

ESTABLISHING SOCIOMATHEMATICAL NORMS IN STATISTICS TEACHING IN LOWER-SECONDARY CLASSROOMS

STINE GERSTER JOHANSEN
University College Copenhagen
stgj@kp.dk

CHARLOTTE KROG SKOTT
University College Absalon
chks@pha.dk

ABSTRACT

Because digital technologies and collection of data about our lives are omnipresent in our society, statistics has become an increasingly important school subject. In this article, we investigate the development of statistical literacy in a class of middle school students over the course of a one-year case study. Previously, the students tended to work procedurally with statistics, but during the study, they increasingly engaged in questioning statistical information, justifying statistical interpretations, and communicating data-based claims about issues within and beyond their school context. We take a social constructivist perspective and analyse which sociomathematical norms related to statistical literacy emerged during the classroom interactions. We identified two such norms—what counts as a valuable description of a data set and what counts as a valuable answer to a statistical investigation—that connected classroom interactions with the aim of supporting students' development of statistical literacy, especially its dispositional elements. The study contributes a rare empirical exploration of how to support students' development of statistical literacy at the classroom level.

Keywords: *Statistics education; Statistical literacy; Sociomathematical norms; Classroom research*

1. INTRODUCTION

Because digital technologies permeate our lives—social, private, and professional—our need to understand their use and impact increases. The number of technologies that collect quantitative data about us and the larger society, including economic and environmental issues, is increasing, as is the total amount of data collected. Data collected from the websites we visit are often used to target individuals with tailored campaigns, including political campaigns, and to communicate messages in ways that may obscure important patterns and trends in the data. It is more pressing than ever for schools to ensure that students learn about statistics, specifically to collect, analyse, represent, interpret, and question data.

Statistics has played an increasingly important role in mathematics curricula around the world in the last two decades, but there is a tendency in countries such as the United States (Visnovska & Cobb, 2019) and Denmark (e.g., Skott & Østergaard, 2018) to teach the subject with a focus on introducing concepts (e.g., the mean) and practising procedures using existing data sets. For example, students calculate the mean as part of a routinised approach to data analysis, without attending to specific data or the context of the statistical investigation. However, learning such routinised and detached procedures does not help students to think critically about statistics (Visnovska & Cobb, 2019) and to become statistically literate. Being statistically literate means, for instance, to understand and interpret quantitative studies encountered in one's everyday life and world, to critically evaluate if data may tell stories other than the one presented, and to assess whether the story told has additional important nuances or is based on insufficient assumptions (Bargagliotti et al., 2020; Gal, 2002, 2021). Statistical literacy goes beyond simply learning statistical

concepts and procedures and represents a broader educational goal: to prepare students to understand and participate in public debates and our democracy, more generally.

The present study aims to foster this broader goal of students becoming statistically literate. The study is a local-knowledge case study (Thomas, 2011) and takes place in an ordinary middle-school class in the Copenhagen area in Denmark. Previously, the class worked with statistics in a mostly procedural manner. However, after the one-year case study, many students showed signs of statistical literacy, such as questioning data and justifying their interpretations. To examine the class's evolution, we used a social constructivist perspective and focused on *sociomathematical norms* (Yackel & Cobb, 1996), that is, norms that are “specific to the mathematical aspects of students’ activity” (p. 458). We did so to capture the multifaceted concept of statistical literacy, especially its dispositional elements, understood here as students’ capacity and willingness to question statistical information, justify statistical interpretations, and communicate data-based claims in meaningful contexts. We address the research question: *Which sociomathematical norms that promote statistical literacy do the teacher and students establish in a statistics middle-school classroom?*

The study aims to contribute an exploration of how to support students’ development of statistical literacy, especially its dispositional elements, and to do so by focusing on classroom interactions between teacher and students, especially how they negotiate and develop sociomathematical norms.

We start by elaborating our approach to statistical literacy and positioning the study. Then, we describe our theoretical and methodological approaches. In the result section, we show how the teacher and student develop two productive sociomathematical norms, and then we discuss these results. We conclude by highlighting the study’s practical implications.

2. THEORETICAL PERSPECTIVES

2.1. STATISTICAL LITERACY

Our approach to teaching statistics and designing teaching units is grounded in the broader goal of developing students’ statistical literacy. Gal (2002) conceptualised statistical literacy as comprising two key components: knowledge elements and dispositional elements. The knowledge elements include literacy skills, statistical and mathematical knowledge, contextual understanding, and the ability to pose critical questions. Dispositional elements refer to inclinations to apply this knowledge, particularly in uncertain situations, and belief in the legitimacy of questioning statistical claims, even without formal expertise. In this study, we use the term *dispositional elements* to refer especially to students’ capacity and willingness to question statistical information, justify statistical interpretations, and communicate data-based claims in meaningful contexts so that students critically engage with data in meaningful ways (Gal, 2002, 2021). Statistical literacy is closely related to statistical reasoning and statistical thinking, but the three constructs are not identical. Ben-Zvi and Garfield (2004) characterised statistical literacy as involving foundational skills for understanding statistical information, statistical reasoning as making sense of statistical ideas and information, and statistical thinking as understanding the broader logic of statistical investigations. In this study, we build on Gal’s (2002, 2021) broader conceptualisation of statistical literacy, in which knowledge elements such as reasoning with data, distributions, and descriptors are necessary, but not sufficient, for statistical literacy. Statistical literacy also involves dispositional elements.

In line with Gal’s conceptualisation, the Guidelines for Assessment and Instruction in Statistics Education II (GAISE II) report (Bargagliotti et al., 2020) identified statistical literacy as the ultimate goal of statistics education across school levels. The report advocated for the early and continuous integration of statistical experiences and underscored the importance of inquiry, including posing questions, analysing data, and interpreting results in context. A core feature of statistical literacy, as articulated in GAISE II, is that students should not simply learn statistical procedures but be supported in becoming critical and informed consumers and producers of data-based arguments.

Weiland (2017, 2019) expanded considerations of statistical literacy through the notion of critical statistical literacy, distinguishing between reading the world with statistics by interpreting and critiquing

statistical information provided by others and writing the world with statistics by producing, analysing, and acting on statistical insights. Building on critical literacy traditions, Weiland positioned statistical literacy as both cognitive and ethical, emphasising that engaging with data entails taking a stance toward the social world. In his later work, Weiland (2019) further clarified that “critical” should not be understood only as disciplinary critical thinking, but also as interrogating assumptions, structures, and uses of statistics in society.

Complementing this perspective, the framework of *civic statistics* (Ridgway, 2022) extends statistical literacy to the societal level by focusing on how citizens interpret and critique complex data about public issues. Gal (2022) emphasised that developing students’ critical understanding of civic statistics requires engagement with meaningful and important contexts, often through statistical texts such as media articles, official reports, and other public data-based messages. In such contexts, students need to interpret statistical information, evaluate what is shown or omitted, consider whether claims are supported by data, and formulate justified interpretations or opinions. This line of research highlights the importance of rich, issue-based contexts that allow students to use statistical information as a basis for interpreting, critiquing, and discussing public phenomena.

In this study, we view statistical literacy as a multidimensional competence that unites knowledge and disposition and entails students’ capacity and willingness to question statistical information, justify statistical interpretations, and communicate data-based claims about issues that matter, both within and beyond school. This view of statistical literacy aligns with Gal’s (2002, 2021, 2022) emphasis on knowledge elements, dispositional elements, contextual understanding, and critical evaluation, including the interpretation of statistical information in meaningful civic contexts. It also aligns with Weiland’s (2017, 2019) idea of critical statistical literacy, which combines reading and writing the world with statistics.

Despite expansive conceptualisations, statistical literacy has proved elusive in practice. Nineteen years after his seminal exposition on statistical literacy, Gal (2021) expressed concern that the notion of statistical literacy is still often reduced to basic knowledge and is not considered as a multi-dimensional competency. Consistent with this concern, Callingham and Watson (2017) found that although students deepened their procedural and conceptual knowledge as they advanced through their middle school years, few reached the higher levels of critical reflection. Students’ progression beyond procedural understanding toward critical reflection therefore appears difficult to support in practice (Callingham & Watson, 2017; Gal, 2021). Critical reflection, in which students draw on dispositional elements of statistical literacy, is seemingly particularly challenging to foster, suggesting a need for more knowledge about the classroom processes through which students may come to question, justify, and communicate about statistical information.

2.2. STUDENTS’ DEVELOPMENT OF STATISTICAL LITERACY

Empirical research examining how statistical literacy develops in authentic classroom contexts is limited. Büscher (2018, 2022, 2024) conducted design-based classroom studies that informed how tasks, tools, and contexts can elicit reflective engagement. In one teaching unit, students analysed sea ice in the Arctic using authentic climate data and a dynamic data-analysis environment. The task juxtaposed conflicting evidence from different years and sources, prompting students to question data selection and interpretation. Across his studies, Büscher demonstrated how critical contexts, purposefully incomplete information, and interactive tools can invite imaginative and selective readings of statistical information. His work illustrates what designs can do to stimulate reflection, but such classroom-based studies remain rare.

Beyond Büscher’s small cluster of studies, several quantitative and assessment-based studies examined students’ statistical literacy in broader educational settings. For example, Callingham and Watson (2017) analysed large-scale assessment data to model students’ statistical-literacy levels across the middle years of schooling. Sabbag, Garfield, and Zieffler (2018) developed and validated the Reasoning and Literacy Instrument to assess students’ statistical literacy and reasoning, and Ziegler and Garfield (2018) introduced the Basic Literacy in Statistics assessment to measure foundational skills for statistical literacy. More recently, Phadke, Beckman, and Morgan (2024) investigated how the contextual realism of assessment

tasks affected students' performance on isomorphic items. Together, these studies extended empirical knowledge about learners' competencies but primarily relied on tests or survey instruments to measure statistical literacy rather than on observations of classroom activity to inform how students develop statistical literacy.

Design-based research studies in statistics education have focused on students' statistical reasoning and how it unfolded in classroom interactions. For example, Cobb and McClain (2005) documented how social and sociomathematical norms evolved in a seventh-grade classroom as students learned to analyse and argue with data using dynamic technology. A key sociomathematical norm that emerged concerned what counted as an acceptable data-based argument because students were expected to justify why their method of structuring the data was relevant to the question under investigation. Similarly, Ben-Zvi and colleagues (2007) traced how students working with TinkerPlots collaboratively built representations and arguments, showing how discourse and tool use supported their development of statistical reasoning. These studies provided detailed insight into the interactional processes through which students learned to reason with data. Although these researchers' conceptualisations of statistical reasoning shared many aspects of statistical literacy as conceptualised for this study, the studies and their outcomes were not framed explicitly in those terms.

These bodies of work, ranging from design-based studies of classroom tasks to civic-statistics approaches to public contexts, share a limitation. Although they provide rich accounts of tasks, tools, and contexts, they provide less insight into how the dispositional dimension of statistical literacy may be supported through the social life of the classroom. In particular, more knowledge is needed about the classroom processes through which students may come to question, justify, and communicate statistical interpretations. Here, the work of Cobb and colleagues (Cobb, 1999; Yackel & Cobb, 1996) is inspiring, especially their concept of *intellectual autonomy*. It concerns students' capacity to draw on their own mathematical capabilities when making mathematical decisions and judgments, rather than relying primarily on the authority of the teacher, to participate as autonomous members of an inquiry-oriented mathematics classroom. However, intellectual autonomy is not equivalent to the dispositional dimension of statistical literacy because the two constructs differ in scope: intellectual autonomy is rooted in classroom participation, whereas statistical literacy extends toward civic and societal contexts. Yet both share a commitment to fostering independent, evidence-based reasoning, and for this reason, we will use the same analytical lens as Yackel and Cobb (1996), that is, we use sociomathematical norms in our analysis of classroom interactions.

2.3. SOCIOMATHEMATICAL NORMS

Mathematics classrooms appear to possess a life of their own. Over time, students and teachers develop stable habits and customs to manage classroom complexities, such as, for instance, when teachers unintentionally reproduce routines from their own schooling (Voigt, 1998). To analyse classroom life, researchers have focused on patterns in the ways participants interact with each other, that is, the norms that are often unconsciously negotiated and established as students and their teacher interact. These researchers viewed a classroom and its microculture as a dynamic system that is continuously constituted by the participants' negotiation of norms and sense-making related to the subject matter (e.g., Voigt, 1998).

In alignment with this view, Cobb and his colleagues (2001) suggested a social constructivist approach in which a classroom's microculture was viewed from two perspectives: a social perspective encompassing the normative and collective processes within the classroom and a psychological perspective focusing on individual students' activities, beliefs, and values as they engaged in social processes (Cobb et al., 2001). The two perspectives are interrelated and are "so fully integrated in each other that none of them can be understood without considering the other" (Skott, 2008, p.53). Specifically, the social processes within a classroom are established through interactions and negotiations between the teacher and their students, and their individual ways of contributing to these interactions. At the same time, these processes significantly influence what is possible to learn in the classroom. Consequently, students' mathematical development is

viewed as an emerging phenomenon, contingent upon the social processes established within the classroom as well as the participants' traditions, beliefs, and values regarding mathematics.

Cobb et al. (2001) identified three important aspects of the social perspective: social norms, sociomathematical norms, and mathematical practices. Social norms refer to general classroom norms that are not unique to mathematics but are applicable across various subjects. Examples of social norms are that students will explain and justify their thinking and attempt to comprehend others' explanations (Cobb, 1999). Sociomathematical norms are related to the mathematical dimensions of students' activities. Examples include what counts as an acceptable statistical explanation and what constitutes a different mathematical solution. Social and sociomathematical norms are closely connected, as exemplified by Yackel and Cobb (1996): a social norm encompasses the understanding that students will explain their solutions and thinking, whereas a sociomathematical norm concerns the understanding of why a mathematical explanation is acceptable. Consequently, sociomathematical norms focus on the criteria that govern mathematical activity and discourse, including the criteria for acceptable explanations (Cobb, 1999). The third aspect, mathematical practice, represents a normative way of acting that is interactively established by the teacher and the students. These actions, such as measuring by iterating a foot length, are accepted or taken as shared by most students as a way of working with a specific mathematical topic (Cobb et al., 2001). Each of these three social aspects has a corresponding psychological counterpart. For example, the psychological counterpart to sociomathematical norms consists of students' individual mathematical beliefs and values.

For our purpose, sociomathematical norms offer promise as an analytical lens. One reason is that statistical literacy is a concept too broad and elusive to be divided into specific mathematical practices that a class should learn. Another reason is that productive sociomathematical norms established within a statistical classroom will influence the opportunities that emerge for students to develop statistical literacy. We base this assumption on a number of studies that have shown that intentionally productive sociomathematical norms, such as norms that foster inquiry-based mathematics teaching, can improve students' learning (van Zoest et al., 2012), for instance, in primary education (Lopez & Allal, 2007), secondary education (Widjaja, 2012), university education (Yackel, 2002), teacher education (van Zoest et al., 2012), and professional development (Clark et al., 2008).

In fostering productive sociomathematical norms, the teacher plays a decisive role (Yackel, 2002; Yackel & Cobb, 1996). For example, a teacher can legitimise certain aspects of students' mathematical activity by facilitating classroom discussions that capitalise and build on student contributions to achieve the learning goals, such as accentuating one student's explanation or certain aspects of it that align with goals and thereby implicitly downplaying other explanations that may not align with goals (Yackel & Cobb, 1996). Consequently, the teacher's way of contributing to the classroom interactions was an important part of our analysis.

3. METHODOLOGY

3.1. A LOCAL-KNOWLEDGE CASE

To address our research question, we conducted a case study, which is part of the first author's PhD project. We selected our case according to Thomas's (2011) case-study typologies and employed what he calls a local-knowledge case. This case type was chosen because of the first author's intimate knowledge of the case setting, which allowed an informed, in-depth analysis. Furthermore, the case can be viewed as interesting—although not necessarily representative—because it presented a unique opportunity to study students' development of statistical literacy, especially the dispositional elements, in an ordinary school setting.

Thomas (2011) explained that the anchor point of case selection is the dynamic of the relationship between a study's subject and object. The subject is a practical, historical unit, whereas the object is an analytical frame within which the case is viewed (Thomas, 2011). In our study, the subject was a teacher teaching statistics to her Year 6 class (ages 12–13) by using three teaching modules (designed by the teacher

in collaboration with the first author) over the course of one year. The object was sociomathematical norms serving as an analytical lens to explore if and how the teacher and students negotiated and developed norms that promoted students' development of statistical literacy.

3.2. DATA GENERATION

The first author acted as a participating observer while the teacher taught the three modules. Although present during classroom discussions, the first author mostly took field notes and refrained from participating in these discussions. During group work, she answered students' questions and entered discussions.

For analysis, we first selected lessons in which extended classroom discussions occurred as our primary data. This was a deliberate methodological choice. Although teachers and students would also negotiate sociomathematical norms during group-work interactions or during other classroom activities, they negotiate norms socially and publicly in classroom discussions. There were seven such lessons, which were video recorded (totalling 2 hr 36 min) using a wide-angle camera that captured the electronic board, the blackboard, the teacher, and the students. We then watched the classroom videos and selected three episodes that comprised 1 hr 31 min of video (Episodes A, B, and C, see Figure 1) for analysis. Each episode corresponded to a module theme, and student activity in the selected episodes was more extensive than in the other episodes, particularly in terms of students questioning statistical information, justifying statistical interpretations, and communicating data-based claims. Each episode consisted of a joint launch of a statistical activity and one or two follow-up classroom discussions. The video recordings were transcribed prior to analysis using pseudonyms for students' names.

In addition to the videos and transcriptions of the episodes, student work was a source of data, consisting primarily of letters to the editor—letters that were written and shared in class but not submitted to or published in a newspaper. These letters were considered only if students read them aloud during classroom discussions.

3.3. ANALYTICAL STRATEGY

We conducted a two-step analysis, described in detail below. First, we identified which sociomathematical norms, if any, emerged as most prominent during classroom interactions. Because these norms and their criteria were primarily initiated by the teacher in the first modules, our second step focused on students. We examined whether students interacted (or began to interact) in alignment with the norms to determine if the norms had truly been constituted.

In both steps, we analysed the data independently and then compared our findings. There was a high degree of consistency between our individual analyses. We resolved any discrepancies through discussion until we reached consensus to ensure a high level of analytical reliability.

In the first step, we were inspired by coding procedures from grounded theory (Charmaz, 2014). We independently analysed the transcripts of the three episodes line by line by constructing initial codes. We focused on criteria for statistical activities in a broad sense. This broad focus was necessary because, in our conception, the dispositional elements of statistical literacy do not appear separately from students' statistical work. Students' capacity and willingness to question statistical information, justify statistical interpretations, and communicate data-based claims presuppose that they are also exploring data, reasoning with descriptors and distributions, and interpreting these in relation to context. To construct the initial codes, we used gerunds to capture these process aspects of teacher–student interactions. Our focus was both on what the teacher and students explicitly said and how the teacher guided classroom discussions by accepting certain student contributions (and not others) and accentuating specific aspects of their contributions (and not others). We constructed focused codes by comparing these initial codes and assembling those which conveyed criteria—possibly evolving ones—for the same statistical activity. From these codes, we identified two sociomathematical norms that were most prominent across the three teaching modules: what

counts as a valuable description of a data set, and what counts as a valuable answer to a statistical investigation.

We provide an example of our analytical identification of each norm. In Episode A, the teacher asked, “Has anyone made considerations that weren’t about the mean, but perhaps about something else ... other than the mean?” From this question, we constructed the initial code ‘Encouraging the use of different descriptors.’ In Episode C, the teacher asked, “Which descriptors do you think are interesting?” to which Emma answered, “If you compared the mean and the median it could be interesting.” From this interaction, we constructed the initial code ‘Suggesting comparison of descriptors.’ By comparing such codes, we realised that these two codes and others conveyed criteria for what became an acceptable description of a data set in the class, and we constructed the focused code ‘Criteria for a valuable statistical description.’ This focused code and its associated initial codes constituted the foundation for our description of the first norm of what counts as a valuable description of a data set (see section 4 for a detailed description).

In a similar way, we constructed the initial code ‘Encouraging contextual interpretation of descriptors’ for an interaction between John and the teacher. John observed, “[the average] for mom was 46.3 and for dad 47.1,” to which the teacher responded, “What do you think? Did you expect this: that the fathers would be older than the mothers?” (Episode A). Later, the teacher asked, “In the rich parishes, there are big differences between the median and the mean. What does this mean? How is the population composition in these areas?” (Episode C). We coded this initially as ‘Encouraging comparison of contextual interpretations of different descriptors.’ As described above, we then assembled these and other initial codes, conveying criteria for the same activity, under the focused code ‘Criteria for a valuable statistical answer.’ Together, they constituted the foundation for our description of the second norm: what counts as a valuable answer to a statistical investigation (see section 4 for a detailed description).

In the second step, we revisited our data to analyse if the students began to interact in line with the two norms across the three teaching modules. We focused on whether the students interacted independently without being prompted or encouraged by the teacher. If so, this would indicate that the norms had truly become norms that regulated the classroom’s statistical activities. We found evidence supporting both norms. One example was Emma’s utterance from Episode C, included above, which suggested that she acted in line with the first norm by mentioning different descriptors. It also indicated how she, as a student, contributed to negotiating and developing this norm.

In the presentation of the results of our analysis, we first briefly present the two norms and then demonstrate how they emerged and evolved across the three teaching modules.

3.4. PRESENTATION OF THE TEACHING MODULES

Our case encompasses three episodes, A, B, and C (Figure 1). Prior to this work, it had been three years since the class had last worked with statistics, and students had not worked with TinkerPlots previously (Konold & Miller, 2022). Students were involved in a new data-exploration activity in each episode. In the first episode, they examined aspects of their parents’ ages; in the second episode, they investigated the indoor climate of their classrooms; and in the third episode, they explored geographically-based income disparity in Denmark.

	Teaching module 1	Teaching module 2	Teaching module 3
Lesson 1	Students participate in pair work using TinkerPlots to compare the age distribution of mothers and fathers in the class.	The teacher introduces the issue of elevated CO ₂ levels in Danish classrooms, and students plan how to collect CO ₂ data in their own classroom.	The students work in small groups with TinkerPlots and practice using boxplots on the CO ₂ data and shoe size data generated for the practice.
Lesson 2	Students engage in pair work using TinkerPlots to compare their findings with national statistics from Statistics Denmark. The focus is on age and gender at the time of their parents' first child. Episode A	Students transfer the collected CO ₂ data to TinkerPlots and begin exploring the data.	The students collaborate in small groups using Tinkerplots to examine the meaning of the boxplot and other descriptors in the CO₂ data. Episode B
Lesson 3	Students are collecting data from their peers on Time spent on sports weekly and Time spent on cell phones daily.	Students engage in tasks involving constructed data from a mathematics test. Students practice calculating the mean, median, range, and quartiles.	The students collaborate in small groups to compose an argumentative letter to the editor of the local newspaper addressing the CO ₂ problem. They select representations from TinkerPlots to bolster their argumentation.
Lesson 4	Students collaborate in small groups, using TinkerPlots to explore and analyse their data on time spent on sports and mobile phones.	Students analyse their CO ₂ data in small groups and prepare presentations of their findings.	Students read the first newspaper article discussing geographically-based income disparity in Denmark. Subsequently, they reflect on the selection of statistical descriptors used in the article.
Lesson 5	Students present their findings in group presentations.	In small groups, students present their findings on the CO ₂ issue to the teacher and the school principal.	Students read a second newspaper article based on the same study, focusing on geographically based income disparity in Denmark. They then reflect on the statistical descriptors chosen for the article and compare the two different approaches to presenting the same data. Episode C

Figure 1. Overview of the three teaching modules, with the three selected episodes framed in pink

The first module was designed to familiarise students with TinkerPlots and let them explore new aspects of a familiar phenomenon through data exploration. When students investigated their parents' ages, they worked in groups and discussed their findings in classroom discussions. During group work, they followed a student guide, which balanced open exploration with an introduction to basic TinkerPlots features. Three main themes structured the guide: 'think and hypothesise,' 'explore data,' and 'draw a conclusion.' Because both TinkerPlots and data exploration were new to the students, the first teaching module was designed with a simpler data set and context so as not to overwhelm them. This initial work aimed to support the students' early development of statistical literacy by giving them opportunities to attend to patterns and variation in their own data, use basic statistical descriptors, and formulate data-based observations. In line with our framing of statistical literacy, these activities addressed both knowledge elements, such as descriptor use, and dispositional elements, such as noticing, questioning, and justifying data-based observations. Episode A derived from the classroom discussion in which students presented and discussed their findings about parents' ages.

The second module was designed to help students see how their data analyses could reveal aspects of their classroom's indoor climate and support participation in a (semi-) public debate. By working with their own data in a meaningful context, the intention was to support students' further development of statistical literacy by requiring them to interpret descriptors in relation to a real phenomenon and consider what their findings meant and why they might matter. To introduce the indoor-climate investigation, students read articles about poor indoor climate in Danish classrooms. A CO₂ meter was installed in the classroom, and students measured the CO₂ levels and documented factors such as time of day, subject, and classroom mood. At the end of the module, the students made an oral presentation to the headmaster that showed the results of their investigations. However, the students found it difficult to closely connect their arguments to the statistical information. As a result, the class revisited the CO₂ data at the beginning of the third teaching module, with the goal of understanding how to base their arguments on their statistical findings. The students wrote letters to the editor of the local newspaper, in which they presented arguments about their indoor climate based on the results of their investigations and included TinkerPlots data representations. The indoor-climate task asked students to develop arguments about the CO₂ levels, but it did not inherently require the use of multiple statistical descriptors. Episode B derived from the class's revisiting of the indoor-climate data at the beginning of the third module.

The third module drew on students' prior experiences to help them interpret statistical descriptors in relation to income disparity. By comparing statistical information and reasoning about a societal issue, this module invited students to engage in more formal and critical interpretations of data. The intention of this work was to support more advanced aspects of statistical literacy. When investigating income disparity in the third module, students compared the ways in which two newspaper articles discussed a study of geographically-based income disparity in Denmark. Instead of analysing the data, students had to use the formal descriptors to analyse income disparity. Furthermore, they were urged to consider the motivations behind the study and its potential audiences. Figure 2 presents screenshots of titles for the two articles. The article on the left (article A) was published by TV2, a national broadcasting company (Neupert, 2021): [Danes' income calculated for each parish – what the richest and the poorest earn](#). The article on the right (article B) was published by the Danish think tank, The Economic Council of the Labour Movement (Caspersen, 2021): [In the richest district, they earn five times more than the bottom earners](#). The students first read article A, which reported fewer details (only the average income of each parish), followed by article B. Article B gave more detailed information, such as the mean income and quartiles, and differentiated between income from those who were employed and those who were unemployed. Episode C derived from students' work with the income-disparity task.



Figure 2. Snapshot of two newspaper articles—Neupert (2021) on the left and Caspersen (2021) on the right—that present investigations of income disparity in Denmark

When the students read article A, they were unaware that another article would follow. They were to reflect on the following questions:

1. What is the primary concern or topic that the newspaper article addresses?
2. What numerical data does the article provide?
3. What is the statistical descriptor(s) in the article?
4. Explain how the statistical descriptor(s) help to describe and explain inequality in our society.
5. Which statistical descriptors with which you are familiar does the article not include?
6. Which of the foregoing descriptors may have provided interesting insights?
7. Why would exploring these additional descriptors have been interesting or valuable to understanding inequality better?

Afterwards, they read article B and reflected on the following guiding questions.

1. What data and descriptors are presented in the new article?
2. What do these data indicate?
3. What do we notice about these data?
4. What do the data tell us about inequality in Denmark?

The progression that connected the three teaching modules had a dual focus. First, it aimed to transition students from engaging in informal explorations and visualisations to a more formal statistical vocabulary. This approach aligned with the curriculum requirements for this age group, which emphasised descriptive statistics (e.g., means, medians, modes, and ranges) between years 4 and 6 (Ministry of Children and Education, 2019). Second, the design of the modules progressed from data directly related to the students' life worlds, such as their parents' ages, to make the data more relatable, to phenomena more distant from their immediate experiences, such as indoor climate and geographically-based income disparity. This structured progression aimed to enhance students' understanding of statistics while expanding their analytical skills by examining various topics that were increasingly distant from their life worlds to open the world to them through statistics teaching.

At strategically relevant points, the teacher introduced formal descriptors using the students' investigations as a point of departure. This was designed to scaffold students' transition to a more sophisticated understanding of statistical descriptors.

4. RESULTS OF THE ANALYSIS

When analysing the three selected episodes, we identified two sociomathematical norms. These norms concern (a) what counts as a valuable statistical description of a data set and (b) what counts as a valuable statistical answer to a statistical investigation. In their final form at the end of the third module, these norms could be characterised as follows. (a) *A valuable description or investigation of a data set involves selecting*

statistical descriptors that are appropriate for the distribution, critically evaluating what each descriptor reveals, comparing descriptors to one another, and considering alternative descriptors when relevant in order to gain deeper insight into the phenomenon under investigation. (b) A valuable statistical answer uses descriptors to critically describe, explain, analyse, and interpret a societal phenomenon; shows how the descriptors reflect the distribution in its context; and justifies the relevance of these descriptors for understanding the phenomenon in question.

We next show how the teacher and students gradually negotiated and developed these two norms. Rather than assuming that these norms were present from the outset, we show how Episode A reflects early, teacher-initiated attempts to establish criteria for statistical descriptions; how Episode B provides evidence of emerging student uptake of these criteria; and how Episode C provides evidence of students increasingly taking up and justifying the use of descriptors in ways that indicate that the norms have become part of the classroom's microculture.

4.1. EPISODE A: STUDENTS DESCRIBE GENDER-BASED DIFFERENCES IN THEIR PARENTS' AGES—MEAN, MODE AND RANGE

Just before the conversation in Episode A, classroom discussion centred around the technical aspects of TinkerPlots. The class discussed which representations best presented the age distribution, and how they determined averages (mean and median, specifically). The teacher asked students for their answers, which led to a discussion about the specific descriptors students used in their data analysis. The parents' ages are presented in Figure 3, which is a reconstruction of the students' original presentation on the electronic board (the quality of the video recording was too poor to use).

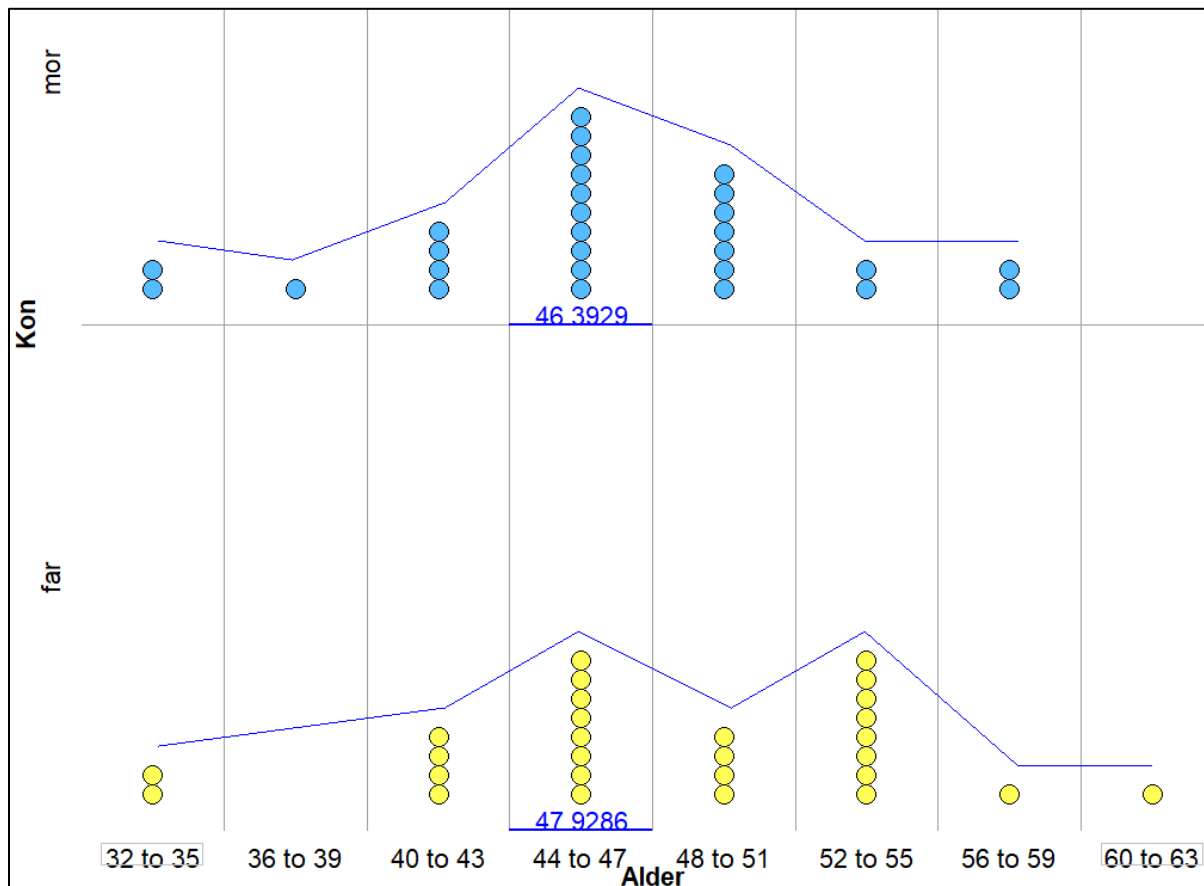


Figure 3. Ages of students' mothers (top) and fathers (bottom)

- Teacher: (...) What did you come up with? Do you remember?
John: You mean moms and dads?
Teacher: Yes.
John: So, mom was 46.3 and dad was 47.9.
Teacher: What do you think? Did you expect that the fathers would be older the mothers, or what? Did you expect that? Do you remember? What did you expect? What did you figure out?
John: I can't remember.
Martin: I was in a group with John, and we thought the fathers were older.
Teacher: Yes, that they were a bit older. So, they were. So, that was the mode, or no, the mean, that we discussed, and you expected the fathers to be slightly older the mothers.

The teacher acknowledged John and Martin's findings and repeated the mean values they reported. By doing so, she highlighted their use of statistical descriptors as a legitimate way of describing the data. Simultaneously, she encouraged the students to relate these numerical results to their expectations when she asked what they had anticipated about mothers' and fathers' ages. In this phase, the criteria for what counts as a valuable description or answer were largely shaped by the teacher's questions and interpretations. Episode A, therefore, illustrates the beginning of a process in which the teacher initiated ways of using descriptors to describe a data set and interpret it in relation to context.

The conversation continued, and the teacher asked for alternative ways to describe the data set.

- Teacher: Has anyone considered anything that didn't involve the mean, but focused on something other just the mean? What else can you see from these (nods towards the board, see Figure 4)? What can you see, John?
John: You can put that line on.
Teacher: The ruler, or what?
John: Yes.
Teacher: What does it show? What do you use that ruler for? Others besides John and Martin also used it, didn't they?
Emma: It's just a guess. How many years of difference there are.
Teacher: Exactly. So, if we placed the ruler starting from the first age at 35, and going all the way up to the last one, which is 61, what would it measure?
Iris: How many years are in between.
Teacher: Yes, so it would measure the difference between the youngest and the oldest. Can you try to do that, John, so we get the exact difference?

John dragged down the ruler on the screen and measured between the points, indicating the oldest and the youngest fathers. The teacher simultaneously commented.

- Teacher: Yes, up there, we have the oldest, and there we have the youngest. Like that.
Emma: That means the oldest could actually be the father of the youngest.
Teacher: Yes, yes, you could actually say that, because there is a 26-year difference in the ages. So, it would be possible. The difference between the oldest and the youngest is called the range of variation, and it's one way to describe your data set. Along with the mean, you can also find the range. And that's what you did there with your ruler, John. I remember you had a bit of trouble getting it to show the age, that 26.0, but you figured out how to do it. That was really impressive.

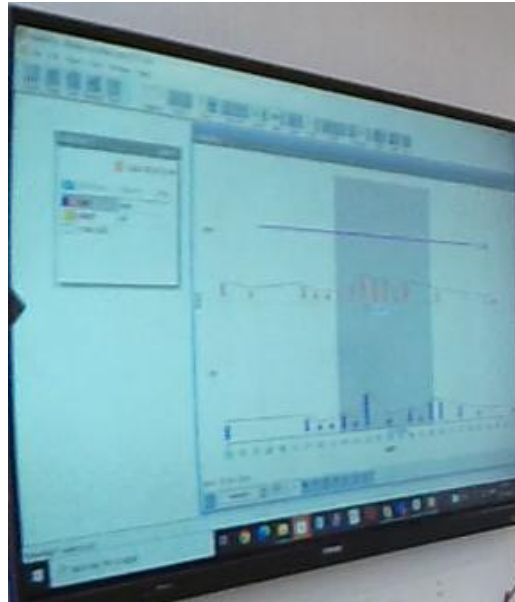


Figure 4. Image of John's computer work where John activated the ruler, the 'draw line' function, and intervals

The teacher requested an alternative descriptor beyond mean age. John suggested using 'that line,' referring to the ruler tool in TinkerPlots that can be used to measure the distance between data points on the dot plot. The teacher prompted the class to consider what the ruler showed, which led to identifying the range of variation as another descriptor. She formalised this by naming the range and recognising John's work. This exchange shows that the teacher was introducing the idea that different descriptors capture different aspects of the distribution. Emma's spontaneous comment about the age difference illustrated that students were beginning to notice contextual implications of the data and to engage in the negotiation of the norms initiated by the teacher.

At the end of the lesson, the teacher asked for yet another way to describe the data set, and as part of her wrap-up, she wrote a list on the blackboard of the formal names of the descriptors discussed during the lesson.

- Teacher: (...) There is one age that is most common. What is it? What is the most common age here?
- William: So ... for which gender?
- Teacher: Is there a difference?
- William: Very little.
- Teacher: Very little? What is the most common age, if you are a father, William?
- William: 46.
- Teacher: 46. What if you are a mother?
- William: Then it is also 46, yes.
- Teacher: (...) That's actually called the mode. It's another way to describe a data set. So, we have (the teacher writes on the blackboard) the mean, the mode, and the range. There. In a moment, when you start working on your new investigation, you should try to find out what the mean, the mode, and the range are. Yes, that's how it is. Is there actually a difference in the range for mothers and fathers? Now, John has put it in here. What did he put it at? Those 26 years. Whose range is that?
- William: That is the dads.
- Teacher: That is the dads. How does the mother's range look? Is it larger or smaller?
- William: It is smaller.
- Teacher: It is smaller, yes. So, there is a difference in the age range depending on whether you are a mother or a father here. Yes, those were the three important concepts for today.

Toward the end of the episode, the teacher shifted the conversation toward naming further descriptors. She guided William through identifying the mode for mothers and fathers and wrote mean, mode, and range on the blackboard. Here, the focus was on compiling descriptors and clarifying terminology, rather than interpreting what they revealed. This suggests that the class was still working to establish basic tools for describing data sets.

In summary, Episode A showed the teacher introducing and highlighting several statistical descriptors and encouraging students to relate them to the context. Students made occasional observations, but the criteria for valuable statistical descriptions and answers remained largely teacher-initiated and teacher-directed at this stage. The episode therefore represents an initial, teacher-led phase of negotiating the foundations for the norms that developed more fully in Episodes B and C.

4.2. EPISODE B: STUDENTS DESCRIBE AND DISCUSS THEIR INDOOR CLASSROOM CLIMATE—MEAN, MEDIAN, MODE, RANGE, QUANTILES AND BOX PLOT

In Episode B, the teacher and students discussed their data sets concerning the CO₂ levels in their classroom. The students referred to the threshold limit value for CO₂ at 1000 ppm.

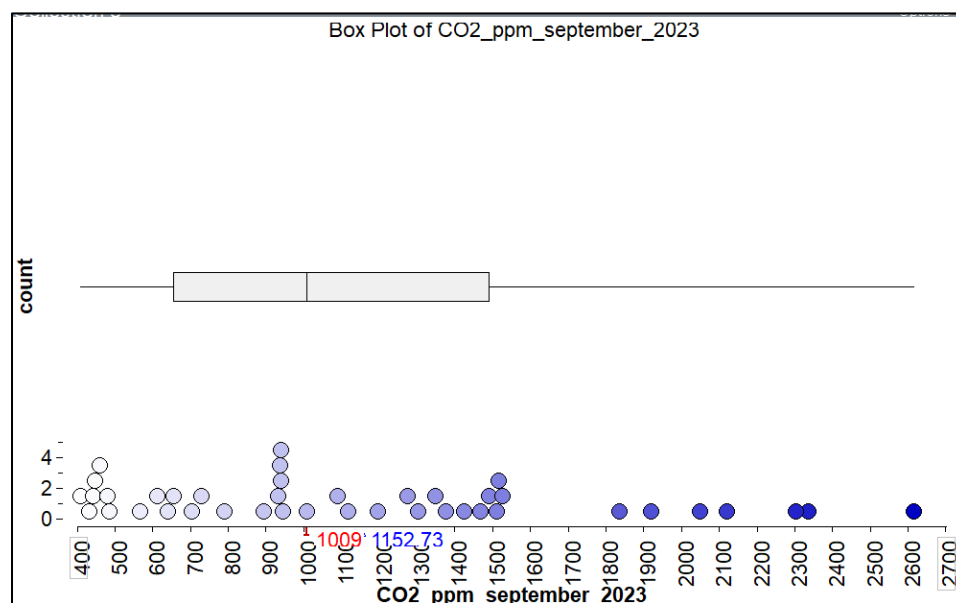


Figure 5. TinkerPlots screenshot showing CO₂ measurements with concentrations varied between 400 and 2600 ppm, mean of 1152 ppm, and median of 1009 ppm

- Benjamin: We have written that our indoor climate is not the best.
 Teacher: No. How so?
 Benjamin: Because half, that is 50% (...) because the one in the middle that shows 50% is at 1009, and that's over 1000, so our indoor climate might not be the best.
 Teacher: Okay. Yes. Have you written more?
 Benjamin: Yes, we also wrote that 25% is very, very high.
 Teacher: So, 25% is that long antenna. Those are incredibly high measurements [with marked emphasis]. That's still a quarter of the time. Have you written anything else, Benjamin?
 Benjamin: Hm... No, not really.
 Teacher: Okay. Emma, Amanda, and René, you are still writing. What have you written?
 Amanda: Yes, kind of ... actually a bit the same as
 Teacher: Yes, but please read it aloud. Go ahead and read it aloud.

- Amanda: Yes, I wrote that if you look at the mean, you can see that it's 1152, which is a bit too high, because, ideally, it should be under 1000. However, it's not extremely high, but it's still on the higher side. And then I also looked at the median, and on the box plot you can see that there are quite a few measurements at the low end, but there are also some at the high end. From 400 to 600, is 25%, which is quite good, but you can also see that from 1000 to 1500, there are also many high measurements.
- Teacher: And that's what you've written?
- Amanda: Yes, pretty much.

Benjamin contributed by using the median to argue that 50 percent of the measurements exceeded the threshold limit, and he supplemented his description with a reference to the upper quartile. These points indicated that he was beginning to use more than one descriptor to support his description of the indoor climate. Amanda combined the mean, the median, and observations from the box plot to describe different parts of the distribution. Her reasoning showed that she was drawing on several descriptors to substantiate her interpretation of the indoor climate. At this point, students' contributions showed that they were beginning to select descriptors they found relevant for describing the data set and forming an argument.

When Benjamin strengthened his argument by mentioning that the upper quartile had exceptionally high values, the teacher took up and reformulated his observation for the class. Her marked emphasis on "incredibly high measurements" and "a quarter of the time" signalled acknowledgement of his descriptor-based observation as relevant for interpreting the indoor climate data.

These interactions support the emerging idea that descriptors were useful for forming arguments about the phenomenon under investigation.

Then, Frida entered the conversation.

- Frida: We have written, well, it has already been said. [Teacher: It doesn't matter.] Well, the different 25% intervals show a wider and wider range of variation as the numbers increase.
- Teacher: How can you see that in the box plot?
- Frida: Well, the first line is shorter, and then the second one gets longer and then longer again.
- Teacher: So, they keep getting longer. And that last whisker there, it's very long, isn't it? So, there's a lot of variation in the measurements. Yes. Okay. Have you written more?
- Frida: What else did we write? Well, the last 50% are higher than the first 50% of the lower values, because the mean is higher than the median.
- Teacher: Okay, they bring it up.
- Frida: Yes, they bring it up more. The others bring it down.
- Teacher: Tomorrow, we will be working a lot on this, and now you have heard a bit about what it shows. You also need to consider your opinion. Do you think the indoor climate you have is okay? Or do you think it's unacceptable that 50% of the time there is too much CO₂ in here? What do you really think? Because you will need your own opinion tomorrow.

Frida's contribution added another dimension beyond description. She observed that the quartile intervals increased as CO₂ levels rose, and she used the box plot structure to justify her interpretation. Although she did not report exact quartile values, she identified a pattern in the distribution and used it to describe the data. Her subsequent comparison of the mean and the median suggests that she was beginning to consider how different descriptors relate to each other, although she did not arrive at a clear interpretation. Overall, these contributions indicate that students were beginning to use descriptors in ways that made sense to them, and that they were exploring how descriptors can support an argument, even if they did not yet fully articulate contextual interpretations.

The teacher ended the conversation by asking students to think about their opinion regarding the indoor climate and whether the CO₂ levels indicated a problem. This set the stage for a writing task the following day in which students were asked to write a letter to the editor, presenting their argument about the indoor climate. Students were encouraged to base their arguments on their own findings. Each student wrote such a letter, and at the end of the lesson the teacher asked volunteers to read theirs aloud.

- Teacher: What, uh, Linda, may we hear what you have written?
- Linda: Yes. We have a problem. Our problem is that our CO₂ level is far too high. This can cause us students to become really tired as the day goes on. We could open the window in the summer, but this can be a problem during the winter. That's what a ventilation system is for. Our problem now is that the school does not have or is not willing to invest in a ventilation system. I think having a sound ventilation system is just as essential as getting new computers, books or other things. They could also put up some CO₂ meters so that students and teachers could be aware of the high readings. In our investigation, we can see that we have very high readings, which is clearly why the school day feels very long, and we get headaches. Our investigation shows that half of the days have CO₂ readings over 1000, which is really high. We can also see that the CO₂ level rises towards the end of the day because we forget to air the room. The mean is 1150, which is way too high, when it should be under 1000. Our median is 1000. Clearly, a solution to our problem is to invest in a ventilation system, to have better indoor air quality, because it is extremely important to have a good indoor climate. Until we can get that, we need to have a CO₂ meter in every classroom to raise awareness of the CO₂ level, so that we can open the windows if it is too high.
- Teacher: Thank you, Linda. That was lovely. I could hear you also included some thoughtful statistical concepts. René, may we listen to yours as the last one?
- René: Okay. I think we have very poor indoor air quality in schools in Denmark. We need to consider how high the CO₂ level is. High CO₂ levels make it harder for us students to concentrate, which makes it more difficult for us to do anything and learn anything in class. We could save an entire school year with better indoor air quality. I mean, eight years of school with good indoor air quality is the same as nine years with poor indoor air quality. I think schools in Denmark should do something about this. Temporarily, we can start by remembering to open the windows and instal CO₂ meters in the classrooms, but over time, the school needs to get a better ventilation system than we have now. This will make a huge difference in the long run. (...) In our study, we see that over 50% of the time, when we are at school, the indoor air quality is too poor. That is half of the days. According to new research, a study shows that 20% of Danish classrooms have such high concentrations that the children's ability to make rational decisions is impaired to a dysfunctional level. I think the school needs to be quicker about taking a decision. Even though it costs money, it's all worth it in the end, when all the students become an entire year smarter because of it. In the meantime, until we get the ventilation system, the teacher just needs to remember to open the windows and install CO₂ meters so we can see when it's time to open a window.
- Teacher: Thank you. We also need to open a window in a bit. The teacher needs to remember that. You too, right? Yes. Listen, this has been a super difficult task you've been working on, and you have all approached it and done an amazing job.

These two students interwove the results of their investigation with personal experiences and opinions in their letters. Linda used the mean and the median to argue that the CO₂ levels were consistently too high. René referred to the median and to external information to support his argument about the seriousness of the situation. In both cases, students independently selected descriptors that they considered relevant to their arguments. Because the task did not require the use of particular statistical descriptors, these choices indicate that students were beginning to see descriptors as useful tools for substantiating claims. The teacher's feedback valued students' use of descriptors and supported their understanding that statistical measures can be used to interpret and communicate findings.

To summarise, Episode B shows that students were beginning to use descriptors in more autonomous ways. They drew on multiple descriptors to support their arguments and connect them to the question of indoor climate quality. These contributions demonstrate that students were starting to negotiate the value of using descriptors to justify claims. Episode B therefore represents an intermediate stage in which students began to use and negotiate the criteria initiated by the teacher in Episode A.

4.3. EPISODE C: STUDENTS DESCRIBE AND EXPLAIN GEOGRAPHICALLY-BASED INCOME DISPARITY IN DENMARK—MEAN, MEDIAN AND MODAL INTERVALS

In Episode C, the teacher facilitated a classroom discussion about what insight different descriptors can provide about the geographically based income disparity in Denmark. The teacher focused on what the descriptors in newspaper article A stated about the phenomenon, and then on alternative, relevant descriptors that might have been interesting to know about.

- Teacher: Yes. What about you, Brian and Jakob?
- Brian: (...) The mean tells us in which cities and parishes people earn most and least, and that can also tell us about inequality.
- Teacher: Yes, so the mean. What ... in which areas and parishes they earn the most and the least and then, yes. Which descriptors that are not presented to you would have been interesting to use, and why? John?
- John: Mode and median and range of variation, and quartile.
- Teacher: So, there is a good deal that is not presented. Are there any of those that you think could be interesting to investigate?

Although the task (see Section 3.4) asked students to identify potentially relevant descriptors, John's contribution shows that he might draw on a range of measures he knew. The next exchange turned from naming possible descriptors to discussing why some of them might be informative for interpreting income disparity.

- Frida: So, if one calculated the mean and such ... That is, if one compared the mean and the median. That could be very interesting. [Teacher: Yes, why?] Because it reveals whether there are any extremes, since that would influence the mean, but not the median. If the mean is higher than the median, then there are some very high ... which are higher than they are low.
- Teacher: Yes, so it could be for each parish, where there (...). It might appear as though they earn an extreme amount of money, but in reality, most people have an ordinary salary, and then there are a just few who earn an extreme amount of money. That would push the mean up, but in that case, the median would reveal whether this was the case. What did you write here?
- Finn: Well, we wrote that the median would be nice to know, and that (...) er, if the mean (...) so that there would not be a single really rich person, and then everybody else was poor. So, if you knew the median, you could see whether they [*the mean and median, ed.*] were close to each other, and then every person would have more or less the same salary. If not, if they, what is it called (...) if they were far apart, then it could be someone being really rich and everybody else who has less money.

The suggestion to compare the mean and median became a focal point in the discussion. Frida explained why this comparison could be informative: extreme incomes would influence the mean but not the median. Finn later extended this reasoning by suggesting that the distance between the mean and median could indicate whether salaries were relatively similar or whether the mean might be affected by a particularly high income. Although his explanation did not capture all aspects of variation, it shows that he was beginning to use the relationship between descriptors to interpret the income distribution. These contributions show that students were beginning to explain not only which descriptors could be relevant, but also why they matter and what they reveal about the distribution.

- Teacher: (...) Carolina, Anna, and Lise, what have you found out?
(The students are talking to each other quietly.)
- Carolina: We came across the median. When you are looking at the mean income of all adults, and then the median income of all adults in rich cities, right. There is a huge difference between the median and the mean. We think that is quite interesting.

- Teacher: Yes. So, where rich people live, there is a big difference between mean and median. What could that tell us? What are the demographics in these areas? We see a significant difference between mean and median. What does that tell us? Emma?
- Emma: I looked at ... It is like ... well, it probably means that in the wealthy [parishes], there is a big difference between median and mean. Still, in the poor parishes, there is almost no difference. So that probably means that ... like... in the wealthy parishes, it is more distributed, so few earn much money. However, there are also more ordinary people who still make a lot, but not like 1½ million. So, that is why the median is dragged down, because some still have an ordinary income. But, in the poor parishes, it is pretty much the same. Everybody earns pretty much the same there.
- Teacher: So, you could say, if you look at inequality, then is the inequality larger in wealthy parishes than in poor parishes? Yes, okay. (...)

After reading article B, which included additional descriptors, Carolina reported that in wealthy parishes there was a large difference between the mean and the median. The teacher asked what this might indicate about the demographics in these areas, and Emma built on the group's observation. She explained that a large difference between the mean and median might suggest that a few individuals earned very high incomes while most earned more moderate amounts. In contrast, she interpreted the small difference between mean and median in poorer parishes as indicating that incomes there were more similar. Emma's explanation shows how descriptors might help make sense of societal patterns such as income inequality.

Across the transcript segments, the students used and discussed statistical descriptors in ways that reflected increasingly autonomous reasoning. When responding to the prompt, students not only listed possible measures but also justified why they might matter. In their subsequent comparisons of mean and median, they explained what these differences could reveal about the distribution. Several students linked the differences between descriptors to plausible interpretations of income patterns in the parishes. The teacher facilitated this development by asking open questions and inviting students to elaborate on their reasoning rather than directing them toward specific answers.

Together, these contributions suggest that selecting, comparing, and interpreting descriptors in context had become part of teacher–student interaction when exploring the use of statistics in the two articles. Across the exchange, the idea of comparing mean and median recurred across several student contributions. Although the teacher continued to initiate and mediate the discussion, students increasingly became oriented to this comparison as a relevant way of interpreting income distributions. The students no longer used descriptors simply to label features of the data but used them as resources for explaining and understanding societal patterns. In this sense, Episode C also shows forms of participation aligned with the dispositional elements of statistical literacy, because students considered what the available descriptors showed or omitted, justified statistical interpretations, and communicated data-based claims about income disparity. Episode C therefore provides evidence that the two sociomathematical norms had begun to function as part of the classroom's microculture.

5. DISCUSSION

This study explored the development of sociomathematical norms that promoted middle-school students' deeper interpretations and articulations of data, indicating the development of their statistical literacy in terms of both knowledge and dispositional elements in an ordinary Danish school setting. The analysis showed that the teacher initiated and guided the negotiation and development of two prominent sociomathematical norms, but that the students also played a role in the negotiation. By analysing the regularities in whole-class interactions, we identified norms of what counts as a valuable statistical description of a data set and what counts as a valuable answer to a statistical investigation. The criteria for activities and discourses related to each norm changed substantially across the three episodes, from simple benchmarks to more advanced criteria for statistical analysis. For example, in assessing or negotiating what constituted a valuable statistical description, the criteria evolved *from* using one (or a few) statistical

descriptor to label features of a data set *to* selecting appropriate descriptors for a distribution and critically evaluating what each descriptor revealed about the phenomenon under investigation.

Notably, the analysis showed that the continued negotiation of these two norms shaped students' opportunities to engage in statistical investigations and to gradually develop aspects of statistical literacy. These opportunities included potential to develop both knowledge elements, such as statistical descriptors, symbolisations and calculation methods, and dispositional elements, such as students' capacity and willingness to question, justify, and communicate about statistical information in meaningful contexts. The study thus contributes to the limited empirical research on how to support students' development of statistical literacy, including its dispositional elements, in authentic classroom settings, and does so from a social constructivist perspective.

Although the first descriptions of statistical literacy are more than 20 years old, and there have been strong visions for students' development of statistical literacy, research shows that students' progressions beyond procedural understanding have proven difficult even in intervention studies targeting statistical literacy (Callingham & Watson, 2017). These intervention studies and other empirical studies have provided rich accounts of tasks, tools, and contexts for teaching statistics, especially about how societal issues can provide a meaningful context for developing the dispositions of statistical literacy (Ridgway, 2022). Our study emphasises that such tasks, tools, and contexts are important for students' development of statistical literacy, which is evident in both the design and enactment of the three teaching modules.

In relation to previous work investigating sociomathematical norms in statistics classrooms (Cobb, 1999; Cobb & McClain, 2005), our findings show similarities and differences. One similarity is that this study also documents how norms for what counts as statistically reasonable descriptions and answers are continuously negotiated between students and their teacher in classroom interactions. A distinction concerns the scope of the intended student autonomy. In Cobb and colleagues' work, intellectual autonomy referred to students' participation in and responsibility for statistical reasoning within the microculture of a classroom, whereas our approach to statistical literacy extended beyond classrooms toward civic and societal contexts with the aim of preparing students to participate in public debates and democratic life. Our analysis showed that the two emerging norms supported students in using statistical descriptors to make sense of societal issues. In the indoor climate activity (Episode B), several students used descriptors to evaluate whether their classroom met acceptable conditions, and some expressed that both the school, the municipality, and the students themselves should take responsibility for addressing this problem. In the income-disparity activity (Episode C), students compared the mean and median to question how newspapers constructed different accounts of inequality by including or omitting information. These forms of reasoning point toward a broader goal of autonomy associated with statistical literacy. Although the students did not reach the fully critical stance described by Weiland (2017), their reasoning aligned with the dispositional elements highlighted by Gal (2002, 2021), including the legitimacy of questioning statistical information and forming justified viewpoints. The sociomathematical norms documented here therefore constitute an important foundation. They supported students' interpretations and articulations related to selecting, comparing, and justifying descriptors in ways that created opportunities to question statistical information, justify statistical interpretations, and communicate data-based claims about real-world conditions that mattered to them. In this way, the norms connected classroom activity to the aims of statistical literacy, particularly its dispositional elements.

However, our study has clear limitations because it is based on a single case. Because this case is neither an outlier (extreme) nor a key case, it does not provide exemplary or generalisable insights according to Thomas (2011). Nonetheless, this local-knowledge case—set in a regular classroom setting at a normal public Danish school with ordinary students and a teacher—demonstrates that it is feasible to shift how students engage with statistics, from procedural to explorative approaches, by fostering productive sociomathematical norms.

Two practical findings of the study might therefore be of value for teachers, teacher educators, and others interested in improving the teaching of statistics. The first practical finding is that the designed tasks seemed to constitute an important instructional support for the teacher. Across the three teaching modules, these tasks progressed *from* engaging students in collecting data directly related to their life worlds, such

as their parents' ages and indoor climate, *to* working with reported statistical summaries (not raw data) about phenomena gradually more distant from their immediate experiences, such as geographically based income disparity. This task progression seemed to help the teacher to support students in shifting *from* using descriptors to label features of a data set, such as the mean of their fathers' ages, *to* use descriptors as elements in their reasoning about a phenomena, which they had access to only through statistical descriptors, such as critically interpreting what the mean of the incomes in a parish can say about the income-distribution in this parish. Such a shift may support students in using descriptors more independently by requiring them to select, compare, and justify descriptors when interpreting reported statistical summaries.

The second practical finding is that the teacher played a key role in continuously negotiating and establishing productive sociomathematical norms. Our analysis suggested that teacher questions were most productive when they prompted students to interpret what descriptors revealed about a data set, encouraged critical reflections on context, and fostered connections between descriptors and their real-world meaning. The importance of such subject-related teacher actions aligns with findings from Yackel (2002).

6. CONCLUSION

In this study, we explored the development of sociomathematical norms that supported middle-school students in developing deeper interpretations and articulations of data, thereby contributing to their development of statistical literacy in terms of both knowledge and dispositional elements in an ordinary Danish classroom. Over the course of the one-year case study, the class shifted from working with statistics in a mostly procedural manner to engaging in statistical investigations, where they collected data and selected, compared, and justified descriptors in ways that seemed to help them question, justify, and communicate about real-world phenomena that mattered to them. We identified the development of two sociomathematical norms as important for these changes: what counts as a valuable statistical description of a data set and what counts as a valuable statistical answer. The teacher and students' negotiation of these norms influenced students' statistical activities and also connected the classroom interactions to the broader aims of statistical literacy. The study contributes to the limited empirical research on how to foster students' statistical literacy—an elusive and complex educational goal—and on how this can be achieved in the social life of a classroom.

Our study suggests two directions for future research: to investigate the mathematical practices that emerge in statistical classrooms targeting statistical literacy to obtain deeper insights into students' learning, and to examine productive teacher actions in such classrooms.

REFERENCES

- Bargagliotti, A. E., 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)*. American Statistical Association and National Council of Teachers of Mathematics. https://www.amstat.org/asa/files/pdfs/GAISE/GAISEIIPreK-12_Full.pdf
- Ben-Zvi, D., & Garfield, J. (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. 3–15). Kluwer Academic Publishers. https://doi.org/10.1007/1-4020-2278-6_1
- Ben-Zvi, D., Gil, E., & Apel, N. (2007). What is hidden beyond the data? Helping young students to reason and argue about some wider universe. In D. Pratt & J. Ainley (Eds.), *Reasoning about informal inferential statistical reasoning: A collection of current research studies* (pp. 29–35). University of Warwick.
- Büscher, C. (2018). *Mathematical literacy on statistical measures: A design research study*. Springer. <https://doi.org/10.1007/978-3-658-23069-2>

- Büscher, C. (2022). Design principles for developing statistical literacy in middle schools. *Statistics Education Research Journal*, 21(1), Article 8. <https://doi.org/10.52041/serj.v21i1.80>
- Büscher, C. (2024). Adapting Habermas' construct of communicative rationality into a framework for analyzing students' statistical literacy. *Educational Studies in Mathematics*, 117(1), 121–141. <https://doi.org/10.1007/s10649-024-10325-5>
- Callingham, R., & Watson, J. (2017). The development of statistical literacy at school. *Statistics Education Research Journal*, 16(1), 181–201. <https://doi.org/10.52041/serj.v16i1.223>
- Caspersen, S. (2021, April 3). *I det rigeste lokalområde tjener de fem gange mere end bunden* [In the wealthiest local area, they earn five times more than those at the bottom]. Arbejderbevægelsens Erhvervsråd. <https://www.ae.dk/analyse/2021-04-i-det-rigeste-lokalomraade-tjener-de-fem-gange-mere-end-bunden>
- Charmaz, K. (2014). *Constructing grounded theory*. SAGE.
- Clark, P. G., Moore, K. C., & Carlson, M. P. (2008). Documenting the emergence of “speaking with meaning” as a sociomathematical norm in professional learning community discourse. *Journal of Mathematical Behavior*, 27(4), 297–310. <https://doi.org/10.1016/j.jmathb.2009.01.001>
- Cobb, P. (1999). Individual and collective mathematical development: The case of statistical data analysis. *Mathematical Thinking and Learning*, 1(1), 5–43. https://doi.org/10.1207/s15327833mtl0101_1
- Cobb, P., & McClain, K. (2005). Principles of instructional design for supporting the development of students' statistical reasoning. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 375–395). Kluwer Academic Publishers. https://doi.org/10.1007/1-4020-2278-6_16
- Cobb, P., Stephan, M., McClain, K., & Gravemeijer, K. (2001). Participating in classroom mathematical practices. *Journal of the Learning Sciences*, 10(1–2), 113–163. https://doi.org/10.1207/S15327809JLS10-1-2_6
- 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. (2021). Promoting statistical literacy: Challenges and reflections with a Brazilian perspective. In C. Monterio & L. Carvalho (Eds.), *Temas emergentes em letramento estatístico* (pp. 37–59). UFPE Press.
- Gal, I. (2022). Critical understanding of civic statistics: Engaging with important contexts, texts, and opinion questions. In J. Ridgway (Ed.), *Statistics for empowerment and social engagement: Teaching civic statistics to develop informed citizens* (pp. 323–343). Springer. https://doi.org/10.1007/978-3-031-20748-8_13
- Konold, C., & Miller, C. D. (2022). *TinkerPlots* (Version 2022) [Computer software]. Learn Troop.
- Lopez, L. M., & Allal, L. (2007). Sociomathematical norms and the regulation of problem solving in classroom microcultures. *International Journal of Educational Research*, 46(5), 252–265. <https://doi.org/10.1016/j.ijer.2007.10.005>
- Ministry of Children and Education. (2019). *Matematik: Fælles mål* [Mathematics: Common objectives]. https://emu.dk/sites/default/files/2020-09/GSK_F%C3%A6llesM%C3%A5l_Matematik.pdf
- Neupert, S. (2021, April 6). *Danskernes indtægt opgjort for hvert sogn - det tjener de rigeste og de fattigste* [Danes' income broken down by parish – what the richest and poorest earn]. TV 2 Nyheder. <https://nyheder.tv2.dk/business/2021-04-06-danskernes-indtaegt-opgjort-for-hvert-sogn-det-tjener-de-rigeste-og-de-fattigste>
- Phadke, S., Beckman, M., & Morgan, K. L. (2024). Examining the role of context in statistical literacy outcomes using an isomorphic assessment instrument. *Statistics Education Research Journal*, 23(1), Article 4. <https://doi.org/10.52041/serj.v23i1.529>
- Ridgway, J. (Ed.). (2022). *Statistics for empowerment and social engagement: Teaching civic statistics to develop informed citizens*. Springer. <https://doi.org/10.1007/978-3-031-20748-8>
- 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>

- Skott, C. K., & Østergaard, C. H. (2018). En matematiklærers professionelle udvikling [A mathematics teacher's professional development]. In J. Bundsgaard, M. Georgsen, S. T. Graf, T. I. Hansen, & C. K. Skott (Eds.), *Innovativ undervisning med it. Forskning i tre demonstrationsskoleforsøg II* (pp. 157–178). Aarhus Universitetsforlag.
- Skott, J. (2008). Introduktion til Paul Cobbs matematikdidaktiske arbejde [Introduction to Paul Cobb's work in mathematics education]. *MONA*, 4, 42–58. <https://doi.org/10.7146/mona.v0i4.36629>
- Thomas, G. (2011). A typology for the case study in social science following a review of definition, discourse, and structure. *Qualitative Inquiry*, 17(6), 511–521. <https://doi.org/10.1177/1077800411409884>
- van Zoest, L. R., Stockero, S. L., & Taylor, C. E. (2012). The durability of professional and sociomathematical norms intentionally fostered in an early pedagogy course. *Journal of Mathematics Teacher Education*, 15(4), 293–315. <https://doi.org/10.1007/s10857-011-9183-y>
- Visnovska, J., & Cobb, P. (2019). Supporting shifts in teachers' views of a classroom statistical activity: Problem context in teaching statistics. *Mathematical Thinking and Learning*, 21(4), 285–305. <https://doi.org/10.1080/10986065.2019.1576003>
- Voigt, J. (1998). The culture of the mathematics classroom: Negotiating the mathematical meaning of empirical phenomena. In F. Seeger, J. Voigt, & U. Waschescio (Eds.), *The culture of the mathematics classroom* (pp. 191–202). Cambridge University Press.
- Weiland, T. (2017). Problematizing statistical literacy: An intersection of critical and statistical literacies. *Educational Studies in Mathematics*, 96(1), 33–47. <https://doi.org/10.1007/s10649-017-9764-5>
- Weiland, T. (2019). Critical mathematics education and statistics education: Possibilities for transforming the school mathematics curriculum. In G. Burrill & D. Ben-Zvi (Eds.), *Topics and trends in current statistics education research* (pp. 391–411). Springer. https://doi.org/10.1007/978-3-030-03472-6_18
- Widjaja, W. (2012). Exercising sociomathematical norms in classroom discourse about data representation: Insights from one case study of a grade 6 lesson in Indonesia. *The Mathematics Educator*, 13(2), 21–38. <http://hdl.handle.net/10536/DRO/DU:30048395>
- Yackel, E. (2002). What we can learn from analyzing the teacher's role in collective argumentation. *Journal of Mathematical Behavior* 21(4), 423–440. [https://doi.org/10.1016/S0732-3123\(02\)00143-8](https://doi.org/10.1016/S0732-3123(02)00143-8)
- Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education*, 27(4), 458–477. <https://doi.org/10.5951/jresmetheduc.27.4.0458>
- Ziegler, L., & Garfield, J. (2018). Developing a statistical literacy assessment for the modern introductory statistics course. *Statistics Education Research Journal*, 17(2), 161–178. <https://doi.org/10.52041/serj.v17i2.164>

STINE GERSTER JOHANSEN
 Department of Teacher Education
 University College Copenhagen
 Humletorvet 3
 1799 Copenhagen V
 Denmark