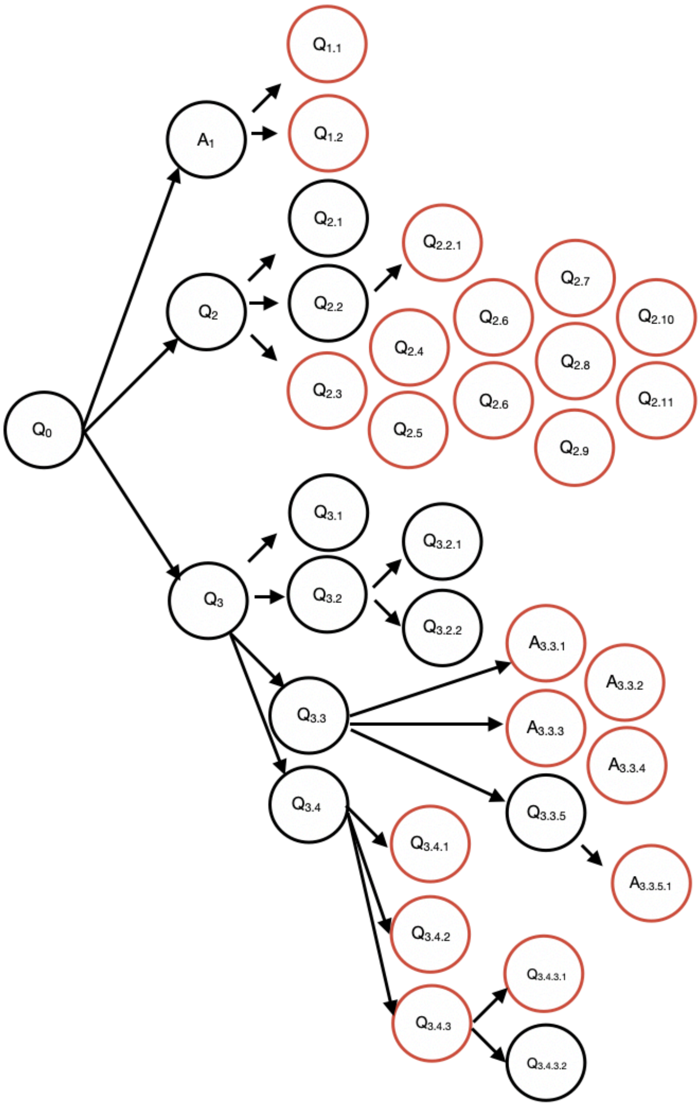


Figure 3. Q&A diagram of a Danish lesson. Contributions from the teacher are in black, while those from the students are in red.

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 trendline and find the linear regression function…"

$Q_{1.1}$: "How do I draw a trendline in Google Sheets?"

$Q_{1.2}$: "Google "how to find trendline 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 trendline 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:

T_1: Conduct an experiment to investigate the correlation between two (more or less strictly defined) observables.

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 trendline 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 trendline. The fourth technique (also instrumental) exhibits the equation. It is visible on the screen next to the trendline, 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 trendline 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 favour of that assumption but, instead, refers to his own subjective judgement: "In my head…" The dialogue about T_2 continues at a concrete level, but further theoretical considerations about the validity of the trendline 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.

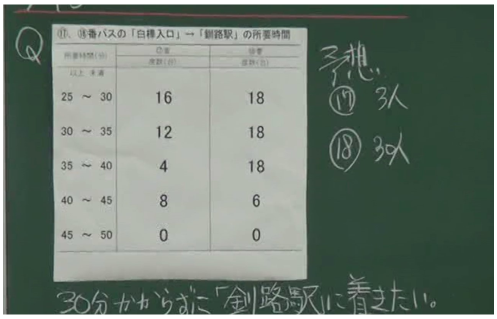


Figure 4. Frequency distribution of bus times.

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.

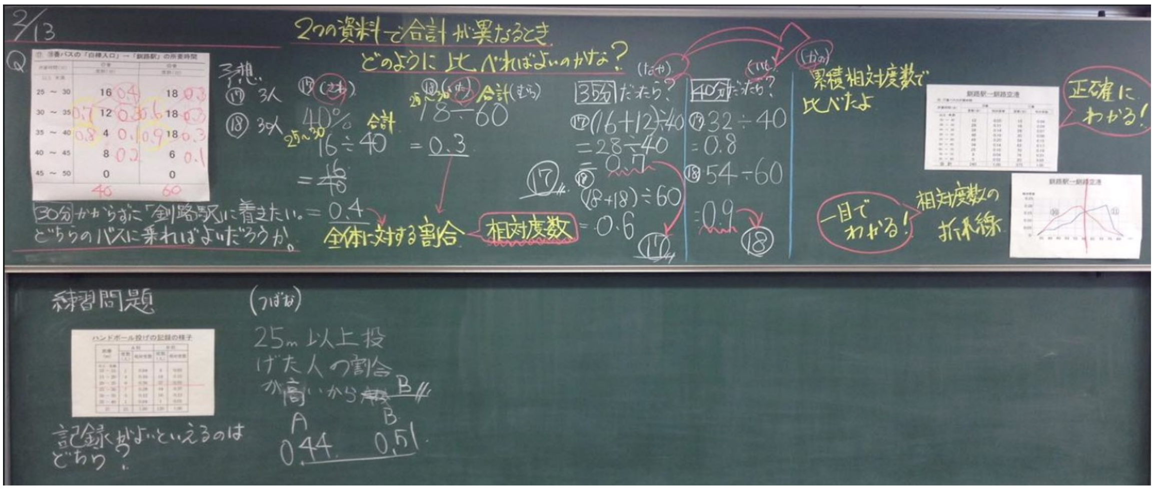


Figure 5. Blackboard with students' and teacher's questions and answer.

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.

The dynamics between questions and answers in whole-class dialogues

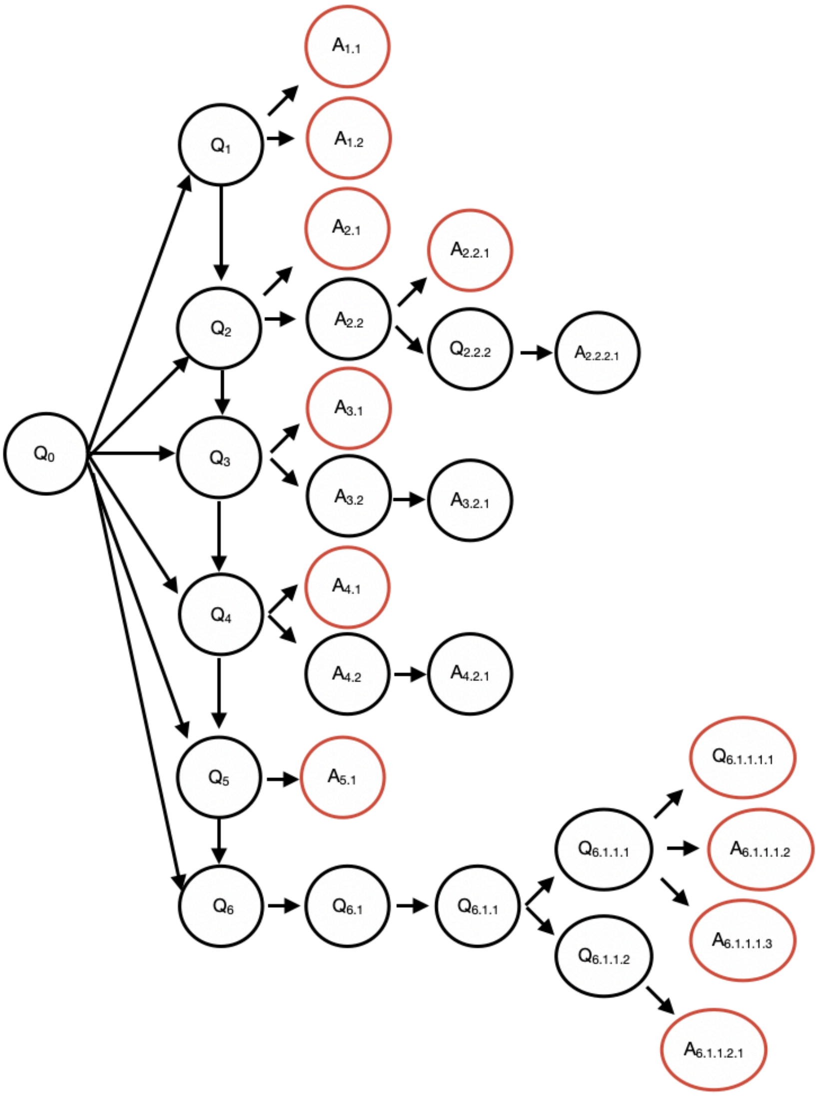


Figure 6. Q&A diagram showing questions and answers by the teacher (black) and students (red) in the Japanese classroom.

Q_0: "Which bus is better for us to ride?"

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 7)"

$Q_{6.1.1.1}$: "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 analyse 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. $A_{2.2.1}$, $Q_{2.2.2}$, and $A_{2.2.2.1}$).

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:

T: Compare two data sets based on grouped frequency tables

T_1: Develop "new" techniques to compare two (grouped) data sets

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 7 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 analysed 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. Analysing 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:

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 aims	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 black board 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 trendline, 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 (Miyakawa & Winsløw, 2019; Jessen et al., 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 organised 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 trendlines. 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).

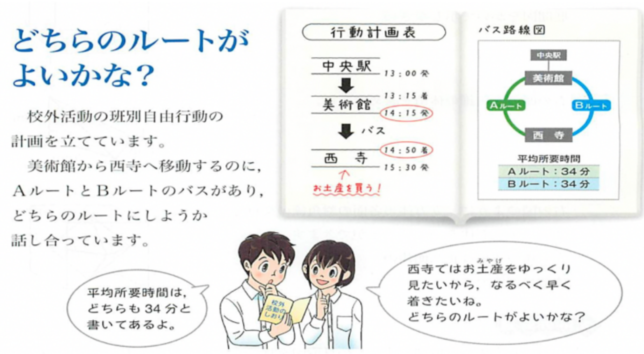


Figure 7. Textbook, *New Mathematics 1* (Fujii et al., 2016, p. 206). Introducing the problem: “Which bus should we ride?”

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**（そうたい）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 visualising the didactical process of the lesson. *Bansho* is a Japanese word meaning "(teachers') blackboard organisation" 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

Analysing 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 centre 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 trendline) 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 (Figures 3 and 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 analysing 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.

Acknowledgements

The author would like to acknowledge the help received from 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.

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 educati*on (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 miruzentangen, 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

8.4 Paper IV: Questioning the world – More than inquiry

QUESTIONING THE WORLD—MORE THAN INQUIRY

CAMILLA HELLSTEN ØSTERGAARD
University of Copenhagen and University College Copenhagen
camilla.h.oestergaard@ind.ku.dk *and* choe@kp.dk

DORTE MOESKÆR LARSEN
University of Southern Denmark and University College Lillebælt
dmla@ucl.dk

ABSTRACT

Based on the theoretical framework of the anthropological theory of the didactic, we present and exemplify a new approach to inquiry in statistics education and analyse how utilising this approach in grade five can contribute to students' development of statistical literacy and statistical reasoning. The analysis draws on a new definition of the theoretical distinction between statistical literacy and statistical reasoning. Our data come from designing and implementing a statistical inquiry approach in collaboration with a team of mathematics teachers. The results indicate that the dialectic of students' research and study of statistical knowledge in a digital milieu may form a recursive sequence of statistical learning situations.

Keywords: *Statistics education research; paradigm of questioning the world; study and research paths; statistical literacy and statistical reasoning*

1. SHIFTING PARADIGM

Statistical literacy is an essential life skill that provides individuals with the ability to act informed in social, political and economic situations (Ben-Zvi & Garfield, 2004). For students to become statistically literate, they must be introduced early to the ideas that underlie statistical literacy (e.g., Langrall et al., 2008), develop and appreciate an understanding of how statistical literacy provides a basis for decision making (Wallmann, 1992) and be engaged in statistical inquiry in explorative learning environments (Ben-Zvi et al., 2018).

In the research literature, there are several inquiry proposals that can be used in statistics education (e.g., Franklin et al., 2007; Turkey, 1977; Wild & Pfannkuch, 1999) and that provide didactic models of how experienced statistical researchers work. In the current article, we present and discuss a new inquiry approach: the paradigm of questioning the world.

Today, the prevailing educational paradigm is the paradigm of 'visiting monuments' (PVM; Chevallard, 2007; 2015). In PVM, the curriculum is formulated as a collection of works, or *monuments,* prescribing what is taught at schools. Here, the process of learning can be understood as a guided tour where the students visit works of knowledge and that involves learning what the works are made of and how the work can be used, for example, how the students can use the work 'to solve some given sets of problems – usually called *applications*' (Bosch & Ruiz-Munzón, 2022, p. 174). A consequence of the PVM is that students see school knowledge according to what Chevallard (2015) calls the *recycle bin/empty recycle bin-principle*, meaning that 'the knowledge taught may be forgotten or ignored, as soon as exams have been passed' (Chevallard, 2015, p. 176). In some cases, PVM develops situations in which students use statistical methods but do so without reflecting on whether the answers obtained are relevant to the question at hand, so the students might consider the visited work as something that has no relevance outside of school. To go beyond PVM, we need to establish situations where students can understand the world around them by addressing meaningful problems and elaborating on new questions.

This alternative paradigm is the paradigm of questioning the world (PQW; Chevallard, 2007, 2015). In PQW, the focus of the curriculum is on important questions to address—questions about the world. The questions are taken seriously and are not mere excuses to visit some explicit pre-established organisations of work (Barquero & Bosch, 2015). The answers in PQW are not known in advance, which means that students and teachers must develop a new kind of didactical contract (in the sense of Brousseau (1997)), in which "it is up to the students led by the teacher to identify the fields of knowledge useful for their inquiry" (Chevallard, 2022, p. 88). In such inquiry processes, students can follow many different paths: they need to inquire and search for established answers, study and validate the knowledge found, develop even more derived questions and articulate arguments and perspectives. PQW also includes some visiting of works because a new generation of students cannot reconstruct works anew, but they must search for and study knowledge in, for example, books and media.

In the present article, we examine how PQW may support the development and interaction of statistical literacy and statistical reasoning. We do so through a case study of how a grade five class explored the following question: ‘Are we physically active?’ This question generates an inquiry process with both derived questions and answers and is called the *generating question* of inquiry. Our study is a design research project (Bakker, 2018) with the double aim of studying didactic phenomena, PQW, and of theoretically developing the new educational approach to statistical inquiry. We identify and analyse certain *dialectics* in the concrete inquiry process, and at the end of the article, we discuss the wider potential of PQW in relation to statistical literacy. Our hypothesis is that students’ engagement in PQW may provide an ideal setting for developing statistical literacy, particularly that this could lead to a new dialectic between statistical literacy and statistical reasoning.

2. STATISTICAL CONCEPTS AND STATISTICAL LEARNING MODELS

To formulate learning goals for students’ statistical activities, there are several concepts and aims described in the literature: statistical literacy, statistical reasoning and statistical thinking. These concepts are often used interchangeably and there is a lack of clarity regarding the relationship among them and the extent to which they overlap (Sabbag et al., 2018). The lack of consistency in the literature shows that the concepts are still evolving in accordance with pace of change in the development of new forms of communication, new technologies and new ways of receiving and presenting data (e.g., Gauld 2017; Wild, 2017). In the current paper, we focus on the concepts of *statistical literacy* and *statistical reasoning*.

2.1. STATISTICAL LITERACY AND STATISTICAL REASONING

One of the fundamental arguments for *statistical literacy* is that individuals face certain common challenges within society. There are, however, many views of what constitutes a literate citizen. Statistical literacy has been described as the *basic skills* of understanding statistical results and the *process* of conducting, organising and representing data (Ben-Zvi &

Garfield, 2004, p. 7), *critical sense* (Gal, 2002), which includes critically interpreting, evaluating and communicating statistical information, and the ability to *appreciate statistical* knowledge (Wallmann, 1992).

There is no simple definition of statistical literacy. A commonly used characterisation—and more than a decade-old distinction—is the dichotomy between the role as data consumers and data producers (Gal, 2002) and the newer approach, that is, reading and writing the word and the world (Weiland, 2017). Reading is both classified as *selective reading*, a process of encoding the phenomenon, abstracting the data into models and further interpreting models, but also a reverse process of *imaginative reading*, where the reader deinterprets, deabstracts and decodes arguments, models and data (Büscher, 2022). The statistical literacy perspectives of consumers/producers and reading/writing both relate to a twofold perspective on statistical literacy, including seeing "statistics as a lens, enable a new view of the world" (Weiland, 2017, p. 37), but also seeing statistics as a possibility to change the 'world' by actively producing and analysing data and creating new contexts.

Several researchers have listed the minimum aspects of statistical literacy that must be mastered, for example, according to the changing role and nature of data. delMas (2002) describes basic literacy as the capacities to identify, describe, rephase, translate, interpret and read. With his separation between reading and writing, Weiland (2017) first includes quantitative argument and an evaluation of statistical information and, second, formulates statistical questions, collecting data, analysing data, interpreting data and discussing and communicating the meanings of statistical information.

Statistical reasoning can be defined as "the way people reason with statistical ideas and make sense of statistical information" (Ben-Zvi and Garfield, 2004, p. 7), including making arguments with statistical concepts, understanding statistical information and connecting to statistical concepts from which inferences can be drawn (Garfield 2002). Statistical reasoning also includes being able to explain statistical processes—why or how statistical results were developed and why a conclusion is justified (delMas, 2002). Statistical reasoning in early school years is connected to the development of informal ideas of inference (e.g., Gil & Ben-Zvi, 2011; Meletiou-Mavrotheris & Paparistodemou, 2015), which includes principles such as 1) making generalisations that extend 'beyond the data', 2) the use of data as evident for generalisations and 3) the use of probabilistic language in generalisations (Makar & Rubin, 2009).

Thus, the separation between the two concepts of statistical literacy and statistical reasoning is difficult. The difference resides in what we ask the students to do with data and statistical methods. delMas (2002) proposes that we look at the nature of the tasks given to students to identify whether the instruction promotes literacy or reasoning. If the goal is to develop students' statistical literacy, then teachers can ask students to identify examples, describe graphs, distributions and relationships, rephrase or translate statistical findings or interpret the results of a statistical procedure. If, instead, students must explain why or how the results are produced or why a conclusion is justified, the teacher should ask the students to develop their statistical reasoning (delMas, 2002).

Later in the article, we redefine statistical literacy and statistical reasoning in terms of praxeologies, a notion from the theoretical framework that we now proceed to present.

3. FUNDAMENTAL ELEMENTS OF THE ANTHROPOLOGICAL THEORY OF THE DIDACTIC

In the present study, we use the theoretical and methodological framework known as the anthropological theory of the didactic (ATD; Chevallard, 1992, 2006, 2007, 2012). ATD allows us to model teaching and learning (Bosch & Gascón, 2014) and develop new theoretical knowledge about these processes. The current article draws on ATD to identify and analyse dialectics in PQW and analyse statistical inquiry.

3.1. STUDY AND RESEARCH PATHS

The notion of *study and research paths* (SRP) is a design tool for teaching within PQW (Bosch, 2018; Chevallard, 2006, 2015). Students are engaged in an SRP through a generating question, Q_0, and are supposed to seek a final answer, $A^♥$. To do so, the students must do the following:

- Pose derived questions, $Q_1 .. Q_n$
- (Re)search for answers, $A^◊$, in media (e.g., the internet, textbook, lecture, news)
- Study the relevant works, w, of others
- Study and validate the answers, $A^◊$, found within these works

- Use different resources (e.g., digital tools, data material)
- Establish the milieus for exploring questions
- Elaborate on, discuss and validate a final answer, $A^{\heartsuit}$

SRP are not linear processes; instead, they unfold in many different (and unknown) paths (Bosch & Winsløw, 2015). In an SRP, the works, w, are not determined in advance, but instead, they must be found relevant to Q_0 and the successive derived questions Q_1.. Q_n.

We hypothesise that, in the domain of statistics, SRP can serve to expand students' statistical literacy because they interpret and evaluate statistical information in a meaningful context and analyse, discuss and communicate their answers (Gal, 2002). As in civic statistics (Nicholcon et al., 2018), the aim of SRPs is not to master certain given statistical skills, but to address questions of relevance. The questions to pose in PQW do not only address issues of public and social concern, but they also focus on contexts that are important and familiar to the students. In statistics, a wide range of disciplines may be drawn on to reach $A^{\heartsuit}$.

When shifting from PVM to PQW, the didactic contract (Brousseau, 1997) between teachers and students changes, and new classroom routines are required (Strømskag & Chevallard, 2022). Students must plan their inquiry process, develop and address questions, (re)search for answers and study the answers obtained.

3.2. DIALECTICS IN STUDY AND RESEARCH PATHS

As an SRP evolves in a classroom, different dialectics emerge. The dialectics are the driving dynamic of the SRP: dialectics of study and research, questions and answers, media and milieu, conjectures and proofs, individual and group work, black boxes and clear boxes, on-topic and off-topic and so forth. Below, we focus on three of these dialectics, which are strongly linked with one another.

Study and research

A study is defined as the activity of accessing existing knowledge, and *research* is the process of autonomous problem solving (Bosch & Winsløw, 2015). For a large part of teaching, study and research tend to evolve more or less independently. On the one hand, there are research-oriented activities (such as task solving), and on the other hand, students study a given piece of work. An SRP aims for students to engage in a dialectic between study and research, where new answers are developed, where students question the answers and where the experience of the study aims at finding answers (Winsløw et al., 2013). These dialectics are often not realised in lessons based solely on a textbook.

Questions and answers

Another dialectic is that of questions and answers. To analyse the question-and-answer dialectics, we define a *question* as a discourse that presents a problem to someone (Østergaard, 2022); thus, the notion of a question is to be understood semantically, not just grammatically (‘What do you think?’ is grammatically a question but not a concrete problem; cf. Bosch & Winsløw, 2016). An *answer* is a discourse that enables us to extend our practical and theoretical knowledge—in short, it represents a *praxeology* (Chevallard, 1999). Some answers are preliminary or partial, while others are final results from our study of Q_0. The process of SRP—and the ensuing dialectic of questions and answers—can be visualised in a tree-like question-and-answer diagram (Q&A diagram) (Jessen, 2014; Winsløw et al., 2013). In the SRP literature, Q&A diagrams are used to visualise the main questions and answers in the SRP, either as a tool for planning or to analyse a realised SRP (e.g., Østergaard, 2022; Winsløw et al., 2013).

Media and milieu

Media are the containers of information and can include books, websites, news media and lectures. Media always carry an intention to communicate some answers (possibly without explicit questions). In contrast, *milieus* are constructed to facilitate the explorations of questions. We distinguish two categories: one that includes already existing answers, $A^{\Diamond}$, from the available media or students already established knowledge and resources, for example, digital tools; and one where the theories, experiments and derived questions are brought into the milieu by the students or teacher (Kidron et al., 2014). Milieus may motivate students to critically assess (or *question*) such elements (Winsløw et al., 2013). The design and organisation of milieus are essential for the study process carried out by students and teachers. Media and milieu do not refer to a property of objects but rather to their specific use in the inquiry process. In PVM, the dialectic of media and milieu is often weakly balanced. The students use existing answers, $A^{\Diamond}$, from media, such as their textbook, without testing the validity of the information; and activities, for example, an exercise in a textbook, rarely require search for other media. This contrasts with how a statistician works (Kidron et al., 2014).

3.3. PRAXEOLOGIES

In ATD, any human activity is modelled as *praxeologies* (Chevallard, 1999). A praxeology models knowledge in terms of a *praxis block*, that is, know-how, and a *logos block*, that is, know-why. The praxis block is formed by various types of tasks and techniques. A task can be small, 'What is the mean?' or larger, 'Can you analyse the data?' The latter calls for a subdivision into smaller tasks belonging to types that are equipped with a specific technique. A technique is a way of carrying out a task. For mathematical tasks, it can be a traditional paper-and-pencil technique or an instrumented technique (based on a digital tool; cf. Lagrange, 2005). Techniques can be developed to be more powerful or easier to perform (Chevallard, 2019). The logos block is formed by technology and theory. The technology is a discourse on techniques that have the purpose of making the technique intelligible and providing an understanding why a technique is required, how it relates to other techniques and so forth. The technology is often built at the same time as the technique and refers directly to the type of task that it allows. The absence of technology leads to 'idiosyncratic' beliefs acting as fake technology (Chevallard, 2019). Theory is a more advanced and more abstract discourse than technology that is focused on the justification and unification of technologies, here independently of specific techniques. Theory "carries the idea of a distanced stance, that of a spectator – the meaning of Greek theoros – towards the activity domain the theory in question is concerned with" (Chevallard, 2019, p. 91). In statistics, it is not enough to reason on a specific set of data; instead, an explanation of abstract statistical processes and notions is needed. Praxis and logos do not exist separately but are mutually constitutive: "Praxis thus entails logos which in turn backs up praxis" (Chevallard, 2006, p. 23). Once answers have been developed, new praxeologies can be elaborated on by integrating established answers into new techniques in a more developed logos.

The following is an example of praxeological modelling of the intended knowledge in a task of finding the mean from a fifth-grade textbook.

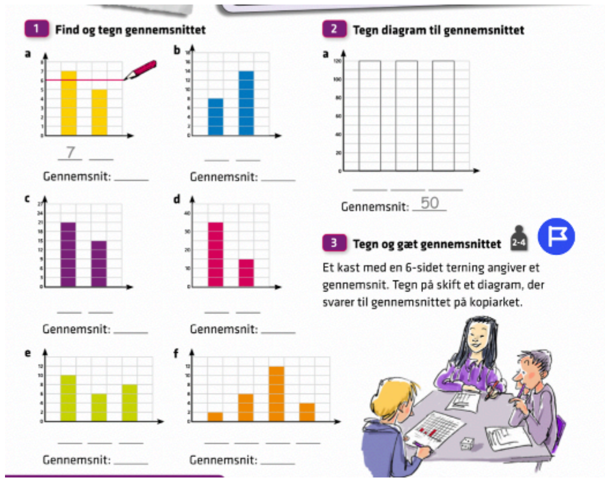


Figure 1. Find and draw the mean. Exercise from a Danish textbook (Anesen & Winth, 2010, p. 58)

Table 1. The praxeology of finding the mean by drawing in a bar graph.

Praxis	Type of task	Find and draw the mean of two observations, as visualised in a bar graph.
	Technique	Find out how much you have to move from one bar to another so that both of the bars have the same length (= the mean). Write down the sizes of the bars. Draw a line symbolising the mean. Write down the size of the mean.
Logos	Technology	There is no encouragement to discuss the technique given in the textbook. *However, the graphical representation can set a stage for dialogues about why we 'move' from one bar to another, illustrate the algebraic representation and give ideas on how to calculate the mean of a set of observations.*
	Theory	

In the current article, we reopen the dialogue about defining statistical literacy and statistical reasoning, suggesting the usefulness of praxeologies to this end. A tempting distinction would be to characterise statistical literacy as merely the capacity for engaging in praxis and statistical reasoning as solely based on logos. However, such a sharp separation would make statistical literacy entirely tacit and, thus, devoid of the important assets usually attributed to it, such as critical sense and an appreciation of statistical knowledge. In the general understanding found in the literature, statistical literacy includes being able to relate critically to specific statistical techniques for solving a problem, a process that requires technology in the sense described above. Statistical reasoning is, on the other hand, clearly resting on the logos block, here on both technology and theory, because reasoning requires not only discussing techniques, but also appealing to the general results and perspectives that go beyond these. Reasoning requires theoretical language, which refers to general statistical concepts, and because it does not directly involve techniques, it cannot be instrumented.

Thus, we define *statistical literacy* as the capacity for engaging with statistical types of tasks, techniques and technology and *statistical reasoning* as a discursive activity based on statistical technology and theory. Thus, we see the two as partially overlapping (Figure 2).

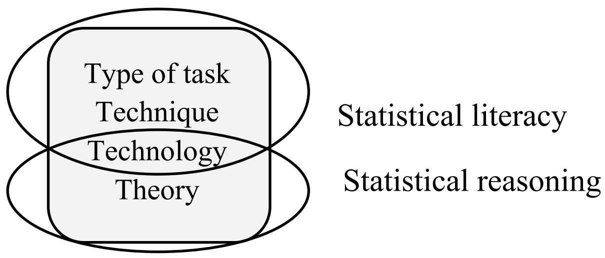


Figure 2. Definition of statistical literacy and statistical reasoning in terms of the praxeological elements on which they focus.

The development of statistical literacy and statistical reasoning is not a closed cycle with a beginning and an end, but rather, it is a process of developing the sequences of statistical praxeologies that become gradually more complex. They are both crucial to building connected and complex statistical knowledge.

4. RESEARCH QUESTION

Most research on SRP has been carried out at high schools (e.g., Jessen, 2014), in the field of teacher education (e.g., Barquero et al., 2018; Rasmussen, 2016) and at universities (e.g., Barquero et al., 2013; Barquero et al., 2018). Only a few studies have explored SRPs in primary and lower secondary schools (including some of Chevallard's personal, unpublished work). In the current article, we address the following research question: How can SRP contribute to middle school students' development of statistical literacy and statistical reasoning?

5. RESEARCH METHODOLOGY

The methodology in the current study follows the research principles of *design research.* (Bakker, 2018; Cobb et al., 2017). The methodology has four steps (Barquero & Bosch, 2015).

(1) Identify a *didactic phenomenon* to be addressed (here, the process of changing PVM to PQW: including how to orchestrate an SRP with numerous shifts between media and milieu, group work and whole-class conferences, change the didactical contract and introduce a new digital tool to analyse data).
(2) *Design of an educational activity* (here, SRP in the domain of statistics, development of *a priori* Q&A diagrams (Winsløw et al., 2013) and a more linear lesson plan to provide a scaffold for the teachers during implementation (Fujii, 2016)
(3) *Implementation, observation and clinical analysis of the educational activity*.
(4) *A posteriori analysis* based on the validation and redesign of the educational activity and drawing up the *corresponding consequences of the didactic phenomenon* initially identified.

In the analysis, the main focus is on steps three and four.

5.1. METHODS AND DATA COLLECTION

In the didactic engineering process, the first author worked with a team of three fifth-grade mathematics teachers to identify, design and implement an SRP. The SRP was implemented and further developed in a cyclic process in three fifth-grade classes. In the current article, we alone use data from the last design and implementation. In the analysis of data, we construct a descriptive case (Yin, 2014) in which we make an in-depth description and analysis of the implementation of an SRP in the fifth-grade class C.

Our background case data consist of video and audio-recorded teacher interviews, teacher workshops, planning SRP design meetings, three SRP and teacher reflection sessions. In addition, we have collected teachers' planning tools in the form of Q&A diagrams and lesson plans and students' productions constructed in *TinkerPlots* (software introduced later in this paper).

In the current paper, we do not analyse the didactic design reserach process as a whole but concentrate on the third implementation of the SRP. The main data consist of eight hours of video recordings. In the classroom, two cameras focused on the students: the interactive whiteboard and students' discussion and presentation of answers. The student group works were video recorded by the computer, recording the students in front of the computer and the screen at the same time. Other main data include the *a priori* Q&A diagram and lesson plan.

For the analysis of the classroom processes, we first constructed a Q&A diagram displaying the development of questions and answers observed in the classroom. To do so, we analysed parts of the whole-class discussions semantically, particularly the conferences. We then modelled the answers developed in the SRP in terms of praxeologies. The praxeological modelling process took its point of departure in the Q&A diagram, where the derived questions were identified as semantic questions. In the third part of the analysis, we interpreted the outcome of the SRP in terms of statistical literacy and statistical reasoning.

6. CONTEXT OF STUDY

The teachers were introduced to the SRP in two workshops, in which they studied different potential generating questions. In the workshops, the teachers were furthermore introduced to the digital tool TinkerPlots.

Statistical digital tools have changed the way students' access, analyse and visualise data. Digital tools offer several affordances, but the use of digital tools also raises new didactical questions of how we understand and improve statistical practices and how we facilitate and enhance statistical learning in the classroom (Fitzallen & Watson, 2010; Noll & Kirin, 2017). It is often hypothesised that digital tools allow for statistical creativity and are a faster way of processing data, especially to produce graphical representations, without students losing touch with the statistical meaning (Pratt et al., 2011). Instead of performing trivial tasks, the idea is for students to operate at the level of theory blocks. Dreyfus (1994) describes the affordances of digital algebra tools in that the students 'can concentrate on the operations that are intended to be the focus of attention and leave the lower-level operations to the computer' (p. 205). We see similar opportunities in dynamic data visualisation and modelling tools.

TinkerPlots is a dynamic data visualisation and modelling tool that supports inquiry-based approaches by scaffolding students' free experimentation with data in a trial-and-error style (Konold & Higgins, 2003). Some of the instrumented techniques are close to paper-and-pencil techniques, including ordering, stacking and separating data, while other techniques are 'push-button' techniques. TinkerPlot encourages students to use informal techniques and build their own intuitive models by searching for trends and patterns. TinkerPlots was chosen because of several valuable criteria: easy to use, representations in multiple forms, facilitate both statistical and natural language, provide extended memory when organising and reorganising data, allow for various entries for abstraction of concepts (Fizallen, 2007) and facilitate the students' thinking about statistics (Konold & Lehrer, 2008) in moving back and forth between different visualisations. In the SRP 'Are we physically active?', TinkerPlots formed a part of the milieu as a tool used to develop and study answers. The aim of the SRP was that the digital tool could be used to bridge among task, technique, technology and theory.

Another part of the milieu was the data that the students collected through a survey data about their physical activities. The data included 15 questions about gender, sleeping time, bike riding time, daily number of steps, time spent on digital devices and so forth. The data led the students to pose new questions and compare their data with international results found in newspapers and other media.

6.1. CONDITIONS AND CONSTRAINTS FOR IMPLEMENTING THE SRP IN PRIMARY SCHOOL

To understand the context of an SRP, it is important to analyse the conditions and constraints for implementing SRP in primary schools. Our baseline for this analysis was the scale of levels of didactic codetermination (Chevallard, 2002; Jessen et al., 2019), where we focused on the levels from school to discipline, particularly on school curricula and conditions for the participating teacher.

The Danish Ministry of Education formulates the Education Act (2006), the overall goal of mathematics and the Common Mathematical Goals (2019), which has, for example, the generic aim of "… experiences, immersion and enthusiasm, so that students develop imagination and gain confidence in their own possibilities and background to take a stand and act" (Danish Ministry of Education, 2019, p. 3) and mathematical aims grounded in "creativity, problem solving, argumentation and communication" (Danish Ministry of Education, 2019, p. 3) and "students must recognise the role of mathematics in a historic, cultural and societal context, and assess the use of mathematics to take responsibility and exert influence in a democratic society" (Danish Ministry of Education, 2019, p. 3). These goals all seem to call for inquiry approaches, such as an SRP.

The Common Goals in mathematics are general guidelines and need to be interpreted in relation to more concrete goals that should be realised in classroom teaching. The aims for the teaching of statistics in grades four to six are as follows: 'Students can do statistical inquiries, including use and interpretation of graphical representations of data; they can explore and present their own statistical inquiry; compare sets of data with the statistical models of frequency, relative frequency and other simple descriptors' (Danish Ministry of Education, 2019, p. 7). These wide goals do not indicate specific modes of teaching but again seem quite consistent with SRP as a didactic approach in statistics.

The institutional context for our case study was three fifth-grade classes, here called A, B and C, at a school in a rural area of Denmark. In the present article, we focus on an SRP carried out in class C with 22 students. The team of mathematics teachers involved two novice mathematics teachers and a mathematics teacher with more than 15 years of teaching experience. The experienced teacher also worked as a mathematics counsellor at the school, supervising other primary-level mathematics teachers. All teachers had little experience with digital tools in statistics (mainly Excel).

In total, the three teachers spent 50 hours (Figure 3) on an initial interview, a workshop introducing SRP and TinkerPlots and, subsequently, on the design, implementation, reflection and revision of the SRP.

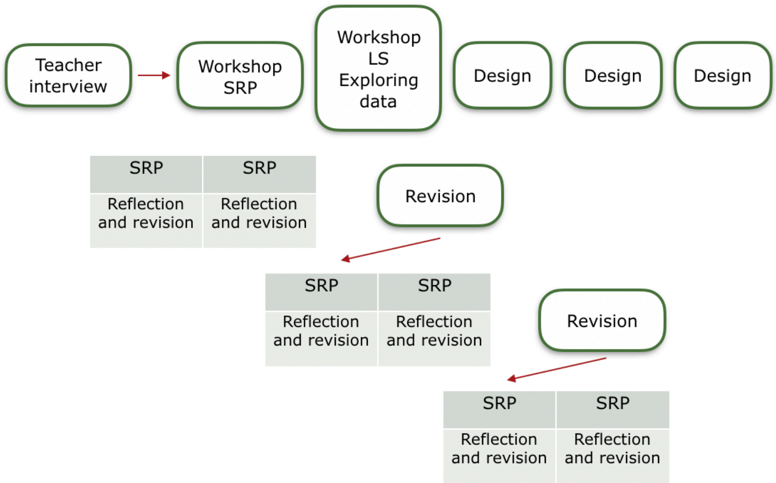


Figure 3. Illustration of the design of the didactic engineering project with the teachers

This setup contrasts with the sparse time teachers normally have to plan mathematics lessons. Two of the teachers had earlier participated in teacher development projects dealing with inquiry, innovation and technologies (one of the teachers is described in Larsen et al., 2022).

In their everyday teaching, the students worked with textbook exercises in smaller groups and with problem-solving tasks developed by the teachers. The students rarely presented their answers to the rest of the class or engaged in whole-class discussions of wider mathematical or statistical points. In general, there are no exams in grade five, but there is a national mathematic test.

To prepare for students' autonomous actions in the classroom, the teachers imagined what kind of work the students might visit and question (*a priori* Q&A diagram) and how they could encourage the study of works as a natural part of the investigation. In the preparation of the SRP, the teachers selected a 'light' science article about children's activity level, a newspaper article and a video clip from Children's News, in which Danish children were interviewed about their activity level. The teachers further discussed how to shift between students' group study and whole-class conferences, where the students presented and discussed questions, works and answers (lesson plan).

In the next section, we outline the study and research process actually realised.

7. REALISED SRP

7.1. 'ARE WE PHYSICALLY ACTIVE?'

The starting point for the SRP was the generating question, Q_0, 'Are we physically active?' Here, the students researched media and defined physical activity for children as 12,000 steps a day (Samvirke, 2014) or one hour of hard physical activity (Danish Health Authority, 2006; World Health Organization, 2020).

At first, the students discussed informatively whether they were active or not. Among the 22 students, 11 believed that they were active. In the process of finding a final answer, $A^{\heartsuit}$, the students searched for answers in the provided media (described above). In the *a posteriori* Q&A diagram (Figure 4), we can see the answers $A^{\Diamond}_2 \ldots A^{\Diamond}_{15}$ that students found in these media, along with how these answers led the students to pose derived questions, $Q_{2.1} .. Q_{15.1}$. To validate the answers, $A^{\Diamond}_2 \ldots A^{\Diamond}_{15}$, no students investigated the underlying data from the media, but instead, they focused on finding answers through their own survey and analysed if these answers, $A_{2.1.1} \ldots A_{2.1.x}$, were in line with the answers found in the media.

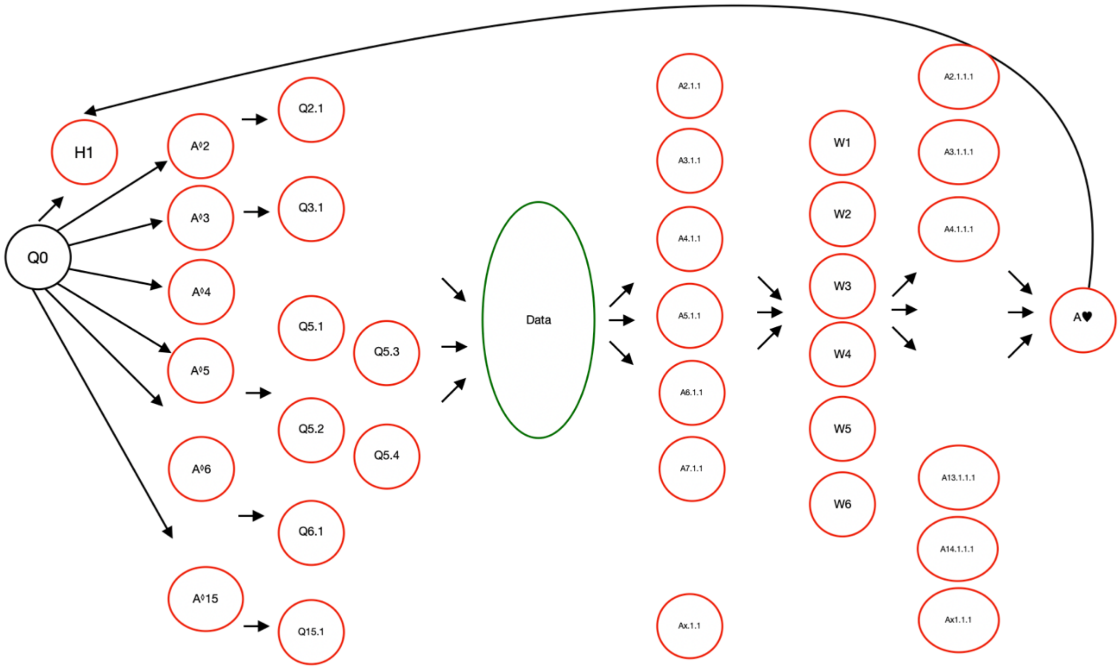


Figure 4. Q&A diagram visualising some of the dialectics of questions and answers. (The teacher's question is black, and students' hypothesis, answers and questions are red).

We now present some of the students' hypotheses and answers found in the media (Figure 3).

Q_0: Are we physically active?

H_1: We are physically active (50% of the students). We are not physically active (50% of the students).

$A^{\Diamond}_2$: 97% of girls in the world invest less than one hour a day in physical activities.

$Q_{2.1}$: How many students (in grade five) play sports?

$A^{\Diamond}_3$: 84.5% of Danish children between 11 and 15 years are active for less than one hour a day – The WHO recommends that children are active one hour a day. The Ministry of Health recommends that children walk 12,000 steps a day.

$Q_{3.1}$: What are the number of steps children walk a day?

$A^{\Diamond}_4$: 8 out of 10 students are not physically active.

$Q_{4.1}$: How many students are physically active in the three classes?

$A^{\Diamond}_5$: How many are physically active?

$Q_{5.1}$: Are boys more active than girls?

$Q_{5.2}$: How much time do girls and boys spend on digital media?

$Q_{5.3}$: What are the daily number of steps for students in the three classes?

Q$_{5.4}$: How much time do students spend on digital media at the three schools?

A$^{\Diamond}_{6}$: 1.6 million children between 11 and 17 years of age have participated in the survey.

Q$_{6.1}$: What is a good question to pose to answer the overall question?

..

A$^{\Diamond}_{15}$: 15-year-old girls are the least active group of children; only 6% of the girls meet the expectations from the WHO.

To explore the questions and develop answers, the students used different resources: the given media; the data collected from students in the three classes; and, finally, TinkerPlot, to analyse and visualise the data (Figure 5).

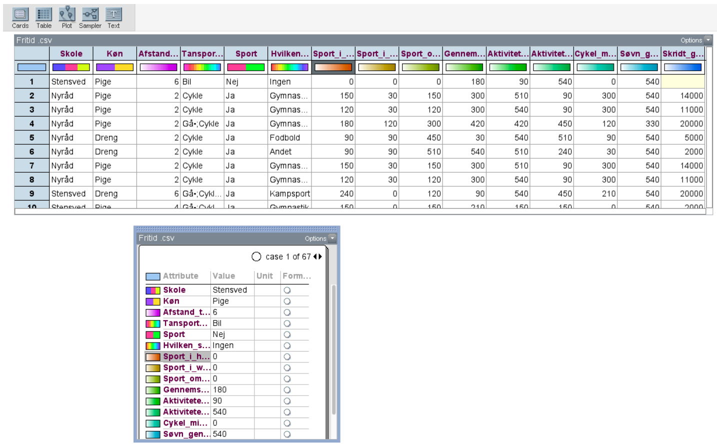


Figure 5. Data in TinkerPlots are represented in a table and data cards. Includes 15 questions, N=67

The data, TinkerPlots and statistical knowledge from both the media just studied and known statistical techniques together formed the essential milieu for the research process to be carried out.

In the SRP, two milieus were established. First, there were *research milieus*, where, in small groups, the students continued to use techniques such as posing questions, sorting data, making dot plots, bar graphs, inquiring types of diagrams, analysing diagrams and interpreting if and how the diagram can give an answer to the derived question and the generative question, Q$_0$. This research activity led to the answers A$_{2.1.1}$…A$_{x.x.x}$, for example, as shown in Figure 6.

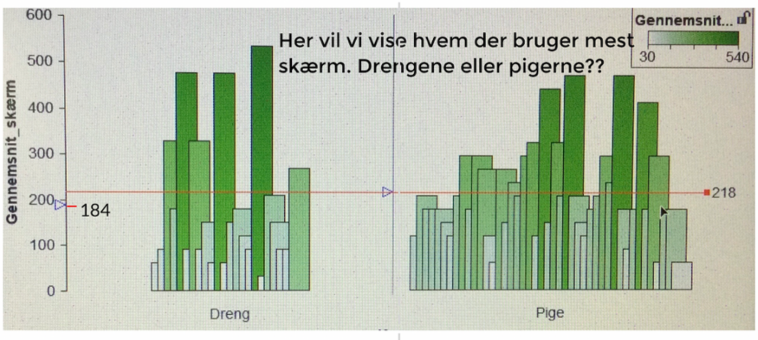


*Figure 6. $Q_{5.2}$: How much time do girls and boys spend on media? $A_{5.2.1}$: Bar graph showing how much time boys (*dreng*) and girls (*pige*) spend on media, including the mean. The students wrote the following: 'It seems like girls spend more time on media. We are surprised because boys play a lot of games, while girls spend time on social media. On the one hand, boys spend 184 minutes (mean) a day on media, and on the other hand, girls spend 218 minutes (mean) a day on media. Girls spend most of their time on media'.*

Second, a *study milieu* took place, in which the students presented their first analysis at whole-class conferences and the students and the teacher together visited some of the questions, answers and techniques used, $w_{w.1} \ldots w_{w.6}$. The two milieus interacted as the students' activity alternated throughout the SRP between the research milieu (where the students performed research on data) and study milieu (where the students together with the teacher studied the statistical models and statistical works used or developed in the group milieu). The dialectic of media and milieu became visible when answers found in media gave rise to new milieus in which the students investigated data and validated answers.

The Q&A diagram, Figure 4, visualises how the students, with their new praxeologies of the studied works, continued to develop more informed answers ($A_{2.1.1.1} \ldots A_{x.2.1.1.1}$, Figure 3) to Q_0, along with how these answers together formed the final answer(s) $A^{\heartsuit}$. At the end of the lesson, the students discussed $A^{\heartsuit}$ in relation to their first informal hypothesis, H_1. In the Q&A diagram, the dialectic of study and research appears indirectly. First, the students researched the media to find answers; second, the students researched their own data; third the students studied works like statistical descriptors to validate the developed answers; and finally, the students' developed new praxeologies to improve their answers. The alternation between research and study became natural because (re)searching for answers calls for a study to validate these answers.

It is worth noting the relatively high autonomy of students' work. In Figure 3, we can see an illustration of how the students posed questions and answered them, here as guided by the teacher. This is in contrast to their regular teaching, in which the teacher posed questions and usually validated the students' answers.

In the following sections, we analyse in depth an episode of the students' study of their works. We use Q&A diagrams to visualise the conferences. At the conferences, the students and teacher visited the works of *frequency and relative frequency*. The works were directly found in the answers developed by the students.

7.2. VISITING THE WORK OF FREQUENCY AND RELATIVE FREQUENCY

With TinkerPlots, the students could produce diagrams of the data and use these to develop their own models for answering a question or exploring informal inferences. The students identified, described and interpreted the data, including the notions of frequency and relative frequency. Some derived questions to Q_0, 'Are we physically active?' included 'how many' (e.g., Figures 7 to 9).

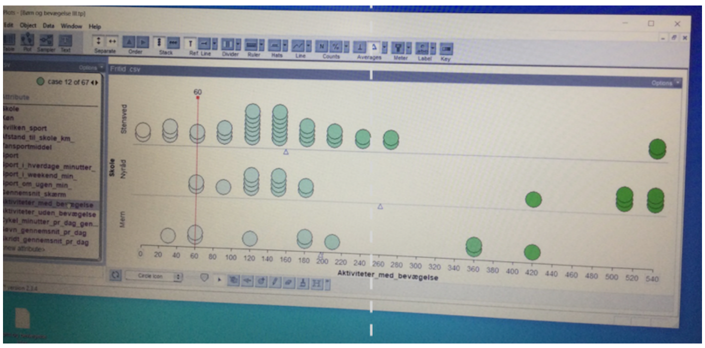


Figure 7. $Q_{4.1}$: How many students are physically active in the three classes? $A_{4.1.1}$: *Frequency graph comparing physical activity minutes per day in the three classes. The students wrote the following: In the diagram, we see the students who are active for more and less than 60 minutes [the red line indicates 60 minutes]. In class C, all the students are active for 60 or more minutes, but in class B, there are five students who are active for less than 60 minutes. In class A, one student is active for less than 60 minutes. Then, class B is the best.*

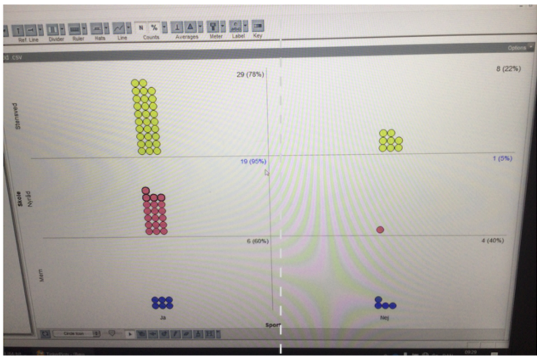


Figure 8. *$Q_{2.1}$: How many students (in grade five) play sports? $A_{2.1.1}$: Diagram comparing participation in sports in the three classes. In the diagram, frequency and relative frequency are visible. The students wrote the following: More students play sports in class B (yellow), but there are also more students in school B. We then found the relative frequency to find the answers. In School C (red), only one student does not play sports; School C (purple) is the school where the most students play sports.*

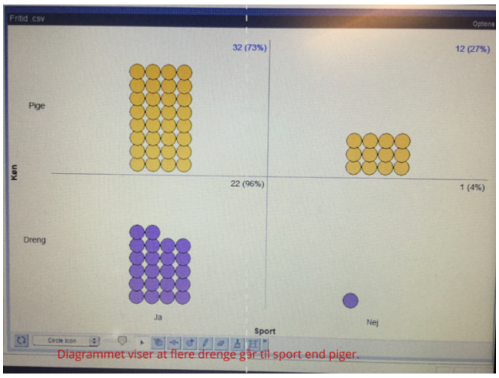


Figure 9. $Q_{5.1}$: Are boys more active than girls? The students wrote the following: $A_{5.1.1}$: The diagram shows that more boys participated in sports activities than girls [96% of the boys to 73% of the girls], here based on a diagram comparing boys (purple) and girls (yellow) and their participation in sports activities.

The instrumented techniques provided by TinkerPlots allowed the students to find frequency and relative frequency along with visual representations of their data. The research process involved separating and stacking observations in TinkerPlots and 'black-box processes' like pressing the button N or % to find the frequency or relative frequency. To understand the statistical answers, the students visited the works, w_3, of frequency and relative frequency.

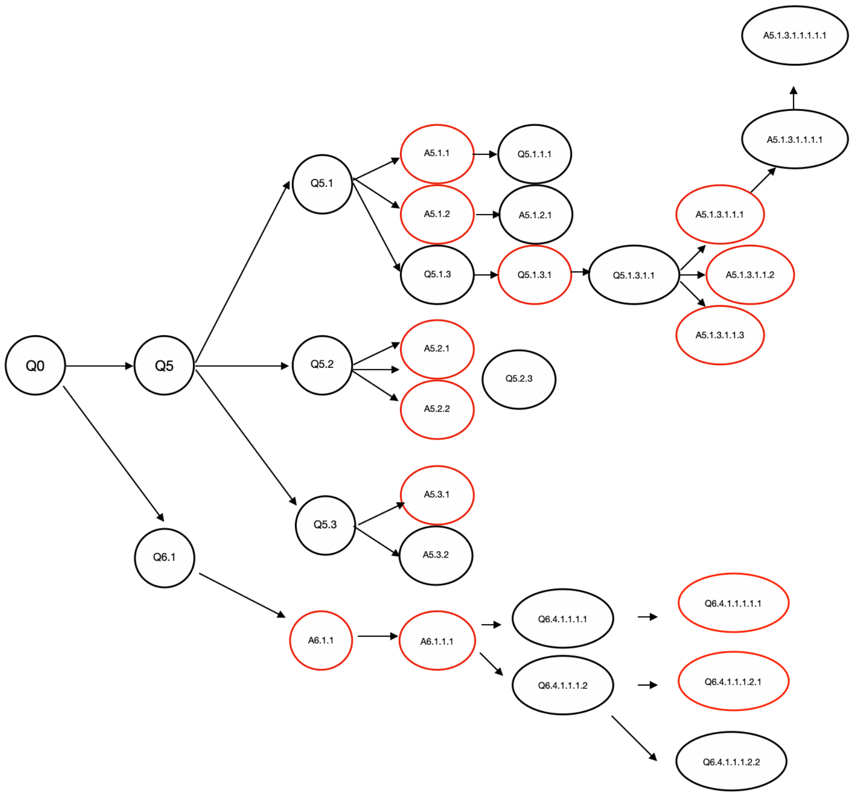


Figure 10. Q&A diagram visualising the visits of frequency and relative frequency (in Figure 3 W3). (The teacher's questions and answers are black, and the students' questions and answers are red)

The following transcription is an unfolding of a visit of work, W3 (Figure 10), and includes students' and teacher's derived questions, preliminary and more established answers.

Q_0: Are we physically active?

Q_5: How many [are physically active]?

A…A (e.g., Figures 6 to 8).

${}^{w3}Q_{5.1}$: Do you know what to call it in 'math language', this question about how many?

${}^{w3}A_{5.1.1}$: Percentage.

${}^{w}Q_{5.1.1.1}$: No.

${}^{w3}A_{5.1.2}$: Frequency.

${}^{w3}A_{5.1.2.1}$: This [frequency] is how many times an observation is shown in a set of data.

${}^{w3}Q_{5.1.3}$: There is also something else … you have all pushed (a button).

${}^{w3}Q_{5.1.3.1}$: Percentage [button].

$^{w3}Q_{5.1.3.1.1}$: What is that [button]? What does % mean? What does 30% mean?

$^{w3}A_{5.1.3.1.1.1}$: The % icon, when you push the % button; then, it will show you how many percentages.

$^{w3}A_{5.1.3.1.1.3.1}$: 37% is not the same as 37 people.

$^{w3}A_{5.1.3.1.1.3.1.1}$: In mathematics, we call it relative frequency (the teacher put up a board explaining the definition of relative frequency and how to calculate relative frequency).

$^{w3}A_{5.1.3.1.1.2}$: How many are there in each class?

$^{w3}A_{5.1.3.1.1.3}$: That there are 37%, 37% of all students are from class C; 37% of the students in the survey.

$^{w3}Q_{5.1.2}$: And what about percentages? What are percentages?

$^{w3}A_{5.1.2.1}$: Out of 100.

$^{w3}A_{5.1.2.2}$: If there were 100 students, 74 students would have 100 Danish kroner.

$^{w3}Q_{5.1.2.3}$: It starts with the letter R (the teacher put up another board on the wall explaining relative frequency).

$^{w3}Q_{5.1.3}$: Do you think it is true if I say that there are more boys than girls playing sports?

$^{w}A_{5.1.3.1}$: No, because you don't know.

$^{w}A_{5.1.3.2}$: I need some data [to answer], and I am not allowed to say, 'I think' or 'I believe'. Instead, I have to write 'in my data I see …'.

$Q_{6.1}$: What is a good question to pose to answer the overall question?

…

$Q_{6.1.1}$: How many students play sports?

$A_{6.1.1.1}$: (The student presenting reads from her notes). In grade five, most students play sports in class B. In class B, 29 students play sports (a new diagram on the white board), and there are eight students who do not play sports. In class C, 19 students play sports, and one student does not play sports. In school A, six students play sports and four do not. In class B, most students play sports, but there are also more students in class. We then calculated it in percentages. The diagram shows that 78% play sports in class B, 95% play sports in class C, and 60% in class A.

$^{w3}Q_{6.1.1.1.1}$: Do you think that class B has most students playing sports?

$^{w3}A_{6.1.1.1.1.1}$ $_{Q6.4.1.1.1.1.1}$: Uhm, it is because the percentage … it is not only about how many students there are. But more like relative frequency. …

$^{w3}Q_{6.1.1.1.2}$: Relative frequency, why? Why is it not possible to compare class A and class B if we do not use relative frequency?

$^{w3}A_{6.1.1.1.2.1}$: Because there are more students in class B.

$^{w3}A_{6.1.1.1.2.2}$: More students in front of the screen (use of digital media), but also more students in the classroom, and you cannot compare the two classes when the classes have a different number of students. But, if you instead calculate the relative frequency and percentage, then you suddenly can compare the two classes. Yes.

The above Q&A diagram (along with an explanation of the questions and answers) shows that visiting work does not entail the generating question, Q_0, being pushed into the background. Q_0 remains the context of students' study and the development of new praxeologies as they visit works.

In the milieu of visiting the notions of frequency and relative frequency, different media are available in the classroom: the students' own graphs and notes from researching the generating question, the teacher and, later on in the study, a board (Figure 1) could explain the definition of frequency as "frequency is how many times an observations is present in a set of data" and how to calculate relative frequency as "relative frequency is frequency divided with the total amount of observations", including visual representations.

Figure 11. Boards with definitions and visualisations of frequency (in Danish hyppighed*) and relative frequency (in Danish* frekvens*)*

Type of task: What is a relevant measure of ‘how many’ in statistics? Here, TinkerPlots can furnish a number of instrumented techniques (see, e.g., Figures 7 to 9), which subsequently can give rise to more theoretical justifications and explanations. At the conference, the students discussed the two different techniques to compare data: “${}^{w3}A_{5.1.2}$: Frequency (obtained by pressing the N-button)” and “${}^{w3}A_{5.1.3.1.1.1}$: The %-icon, when you push the %-button, then it will show you how many percentages” (obtained by pushing the %-button). The teacher further added, “${}^{w3}A_{5.1.3.1.1.3.1.1}$: In mathematics, we call it relative frequency”.

Technology: At the conference, the teacher explicitly pointed out how to compute frequency, that is, ${}^{w3}A_{5.1.2.1}$. At first, there was no explicit definition of relative frequency, except that the teacher linked it to percentage and opposed it against a frequency: “${}^{w3}A_{5.1.3.1.1.3.1}$: 37% is not the same as 37 people”. The students also seemed to realise the difference between frequency and relative frequency: “${}^{w3}A_{6.4.1.1.1}$: Uhm, it is because the percentage … it is not only about how many students there are. But more like relative frequency”. Later, the teacher presented a definition of relative frequency on the board and how to calculate relative frequency, but this was not further explained or discussed with the students.

Theory: The teacher set the stage to make the students reason about statistics, for example, when she asked ‘*why*’ in ${}^{w3}Q_{6.1.1.1.2}$ and then, in ${}^{w}A_{5.1.3.2}$, recommended the students build their justifications and argumentations on data, not mere beliefs. The students in ${}^{w3}A_{6.1.1.1.2.1}$ justified the use of percentage, and at the end of the visit, they explained why they need to know the concept of percentage “${}^{w3}A_{6.4.1.1.1}$: Uhm, it is because the percentage … it is not only about how many students there are. But more like relative frequency”. In the theoretical discussion, some arguments were a mix of technology and theory, for example, when the students use the context of the survey in their arguments: “${}^{w3}A_{6.1.1.1.2.1}$: Because there are more students in class B”. The teacher ended the lesson with the following conclusion: “${}^{w3}A_{6.1.1.1.2.2}$: More students in front of the screen (use of digital media), but also more students in the classroom, and you cannot compare the two classes when the classes have a different number of students. But if you instead calculate the relative frequency and the percentage, then you suddenly can compare the two classes. Yes”.

In the Q&A diagram (Figure 10), we can see how the teacher drove the conversation and posed most of the questions, for example, in ${}^{w3}Q_{5.1}$, ${}^{w3}Q_{5.1.2}$ and $Q_{6.1}$, along with how the teacher validated the students’ answers. The dialectics of questions and answers in the milieu differed from the general view of the whole SRP and the autonomy of the students. The teacher played a key role in the conference and in the development of the students' development of technology and theory.

7.3. VISITING MORE WORKS OF STATISTICS

The students visited six pieces of school statistical works (w1... w6; Figure 3): mode, mean, minimum and maximum value, range, frequency and relative frequency and types of diagrams. Some of the works were known to the students, but their relation to the works developed during the SRP. Other works, such as range and relative frequency, were new to the students and appeared during the study and research process. Both known and new works were important in the students' inquiries into Q_0.

The starting point of a visit of works was the students' or teacher's questioning of the statistics used, e.g., "What is the mean? What is that?". In the discussions, the students explained statistical processes and gave examples like the mean being computed by processes such as "add and divide" and "give and take". The students further explained how a visual representation of the data also visualised the algorithm of the mean (Figure 12, right).

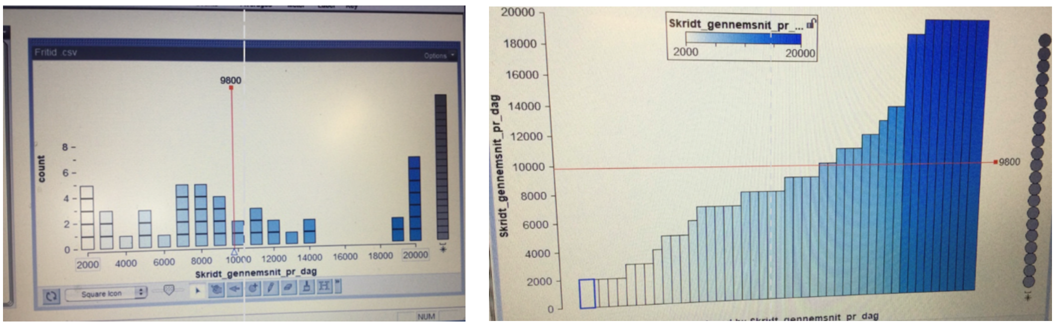


Figure 12. Q: What are the number of steps children walk a day? (Left) Frequency graph showing steps per day. The red line demonstrates the mean. (Right) Bar graph showing students' steps per day. The red line demonstrates the mean. A student explained how the 'long' bars can give to the 'short' bars, which will result in all bars being of the same size.

The theoretical considerations were not just to explain the algorithms or '*how*' to calculate a descriptor. Theory was also used to explain why a conclusion was justified using the given descriptor.

The visits of work were motivated by realisations that certain praxis blocks (involving black-box instrumented techniques) called for an explanation and justification (logos) and represented the dialectic between the students' initial research and subsequent study of works. The visits of works did not stand alone in the SRP but were always initiated by the needs experienced in the study processes to develop or extend new praxeologies and reach the final answer(s) $A^{\heartsuit}$.

8. DISCUSSION AND CONCLUSION

Defining statistical literacy and statistical reasoning in terms of praxeologies makes it possible to not only see the results of statistical analysis, but also those aspects behind these statistical processes. These new definitions of statistical literacy and statistical reasoning can complement former definitions because they offer a direct way to detect them as we analyse statistical activity and discourse.

In the present analysis of the SRP 'Are we physically active?', we see how the process of developing answers has unfolded in many ways, here with a radical change of didactic contract. The students developed statistical literacy as they inquired and elaborated on a real word problem and experienced that statistical techniques and technologies could provide answers that are useful not only in a school context, but that also allow for answering questions of relevance to real-life situations, which are invariably more complex than the statistical problems presented in textbooks.

Furthermore, the students developed statistical reasoning in discussions and elaborations of statistical works, here being motivated by statistical calculations and visual representations and being facilitated by instrumented techniques. The connection of technology and theory enabled the students to make reasoned conclusions based on the data and to explain why a given analysis was valid. This was crucial for statistical reasoning. In our analysis, TinkerPlots supported the students' learning processes; thus, the use went much beyond providing fast ways of processing data into graphical representations. Statistical reasoning appeared in the SRP only when the teacher and students together visited statistical content knowledge. The teacher bridged statistical literacy and statistical reasoning by asking 'how' or 'why' questions, which led the students and teachers to develop technology and theory together. The students' development of statistical techniques, technology and theory led to more complete statistical praxeologies, illustrating why statistical literacy and statistical reasoning cannot be seen as disjoint concepts but instead as mutually supportive and complementary. The dialectic between students' research and study can even be seen as a kind of recursive sequence, in the sense that research (building and drawing on statistical literacy) requires study processes (required by and developing statistical reasoning), which again entails new research and so on. The recursive process is illustrated in Figure 4 in terms of students' alternation between medias and milieus.

Changing from PVM to PQW meets the expectations of an authentic purpose in an authentic situation and nurtures students in understanding, analysing, producing and communicating statistical results. However, authenticity also places the specific SRP 'Are we physically active?' in a broader context. Besides statistical praxeologies, this context also includes broader issues such as open data, fake news and well-being. We claim that working in an unpredictable, maybe even chaotic, and messy context of an SRP can cultivate an adaptive mindset among students, one that will equip the students with the competence to engage in research and study on their own. This is an increasingly crucial competence because "in a rapidly changing environment, the requirement for particular technological skills may change or evolve, whereas the core competencies will not" (Damouras et al., 2021, p. 896).

9. REFERENCES

Anesen, L., & Winth, N. (2010). *Format 5, elevbog*. Alinea.

Bakker, A. (2018). What is design research in education? 1. In *Design research in education* (pp. 3-22). Routledge.

Barquero, B., & Bosch, M. (2015). Didactic engineering as a research methodology: From fundamental situations to study and research paths. In A. Watson & M. Ohtani (Eds.), *Task design in mathematics education: An ICMI study 22* (pp. 249–272). Springer. https://doi.org/10.1007/978-3-319-09629-2_8

Barquero, B., Bosch, M., & Gascón, J. (2013). The ecological dimension in the teaching of mathematical modelling at the university. *Recherches en Didactique des Mathématiques*, *33*(3), 307–338

Barquero, B., Bosch, M., & Romo, A. (2018). Mathematical modelling in teacher education: Dealing with institutional constraints. *ZDM*, *50*(1), 31–43. https://doi.org/10.1007/s11858-017-0907-z

Barquero, B., Monreal, N., Ruiz-Munzón, N., & Serrano, L. (2018). Linking transmission with inquiry at university level through study and research paths: The case of forecasting Facebook user growth. *International Journal of Research in Undergraduate Mathematics Education*, *4*(1), 8–22. https://doi.org/10.1007/s40753-017-0067-0

Ben-Zvi, D., & Garfield, J. B. (2004). Statistical Literacy, Reasoning, and Thinking: Goals, Definitions, and Challenges. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 3-16). Kluwer Academic Publishers. https://doi.org/10.1007/1-4020-2278-6_1

Ben-Zvi, D., Gravemeijer, K., & Ainley, J. (2018). Design of statistics learning environments. In *International handbook of research in statistics education* (pp. 473–502). Springer. https://doi.org/10.1007/978-3-319-66195-7_16

Bosch, M. (2018). Study and research paths: A model for inquiry. In B. Sirakov, P. N. De Souza, & M. Viana (Eds.), *Proceedings of the International Congress of Mathematicians 2018* (Vol. 3, pp. 4001–4022). World Scientific. https://doi.org/10.1142/9789813272880_0210

Bosch, M., & Gascón, J. (2014). Introduction to the anthropological theory of the didactic (ATD). In A. Bikner-Ahsbahs & S. Prediger (Eds.), *Networking of theories as a research practice in mathematics education* (pp. 67–83). Springer. https://doi.org/10.1007/978-3-319-05389-9_5,

Bosch, M., & Ruiz-Munzón, N (2022). The curriculum problem and the paradigm of questioning the world. In Y. Chevallard, B. Barquero, M. Bosch, I. Florensa, J. Gascón, P. Nicolás, & N. Ruiz Munzón (Eds.), *Advances in the anthropological theory of the didactic* (pp. 173–177). Birkhäuser. https://doi.org/10.1007/978-3-030-76791-4

Bosch, M., & Winsløw, C. (2015). Linking problem solving and learning contents: The challenge of self-sustained study and research processes. *Recherches En Didactique Des Mathématiques*, *35*(3), 357–399.

Brousseau, G. (1997). *The theory of didactical situations in mathematics*. *Didactique des Mathématiques 1970–1990.* (N. Balacheff, M. Cooper, R. Sutherland, & V. Warfield, Eds. and Trans.). Kluwer. https://doi.org/10.1007/0-306-47211-2

Chevallard, Y. (1992). Fundamental concepts in didactics: Perspectives provided by an anthropological approach. In R. Douady & A. Mercier (Eds.), *Research in didactique of mathematics. Selected papers* (pp. 131–167). La Pensée Sauvage.

Chevallard, Y. (1999). El análisis de las prácticas docentes en la teoría antropológica de lo didáctico. *Recherches en didactique des mathématiques*, *19*(2), 221–266.

Chevallard, Y. (2002). Organiser l'étude 3. Ecologie et regulations. In J.-LM. Dorier, M. M. Artaud, R. M. Artigue, R. Berthelot & R. Floris (Eds.), *Actes de la 11e École d'Été de Didactique des Mathématiques* (pp. 41–56). La Pensée Sauvage

Chevallard, Y. (2006). Steps towards a new epistemology in mathematics education. In M. Bosch (Ed.), *Proceedings of the IVth Congress of the European Society for Research in Mathematics Education* (pp. 22–30). Fundemi IQS.

Chevallard, Y. (2007). Readjusting didactics to a changing epistemology. *European Educational Research Journal, 6*(2), 131–134.

Chevallard, Y. (2012). *Teaching mathematics in tomorrow's society: A case for an oncoming counterparadigm*. Plenary talk at the 12th International Congress on Mathematical Education, Seoul, Korea.

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

Chevallard, Y. (2019). Introducing the anthropological theory of the didactic: An attempt at a principled approach. *Hiroshima Journal of Mathematics Education*, *12*, 71–114.

Chevallard, Y. (2022) Challenges and advances in teacher education within the ATD. In Y. Chevallard, B. Barquero, M. Bosch, I. Florensa, J. Gascón, P. Nicolás, & N. Ruiz Munzón (Eds.), *Advances in the anthropological theory of the didactic* (pp. 81–89) Birkhäuser. https://doi.org/10.1007/978-3-030-76791-4

Cobb, P., Jackson, K., & Dunlap, C. (2017). Conducting design studies to investigate and support mathematics students' and teachers' learning. *Compendium for research in mathematics education*, 208–233

Damouras, S., Gibbs, A., & MacFeely, S. (2021). Training official statisticians for adaptive statistical practice. *Statistical Journal of the IAOS*, *37*(3), 887–898. https://doi.org/10.3233/SJI-210851

Danish Health Authority. (2005). *Fakta- børn og bevægelse* (*Facts - children and movement).* https://www.sst.dk/~/media/41F8AF24985D414097A27B16E0FA634B.ashx

Danish Ministry of Education. (2006). *Folkeskoleloven (Public school act). https://www.uvm.dk/folkeskolen/folkeskolens-maal-love-og-regler/om-folkeskolen-og-folkeskolens-formaal/folkeskolens-formaal*

Danish Ministry of Education. (2019). *Matematik - fælles mål (Goals of mathematics).* https://emu.dk/sites/default/files/2020-09/GSK_FællesMål_Matematik.pdf

delMas, R.C. (2002). Statistical literacy, reasoning, and thinking: A commentary. *Journal of Statistics Education*, *10*(2), 1–11. https://doi.org/10.1080/10691898.2002.11910674

Dreyfus, T. (1994). The role of cognitive tools in mathematics education. In *Didactics of mathematics as a scientific discipline* (pp. 201–211). Kluwer Academic Publishers.

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). *Guidelines for assessment and instruction in statistics education (GAISE) report: A pre-K–12 curriculum framework.* American Statistical Association.

Fitzallen, N., & Watson, J. (2010, July). Developing statistical reasoning facilitated by TinkerPlots. In *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics.* (ICOTS8).

Fujii, T. (2016). Designing and adapting tasks in lesson planning: A critical process of lesson study. *ZDM Mathematics Education*, 48, 411–423. https://doi.org/10.1007/s11858-016-0770-3

Gal, I. (2002). Adults' statistical literacy: Meanings, components, esponsibilities. *International Statistical Review*, *70*(1), 1–25 https://doi.org/10.1111/j.1751-5823.2002.tb00336.x

Garfield, J. (2002). The challenge of developing statistical reasoning. *Journal of Statistics Education*, *10*(3), 1–12. https://doi.org/10.1080/10691898.2002.11910676

Gil, E., & Ben-Zvi, D. (2011). Explanations and context in the emergence of students' informal inferential reasoning. *Mathematical Thinking and Learning*, *13*, 87–108. https://doi.org/10.1080/10986065.2011.538295

Jessen, B. E. (2014). How can study and research paths contribute to the teaching of mathematics in an interdisciplinary setting? *Annales de Dedactique et de Science Cognitives*, *19*, 199–244.

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: The cases of

Denmark and Japan. In M. Bosch, Y. Chevallard, F. García, & J. Monaghan (Eds.), *Working with the anthropological theory of the didactic in mathematics education* (pp. 118–138). Routledge. https://doi.org/10.4324/9780429198168

Kidron, I., Artigue, M., Bosch, M., Dreyfus, T., & Haspekian, M. (2014). Context, milieu, and media-milieus dialectic: A case study on networking of AiC, TDS, and ATD. In A. Bikner-Ahsbahs & S. Prediger (Eds.), *Networking of theories as a research practice in mathematics education* (pp. 153–177). Springer. https://doi.org/10.1007/978-3-319-05389-9_10

Konold, C., & Higgins, T. L. (2003). Reasoning about data. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 193–215). National Council of Teachers of Mathematics.

Larsen, D. M., Østergaard, C. H., & Rasmussen, K. (2022). Contradicting activity systems – Learning from large-scale interventions that fail to change mathematics teaching practice as intended. *Journal of Mathematics Teacher Education*, 1–24. https://doi.org/10.1007/s10857-022-09548-1

Langrall, C., Mooney, E., Nisbet, S., & Jones, G. (2008). Elementary students' access to powerful mathematical ideas. In L. English (Ed.), *Handbook of international research in mathematics education* (pp. 109–135). Routledge.

Lagrange, J.-B. (2005). Using symbolic calculators to study mathematics: The case of tasks and techniques. In D. Guin & L. Trouche (Eds.), *The didactical challenge of symbolic calculators* (pp. 113–135). Springer. https://doi.org/10.1007/0-387-23435-7_6

Makar, K., & Rubin, A. (2009). A framework for thinking about informal statistical inference. *Statistics Education Research Journal*, *8*, 82–105

Meletiou-Mavrotheris, M., & Paparistodemou, E. (2015). Developing students' reasoning about samples and sampling in the context of informal inferences. *Educational Studies in Mathematics*, *88*(3), 385–404. https://doi.org/10.1007/s10649-014-9551-5

Nicholson, J., Gal, I., & Ridgway, J. (2018). *Understanding civic statistics: A conceptual framework and its educational applications. A product of the ProCivicStat project*. https://iase-web.org/islp/pcs/documents/Conceptual_framework_brief.pdf

Noll, J., & Kirin, D. (2017). TinkerPlots ™ model construction approaches for comparing two groups: Student perspectives. *Statistics Education Research Journal*, *16*(2), 213–234. http://iase-web.org/Publications.php?p=SERJ

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), 1–17. https://doi.org/10.30935/scimath/12401

Pratt, D., Davies, N., & Connor, D. (2011). The role of technology in teaching and learning statistics. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics-challenges for teaching and teacher education: A joint ICMI/IASE study,13* (pp. 97–107). Springer. http://doi.org.10.1007/978-94-007-1131-0_

Rasmussen, K. (2016). The direction and autonomy of interdisciplinary study and research paths in teacher education. *REDIMAT*, *5*(2), 158–179. http://dx.doi.org/10.4471/redimat.2016.1753

Sabbag, A., Garfield, J., & Zieffler, A. (2018). Assessing statistical literacy and statistical reasoning: The REALI instrument. *Statistics Education Research Journal*, *17*(2), 141–160. https://doi.org/10.52041/serj.v17i2.163

Samvirke. (2014). *How to easily walk 10,000 steps a day* (Sådan går du nemt 10.000 skridt om dagen). https://samvirke.dk/artikler/saadan-gaar-du-nemt-10000-skridt-om-dagen

Strømskag, H., & Chevallard, Y. (2022). Breaches of the didactic contract as a driving force behind learning and non-learning: A story of flaws and wants. *Teaching Mathematics and its Applications: An International Journal of the IMA*, *0*, 1–13. https://doi.org/10.1093/teamat/hrac003

Tukey, J. W. (1977). *Exploratory data analysis vol 2.* Addison-Wesley Publishing Company.

Wallmann, K. K. (1993). Enhancing statistical literacy: Enriching our society. *Journal of the American Statistical Association, 88*(421), 1–8. https://doi.org/10.1080/01621459.1993.10594283

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, *67*(3), 223–248. https://doi.org/10.1111/j.1751-5823.1999.tb00442.x

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. https://doi.org/10.1007/s10649-012-9453-3

World Health Organization. (2020). *Physical activity*. https://www.who.int/initiatives/behealthy/physical-activity

Yin, R. K. (2014). *Case study research: Design and methods* (5th ed.). Sage Publication.