Chapter 5
Teaching Quality and Assessment Practice: Trends Over Time and Correlation with Achievement

Nani Teig and Jennifer Maria Luoto

5.1 Introduction

Examining teaching quality and assessment practices in primary schools is of paramount importance, particularly in the context of mathematics and science education. These subjects play a significant role in fostering the development of problem-solving and critical thinking skills, which are crucial for students' academic and long-term success (Delahunty et al., 2020). Although there has been a growing interest in examining teaching quality and assessment practice as key factors influencing student learning outcomes (Andrade, 2019; Klieme & Nilsen, 2022), few studies have compared these constructs in primary mathematics and science classrooms across Nordic countries. This chapter aims to contribute to this expanding field of research by investigating the trends in teaching quality and assessment practice over time as well as their relations to student achievement in mathematics and science across Nordic countries (i.e., Denmark, Finland, Norway, and Sweden). The findings can offer valuable insights for policymakers and practitioners in designing and implementing evidence-based policies and interventions that promote high-quality teaching and assessment practices. Ultimately, this research seeks to support the continuous improvement of education systems in Nordic countries, enabling students to reach their full potential through mathematics and science classrooms.

N. Teig (✉) · J. M. Luoto
Department of Teacher Education and School Research, University of Oslo, Postboks 1099, Blindern, Oslo, Norway
e-mail: nani.teig@ils.uio.no

J. M. Luoto
e-mail: j.m.luoto@ils.uio.no

N. Teig et al. (eds.), *Effective and Equitable Teacher Practice in Mathematics and Science Education*, IEA Research for Education 14,
https://doi.org/10.1007/978-3-031-49580-9_5

5.2 The Nordic Educational Contexts: Mathematics and Science Education

Nordic countries share similar historical, cultural, and economic characteristics, suggesting that schools and teachers operate under relatively comparable conditions (Teig & Steinmann, 2023). Despite these similarities, significant differences exist in the trends of average mathematics and science achievement across Nordic countries (see Chap. 1). While trends in average achievement have largely remained stable in Norway and Denmark, these trends have been decreasing in Finland and increasing in Sweden over several of IEA's Trends in International Mathematics and Science Study (TIMSS) cycles (Mullis et al., 2020).

Previous research has highlighted performance level differences across schools and classrooms within Nordic countries (Yang Hansen et al., 2014). In Finland, no school-level differences existed in grades four and eight, while Norway and Sweden displayed substantial school-level differences for both grades (Yang Hansen et al., 2014). Classroom-level differences also existed in these countries, with Finland showing considerable variation in performance compared to Norway and Sweden (Yang Hansen et al., 2014). While classroom differences may arise from the sorting of students into various classes, they could also reflect the disparities in classroom activities and student experiences (Creemers & Kyriakides, 2008). Teachers play a crucial role in providing learning opportunities through instruction and assessment, allowing them to monitor and improve student performance. Given the differences in achievement levels and the potential role of teaching quality and assessment practices in shaping student performance, it is increasingly important to examine these factors over time in Nordic countries.

5.3 Teaching Quality and Student Achievement

Teaching quality is a multidimensional construct and generally considered to be teaching practices that are related to some types of students' cognitive and non-cognitive outcomes (Baumert et al., 2010; Klieme et al., 2009; see Chap. 2 Theoretical Framework of Teacher Practice for further details). The conceptualization of teaching quality in this book is closely aligned with the Three Basic Dimensions (TBD) of teaching quality: classroom management, supportive climate, and cognitive activation (Klieme et al., 2009).

Classroom management refers to the strategies, techniques, and processes that teachers use to create and maintain a well-organized, focused, and orderly learning environment with minimal disruptions (Praetorius et al., 2018). This includes managing instructional time effectively, ensuring that students stay on task, implementing clear rules, and maintaining order and discipline (Marder et al., 2023).

Supportive climate pertains to the quality of interactions in the classroom and encompasses various aspects, including teacher support, classroom interaction (teacher–student and student–student relationships), and instructional clarity (Nilsen et al., 2016; Praetorius et al., 2018). Supportive climate includes addressing individual student needs, offering various learning opportunities and engaging materials, helping students to understand and link new concepts, clarifying conceptual misunderstandings, and setting clear expectations (Nilsen et al., 2016). This chapter focuses specifically on teacher support and instructional clarity.

Cognitive activation involves instructional approaches and learning tasks that stimulate students' cognitive processing, promote conceptual understanding, and encourage students to engage in higher-order thinking (Baumert et al., 2010; Förtsch et al., 2017; Klieme et al., 2009). The level of cognitive activation often depends on task selection and implementation in the classrooms (Baumert et al., 2010; Lipowsky et al., 2009).

Previous studies investigating the correlation between teaching quality and student achievement in mathematics and science using survey data have produced mixed results. Reviewing studies that investigate these correlations with TIMSS and the OECD's Programme for International Student Assessment (PISA) data, Klieme and Nilsen (2022), highlighted that most studies demonstrated a small positive correlation between teaching quality and achievement. (e.g., Bellens et al., 2019). For classroom management, findings have been inconsistent when correlating with student achievement (e.g., Bellens et al., 2019). Supportive climate displayed the strongest correlation with achievement among the TBD dimensions, while cognitive activation showed a positive correlation with mathematics achievement but often had a negative or insignificant correlation with science achievement when considering inquiry practices. In another review by Klieme (2019), classroom management was positively related to achievement in both subjects across all countries, after controlling for student background and school composition. In contrast, teacher support was not significantly related or only spuriously related to achievement. Cognitive activation showed a slightly positive correlation with mathematics achievement in most countries. However, inquiry-based teaching, as a form of cognitive activation, was negatively associated with science achievement. Both review studies also showed that the relations between TBD dimensions and student outcomes may vary across the teaching activities representing the TBD dimensions, subjects, and countries (Klieme & Nilsen, 2022). Hence, studies examining these relationships should account for and clearly explain these variations.

Focusing on Nordic countries (excluding Iceland), Nilsen et al. (2018) examined teachers' perceptions of teaching quality related to cognitive activation and teacher support in science for fourth and eighth grades using TIMSS 2015 data. While teaching quality was found to have a positive correlation with student achievement, the strength of the correlation, however, varied across the Nordic countries. In Denmark, science teachers' self-reported instructional quality showed no significant relationship with student achievement, while in Finland, Sweden and Norway, teaching quality was positively and significantly correlated with fourth-grade science achievement. For grade eight,[1] Nilsen et al. (2018) revealed a significant correlation between teaching quality and student achievement in science in Norway and Sweden. Furthermore, Teig and Nilsen (2022) investigated the profiles of science teaching quality focusing on teacher support and instructional clarity by exploring Norwegian students' perceptions of teaching quality in grades five and nine using TIMSS 2015 data. They found that teaching quality patterns varied across both grades. In general, students who perceived their teachers as having high teaching quality were somewhat more likely to have higher science achievement.

These studies also highlight the importance to exercise caution when interpreting the relationships between teaching quality and learning outcomes due to the potential for reverse causality (e.g., Nilsen et al., 2018; Teig & Nilsen, 2022). For instance, students with low achievement and negative attitudes towards schooling might perceive their teachers' instruction as low quality or more cognitively challenging compared to others. This complexity highlights the need to further explore the intricate relationship between teaching quality and student outcomes.

5.4 Teacher Assessment Practice and Student Achievement

Teacher assessment practice encompasses various methods and strategies that teachers use to gather evidence of students' current understanding and use it to inform educational decisions, such as in planning lessons, adapting instruction, selecting assignments, providing feedback, and assigning grades (Black & Wiliam, 2009; Gardner et al., 2010; Herppich et al., 2018). These assessments can take multiple forms, including formal assessments such as tests and quizzes, as well as informal assessments like observations and discussions with students.

The primary goal of classroom-level assessments is to provide support to both teachers and students as they work towards determining, monitoring, and enhancing performance (Andrade & Brookhart, 2020; Gardner et al., 2010). By employing effective assessment practices, teachers can better identify their students' strengths and weaknesses, gain valuable insights for adapting their instructional approaches, and empower students with a clearer understanding of the steps they need to take in order to improve their learning (Andrade & Brookhart, 2020; Kanjee, 2009). A holistic assessment practice facilitates a positive collaboration between teachers and students, leading to the development of engaging and effective learning environments.

Research has demonstrated that teacher assessment practice is an essential component of effective teaching and learning, as it often significantly impacts student outcomes (Andersson & Palm, 2017; Hattie, 2009; Palm et al., 2017; Panadero et al., 2017). Palm et al. (2017) conducted a review that revealed a positive relationship between student achievement in mathematics and three types of teacher assessment practice: feedback, student self-assessment, and teacher assessment with subsequent

[1] Only Norway and Sweden participated grade eight survey in TIMSS 2015.

instructional actions. In contrast, Mostafa et al. (2018) discovered a negative relationship between teacher feedback and science performance in nearly all countries participating in PISA 2015. When it comes to students' perceptions of teacher feedback in Nordic countries, students in Denmark, Finland, and Iceland seemed to perceive feedback less frequently than those in Norway and Sweden (Sortkær, 2019). Additionally, Nordic students perceived less feedback than students from other OECD countries (Sortkær, 2019). A recent study also showed that a high frequency of teacher feedback in Nordic countries was more commonly reported in low-achieving students and schools (Rohatgi et al., 2022).

Homework is a common practice that holds significant potential as an assessment tool when used for monitoring student learning in mathematics and science (Martin et al., 2016). While it may not be immediately apparent, homework can indeed be considered an assessment practice. This is because homework assignments allow teachers to evaluate a student's understanding and application of the content covered in the classroom. Moreover, it provides an opportunity for students to assess their own learning progress, identify gaps in their understanding, and practice problem-solving skills (Fan et al., 2017; Fernández-Alonso & Muñiz, 2022). Despite its widespread use, homework showed limited effects on student learning (Fan et al., 2017; Scheerens, 2016). This could be attributed to whether or not homework is used strategically to assess students' developing knowledge, such as identifying the specific types of tasks that may pose challenges.

The use of homework in Nordic classrooms has not been extensively studied, yet it is often a topic of political debate. In Finland, homework in mathematics is an integrated classroom practice at the lower secondary level, with lessons typically starting with homework review and ending with new homework assignments (Krzywacki et al., 2016; Luoto et al., 2022). This pattern suggests a cultural tradition of homework routines (Fernández-Alonso & Muñiz, 2020). Investigating homework as a potentially important part of teacher assessment practice can provide valuable insights into its cultural tradition and contribute to the ongoing debate on the efficacy of homework in promoting student learning.

Previous research highlights the complex relationship between assessment practices and student achievement, emphasizing the importance of context and subject matter. There is a clear need for conducting research that examines teacher assessment practice over time, especially in Nordic countries, to better understand the nuances of these relationships and to develop optimal assessment strategies for various subjects and educational contexts. Investigating the relations between teacher assessment practices and student achievement is also crucial, as it can provide invaluable insights into the effectiveness of current practices and help identify areas for improvement.

5.5 The Present Study

In the present study, we utilize data from TIMSS 2011, 2015, and 2019 in mathematics and science for grade four across Nordic countries to address the following research questions (RQs):

RQ 1. What are the trends in teaching quality and assessment practice over time?
RQ 2. What is the relationship between teaching quality and student achievement?
RQ 3. What is the relationship between teacher assessment practice and student achievement?

5.6 Methods

5.6.1 Data and Variables

Data for the analyses were drawn from TIMSS, a large-scale international survey that assesses student performance in mathematics and science at the fourth- and eighth-grade levels in participating countries every fourth year. To examine trends in teaching quality and assessment practices over time (RQ 1), data from TIMSS 2011 to 2019 were analyzed. Additionally, the study focused on the TIMSS 2019 grade four data to investigate the relations between teaching quality and achievement (RQ 2) and the relations between teacher assessment practice and student achievements (RQ 3).

Table 5.1 summarizes the different aspects of teaching quality and assessment practice that were addressed in the specific RQs and TIMSS data. It is important to note that only a few items measuring teaching quality and assessment practice were consistent across TIMSS 2011 to 2019. As a result, RQ 1 had less comprehensive coverage of these constructs compared to RQs 2 and 3, which examined the most recent TIMSS cycle in 2019. Further details on the specific items used to measure the constructs are presented in Sect. 5.7 Findings and Appendix 1.

Teaching quality

TIMSS assessed teaching quality using both student and teacher background questionnaires, examining various aspects of classroom management, teacher support and instructional clarity, and cognitive activation. To assess *classroom management*, the student questionnaire measured the frequency of various disruptive and disorderly behaviors in the mathematics classroom using six items (e.g., "my teacher has to keep telling us to follow the classroom rules" or "students interrupt the teacher") with a response scale: every or almost every lesson, about half the lessons, some lessons, and never. *Teacher support and instructional clarity*, as important aspects of supportive climate, were measured using student agreement on various statements (e.g., "my teacher does a variety of things to help us learn" or "my teacher is easy to understand") with a response scale that ranges from agree a lot to disagree a lot.

The manifestation of *cognitive activation* varies depending on the subject. Since a distinction can be made between general and subject-specific cognitive activation (Teig et al., 2019), this study operationalizes cognitive activation differently for mathematics and science. In mathematics, teachers were asked on how often they engaged students in generic cognitive activation (e.g., "relate the lesson to students' daily lives") or problem-solving (e.g., "apply what students have learned to new

Table 5.1 The aspects of teaching quality and assessment practice across RQs and TIMSS data

Construct	Questionnaire		RQ 1 Trends in teaching quality and assessment practice (TIMSS 2011–2019)	RQ 2 Relations between teaching quality and student achievement (TIMSS 2019)	RQ 3 Relations between assessment practice and student achievement (TIMSS 2019)
	Student	Teacher			
Teaching quality					
Classroom management	√			√	
Teacher support and clarity of instruction	√		√	√	
Cognitive activation	√	√	√	√	
Assessment practice					
Homework frequency		√	√		√
Homework time		√	√		√
In-class homework discussion		√	√		√
Teacher emphasis on assessment strategies		√			√

problem situations on their own”). In science, teachers were asked how often they engaged students in various inquiry-based cognitive activation (e.g., “design or plan experiments or investigations”). The items related to cognitive activation employ a frequency-based response scale, ranging from “every or almost every lesson” to “never”.

To supplement the teacher questionnaire, this study also used students’ responses on the frequency with which they conduct experiments in their science lessons to represent inquiry-based cognitive activation. The response scale for this component ranged from “never”, “a few times a year”, “once or twice a month”, to “at least once a week”. This approach provides a more comprehensive understanding of the role of cognitive activation in different subject areas.

Teacher assessment practice

As shown in Table 5.1, only homework frequency, the time needed to complete homework, and in-class homework discussion were measured repeatedly across TIMSS 2011–2019. New items representing how much importance teachers place on various assessment strategies in mathematics and science were first introduced in TIMSS 2019 (e.g., “asking students to answer questions during class”). Further details on the specific items used are presented in Sect. 5.7 Findings and Appendix 2.

Student achievement in mathematics and science

TIMSS assessed student achievement with a standardized test that covers cognitive domains and subject-specific content domains. Student achievement was estimated via a measurement model that produced a set of five plausible values to represent the likely distribution of student performance. All plausible values were incorporated into the analyses to produce an average of the model estimates and adjusted standard errors (see Chap. 3 Analytical Framework).

5.6.2 Data Analysis

Data were first prepared using the IDB Analyzer 4.0, while the main analyses were conducted using Mplus 8.8 (Muthén & Muthén, 2022) and IBM SPSS Statistics 28.0. To address RQ 1 about the trends in teaching quality and assessment practice over time, descriptive statistics and one-way ANOVAs were performed to test for significant mean differences within-country across three pairwise comparisons: TIMSS 2011 versus 2015, 2015 versus 2019, and 2011 versus 2019.

To investigate the relationship between teaching quality and student achievement (RQ 2) and the relationship between teacher assessment practice and student achievement (RQ 3), we implemented a two-level approach to account for TIMSS’ cluster sampling design in which students were nested within classrooms. Specifically, we employed multilevel structural equation modelling (MSEM) with students nested in classrooms. The MSEM approach served two main purposes: (a) established measurement models to represent teaching quality dimensions; (b) examined the relations between teaching quality, assessment practice, and student achievement. To accomplish (a), we used multilevel confirmatory factor analysis (MCFA)—an extension of CFA, to multilevel situations that have proved useful to study the factor structure of instructional practices at two levels (Brown, 2015; Morin et al., 2014). Model fit was evaluated using Ryu’s (2014) partial saturation approach by obtaining test statistics and fit indices for each level separately. We referred to common guidelines for an acceptable model fit (i.e., CFI $\geq$ 0.95, TLI $\geq$ 0.95, RMSEA $\leq$ 0.08, and SRMR $\leq$ 0.10; Marsh et al., 2005).

We further performed MSEM with a multi-group approach to separately examine the distinct relationships between teaching quality, assessment practice, and student achievement in each of the Nordic countries. Measurement invariance was conducted to ensure valid comparisons across groups (Sass & Schmitt, 2013). The data generally supported sufficient levels of measurement invariance (further details are presented in Chap. 3 Analytical Framework).

Analyses for RQ 2 were conducted at both the student and classroom levels, with two exceptions: (a) the relationship between the construct cognitive activation and student achievement, and (b) the relationship between assessment practice and student achievement, both of which were analyzed only at the classroom level.

5.7 Findings

5.7.1 The Trends in Teaching Quality and Assessment Practice Over Time

Figures 5.1, 5.2, 5.3, and 5.4 provide a graphical representation of the mean differences of these items across two-time TIMSS cycles (i.e., TIMSS 2011 versus 2015, 2015 versus 2019, and 2011 versus 2019).

Teaching quality

Classroom management was first introduced in TIMSS 2019 and is not included in the analyses of the mean differences between 2011 and 2019.

Two items were related to supportive climate and directly related to teacher support and instructional clarity in mathematics and science. As shown in Fig. 5.1, Nordic countries experienced an overall decline in teacher support and instructional clarity from 2011 to 2019. This decrease was more pronounced between 2015 and 2019 than in 2011 to 2015 in Denmark and Norway, while in Sweden, a different pattern was observed. Finland exhibited a significant increase between 2011 and 2015, followed by a decrease from 2015 to 2019. Similar patterns were identified in both mathematics and science.

Cognitive activation was measured using two items for general cognitive activation, three items for mathematics, and five items for science. The mean differences showed mixed patterns in cognitive activation across cycles, countries, and activities in both mathematics and science (Fig. 5.2). For instance, between 2011 and 2015, students' opportunities to observe natural phenomena and describe their observations decreased in Denmark, increased in Finland and Norway, and remained unchanged in Sweden. Between 2015 and 2019, the same activity decreased in Finland, increased in Norway and Sweden, and showed no significant changes in Denmark.

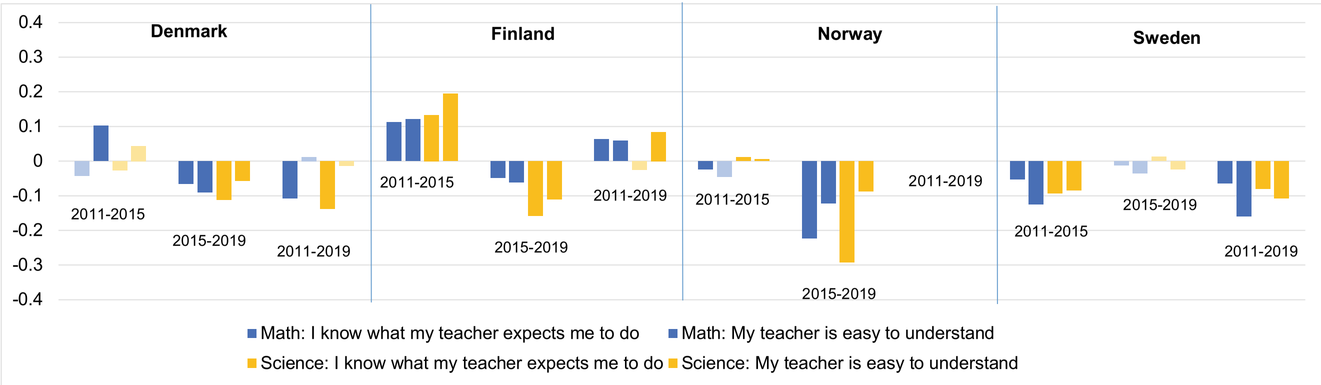


Fig. 5.1 The mean differences in students' perceived teacher support and instructional clarity. *Note* Dark bars = significant difference between cycles (p-value < .05), light bars = non-significant difference between cycles (p-value < .05), positive value = higher mean in the more recent TIMSS cycle of the comparison, negative value = lower mean in the more recent TIMSS cycle of the comparison

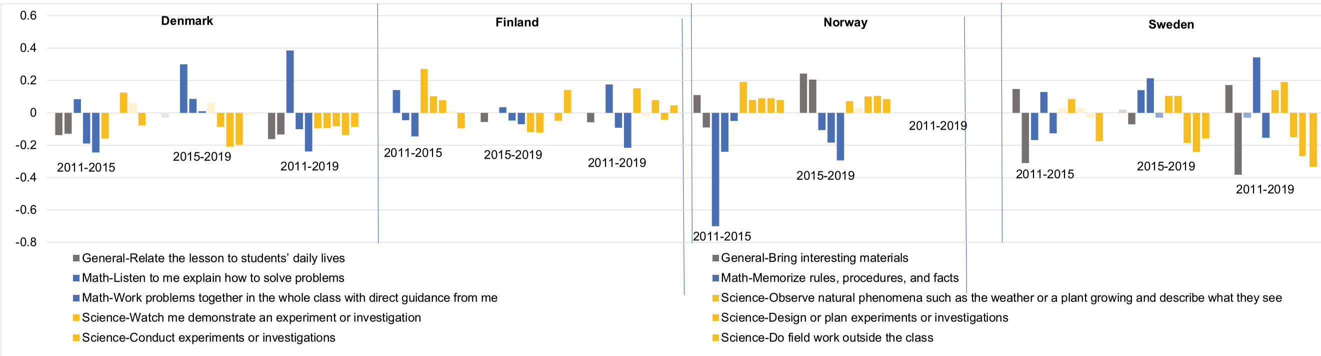


Fig. 5.2 The mean differences in teachers' perceived cognitive activation. *Note* Dark bars = significant difference between cycles (p-value < .05), light bars = non-significant difference between cycles (p-value < .05), positive value = higher mean in the more recent TIMSS cycle of the comparison, negative value = lower mean in the more recent TIMSS cycle of the comparison

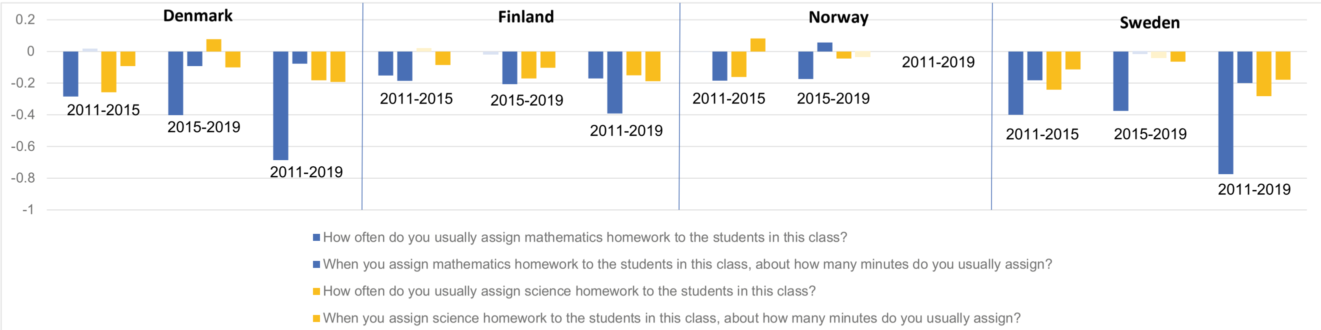


Fig. 5.3 The mean differences in teachers' self-reported assessment practice (homework). *Note* Dark bars = significant difference between cycles (p-value < .05), light bars = non-significant difference between cycles (p-value < .05), positive value = higher mean in the more recent TIMSS cycle of the comparison, negative value = lower mean in the more recent TIMSS cycle of the comparison

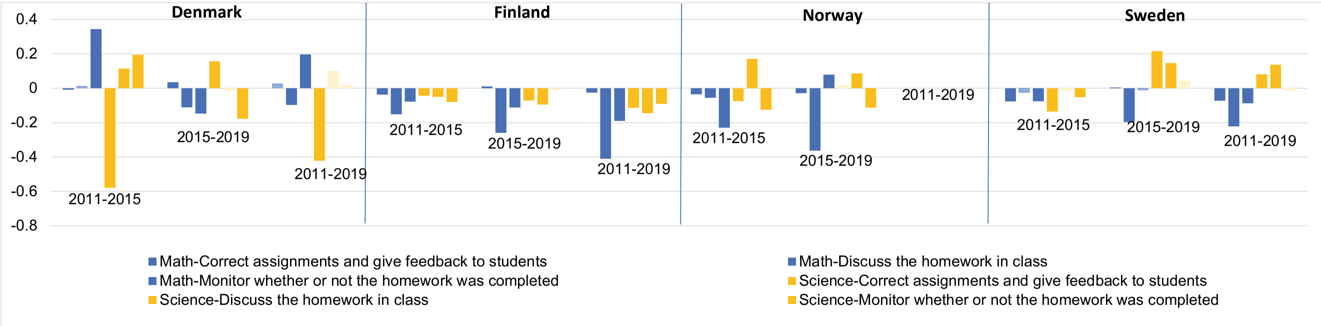


Fig. 5.4 The mean differences in teachers' self-reported assessment practice (the frequency of in-class homework discussion. *Note* Dark bars = significant difference between cycles (p-value < .05), light bars = non-significant difference between cycles (p-value < .05), positive value = higher mean in the more recent TIMSS cycle of the comparison, negative value = lower mean in the more recent TIMSS cycle of the comparison

Assessment practice

Two aspects of teachers' self-reported assessment practice were consistent across the cycles: homework and in-class homework discussion. Figure 5.3 shows an overall decrease in homework frequency and time to complete homework for both subjects across Nordic countries. Another aspect of assessment practice concerns the integration of homework into classroom instruction, represented by three items. Mixed findings were observed for the frequency of in-class homework discussions in the Nordic countries (Fig. 5.4). In Finland, there was a clear decrease in in-class homework discussions from 2011 to 2019. In Sweden, a similar decline occurred in both subjects between 2011 and 2015; however, while in-class homework discussion in mathematics (i.e., monitor whether or not the homework was completed) decreased between 2015 and 2019, the same activity increased in science. The changes varied across cycles, subjects, and even different activities within the same subjects in Denmark and Norway.

5.7.2 The Relationship Between Teaching Quality and Student Achievement

Findings from TIMSS 2019 revealed similar response patterns of teaching quality across Nordic countries (see Appendix 1 for further details on the items used to measure teaching quality). With respect to assessment practice, some variations were found in homework frequency, time to complete homework, and in-class homework discussion, but a similar response pattern was observed in terms of how much teachers place emphasis on various assessment strategies in mathematics and science. As shown in Appendix 2, among Nordic countries, the highest proportion of students taught by teachers who assigned homework was found in Finland. Nevertheless, the majority of Finnish students received homework that took only 15 min or less to complete. Homework assignments were more prevalent in mathematics than in science classrooms across these countries. Additionally, longer tests were more commonly used as assessment strategies in mathematics rather than science lessons, while the opposite was true for long-term projects.

The relations between teaching quality and student achievement

Findings from the MSEM analyses revealed stronger relationships between teaching quality and student achievement in mathematics than science across Nordic countries (Table 5.2). In mathematics, students' perceptions of classroom management were the most robust predictor of achievement compared to other dimensions of teaching quality. Note that classroom management was measured for the first time in TIMSS 2019 and only in mathematics. Perceived teacher support and instructional clarity were related to student achievement in mathematics, especially at the student level. These relations were only significant at the student level for science achievement in Denmark and Finland.

Table 5.2 The relations between teaching quality and student achievement at the student classroom levels

Predictors	Denmark		Finland		Norway		Sweden	
	Student	Class	Student	Class	Student	Class	Student	Class
	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)
Mathematics								
Classroom management	0.07** (0.03)	0.31** (0.11)	0.08** (0.02)	0.14 (0.12)	0.11** (0.03)	0.34** (0.11)	0.08** (0.03)	0.37** (0.09)
Teacher support and instructional clarity	0.22** (0.03)	0.13 (0.13)	0.09** (0.02)	0.20* (0.10)	0.10** (0.02)	0.15 (0.16)	0.06* (0.03)	0.20* (0.10)
Cognitive activation	–	0.22 (0.12)	–	0.14 (0.10)	–	0.02 (0.14)	–	0.12 (0.16)
Science								
Teacher support and instructional clarity	0.09** (0.02)	0.09 (0.12)	0.08** (0.02)	0.09 (0.10)	– 0.02 (0.53)	0.02 (0.91)	– 0.01 (0.03)	0.22 (0.12)
Cognitive activation	–	0.20 (0.18)	–	0.08 (0.11)	–	– 0.09 (0.13)	–	0.10 (0.12)

Note * $p < 0.05$, ** $p < 0.01$

Teachers' perception of cognitive activation was not related to student achievement in both subjects (Table 5.2). However, the findings revealed that the frequency of conducting experiments, measured using the student questionnaire, was related to science achievement in non-linear (inverted U-shape) rather than linear patterns. These relations were observed at the student level in all Nordic countries and at the classroom levels in Finland, Norway, and Sweden.

5.7.3 The Relationship Between Assessment Practice and Student Achievement

As demonstrated in Table 5.3, there were limited associations between teacher assessment practices and student achievement. Homework frequency was positively related to mathematics achievement in Denmark and Sweden, while the time needed to complete homework had a negative relationship with mathematics achievement in Finland. Correcting assignments and providing feedback to students was negatively related to mathematics and science achievement in Finland and Sweden. Notably, no correlations were found between the amount of emphasis teachers placed on various assessment strategies and student achievement in either mathematics or science.

5.8 Discussion and Conclusion

This study examines teaching quality and assessment practices in primary mathematics and science classrooms across Nordic countries (Denmark, Finland, Norway, and Sweden) using TIMSS 2011 to 2019 data. It investigates the trends in teaching quality and assessment practices over time, as well as the relationships of these aspects of teacher practice with student achievement.

5.8.1 The Trends in Teaching Quality and Assessment Practice Over Time

Analyses of TIMSS data from 2011 to 2019 have revealed a decline in aspects of teaching quality related to teacher support and instructional clarity across Nordic countries. One possible explanation for this decline is the changing characteristics of student populations. The TIMSS data showed an increase in the percentages of low socioeconomic status (SES) students and limitations to teaching (Mullis et al., 2020). The percentages of low SES students, as indicated by those who responded to having none or only 1 to 10 books at home, have increased from 7.3 percent in 2011 to 10.2 percent in 2019 (see Chap. 1). Teachers have reported increased limitations

Table 5.3 The relations between assessment practice and student achievement at the classroom level

Predictors	Denmark	Finland	Norway	Sweden
	β (SE)	β (SE)	β (SE)	β (SE)
Mathematics				
Homework frequency	0.23* (0.11)	0.74 (0.09)	0.32* (0.14)	0.09 (0.11)
Homework time	0.07 (0.11)	– 0.24* (0.08)	0.03 (0.13)	– 0.02 (0.08)
In-class homework discussion				
• Correct assignments and give feedback to students	0.07 (0.04)	– 0.09* (0.04)	– 0.07 (0.04)	– 0.07 (0.06)
• Discuss the homework in class	0.02 (0.05)	0.04 (0.04)	– 0.03 (0.04)	– 0.15 (0.06)
• Monitor whether or not the homework was completed	0.07 (0.05)	0.06 (0.11)	0.09 (0.08)	– 0.02 (0.07)
Emphasis on assessment strategies				
• Observing students as they work	– 0.02 (0.08)	– 0.11 (0.06)	– 0.02 (0.08)	– 0.10 (0.07)
• Asking students to answer questions during class	– 0.04 (0.05)	0.02 (0.05)	0.00 (0.05)	– 0.07 (0.06)
• Short, regular written assessments	0.02 (0.04)	– 0.03 (0.05)	0.07 (0.07)	0.02 (0.05)
• Longer tests (e.g., unit tests or exams)	0.05 (0.04)	0.09 (0.05)	0.05 (0.04)	0.09 (0.06)
• Long-term projects	0.05 (0.04)	– 0.01 (0.04)	– 0.05 (0.08)	– 0.03 (0.05)
Science				
Homework frequency	0.05 (0.14)	0.11 (0.09)	0.00 (0.18)	0.00 (0.16)
Homework time	0.29 (0.26)	– 0.12 (0.09)	– 0.08 (0.22)	– 0.12 (0.26)
In-class homework discussion				
• Correct assignments and give feedback to students	– 0.04 (0.11)	– 0.08* (0.03)	– 0.04 (0.07)	– 0.20* (0.07)
• Discuss the homework in class	0.02 (0.21)	0.06 (0.06)	– 0.05(0.07)	0.20 (0.30)
• Monitor whether or not the homework was completed	0.01 (0.21)	– 0.02 (0.07)	0.04 (0.07)	– 0.18 (0.10)
Emphasis on assessment strategies				
• Observing students as they work	– 0.03 (0.07)	– 0.16 (0.05)	– 0.00 (0.09)	– 0.02 (0.09)
• Asking students to answer questions during class	0.03 (0.07)	– 0.09 (0.06)	0.04 (0.11)	– 0.14 (0.09)
• Short, regular written assessments	– 0.05 (0.07)	– 0.03 (0.04)	– 0.04 (0.07)	– 0.07 (0.07)
• Longer tests (e.g., unit tests or exams)	0.04 (0.07)	0.08 (0.04)	0.02 (0.08)	0.00 (0.07)
• Long-term projects	0.04 (0.06)	0.03 (0.04)	– 0.07 (0.08)	– 0.01 (0.05)

* $p < 0.05$,

to teaching, including students lacking prior knowledge, being tired, hungry, and causing more disturbances (see Chap. 7). The increasing diversity of student populations may pose greater challenges for teachers in meeting the needs of all students, which could contribute to lower perceived teacher support and instructional clarity.

Additionally, changes in the curricula could be another factor that influences students' perceptions of teacher support and instructional clarity. The current curricula have become more demanding in terms of what is expected from students, and teachers are now expected to act as facilitators rather than giving strict directions (Carlgren et al., 2006). This shift in instructional practice places more responsibility on the students, which could lead to more students perceiving less teacher support. If teachers are not adequately trained or supported in providing effective instruction, students may perceive less support and instructional clarity in the classroom (Creemers et al., 2012; Darling-Hammond et al., 2017).

Although there is a lack of Icelandic studies investigating trends in teaching quality and assessment practice over time, prior research provides some insight into the prevalence of such practices. For example, studies from lower secondary schools indicate an overall low teaching quality in mathematics related to cognitive activation (Sigurjónsson, 2023) and instructional clarity related to feedback and clear learning goals (Svanbjörnsdóttir et al., 2023). While homework is generally considered an important practice by Icelandic school teachers, low-achieving students and their parents view it as too demanding, and these are the students that spend most time on homework (Sigurgeirsson & Björnsdóttir, 2016). This perception is unlikely unique to Iceland, and in what way such findings are related to achievement across diverse student groups needs to be investigated in future research.

The study also found an overall decline in homework frequency in mathematics and science between 2011 and 2019 across Nordic countries. This trend may be due to the growing emphasis on providing students with more meaningful and relevant learning experiences (Clement, 2010; Remmen & Iversen, 2022). Equity is another possible explanation for the decline in homework frequency in Nordic countries. Homework assignments may not always align with the needs and interests of all students, and low SES students may be less likely to have access to resources and parental support that can help them complete assignments effectively (Bempechat et al., 2011; Rønning, 2011). Consequently, some schools and educational systems may have shifted towards alternative forms of assignments that are more flexible and better aligned with students' backgrounds.

It is important to note that the reasons for the decline in the aspects of teacher support and instructional clarity, as well as homework frequency, are likely complex and multifaceted. Further research is necessary to identify the underlying factors contributing to these trends in order to inform policy and practice aimed at improving teaching quality and assessment practices in Nordic countries.

5.8.2 The Relationship Between Teaching Quality and Student Achievement

Classroom management was identified as having the largest correlation with student achievement in mathematics compared to other dimensions of teaching quality, which is consistent with previous research (e.g., Senden et al., 2023). Good classroom management is considered a prerequisite to facilitate other dimensions of teaching quality, such as in creating a supportive classroom climate and implementing cognitive activation (Charalambous & Praetorius, 2020), which can promote student achievement (Wolff et al., 2021). For example, by minimizing disruptions during learning, teachers can create a structured environment that fosters positive relationships between students and teachers and among students and maximize the amount of time available for cognitively challenging instruction.

This study also suggests that teacher support and instructional clarity seem to be better predictors of achievement at the student level compared to the classroom level. This could be due to individual differences in students' characteristics and background, such as their language abilities and socioeconomic status, which may impact the level and type of support they require from their teachers. Previous research has shown that students' perceptions of teaching quality can vary across diverse groups of students (e.g., Senden et al., 2023; Teig & Nilsen, 2022; Wang et al., 2018). Therefore, it is important to consider individual student characteristics when examining the relationships between teacher support, instructional clarity, and achievement. Measuring these factors at the student level may provide a more accurate understanding of these relationships.

Although previous studies have suggested that cognitive activation is an important dimension of teaching quality related to student learning outcomes (Baumert et al., 2010; Klieme et al., 2009; Lipowsky et al., 2009), the current study did not find any significant relationship between cognitive activation and achievement in mathematics and science across the Nordic countries. One possible explanation for this discrepancy could be related to the way cognitive activation was operationalized and measured in TIMSS, as some items were more related to low- rather than high-level of cognitive activation. For example, in mathematics, the items include memorizing rules, procedures, and facts. Therefore, this finding highlights the need for future research to explore the effects of various conceptualizations of cognitive activation and to what extent these variations matter for student achievement.

Measuring high-level cognitive activation can be challenging via the student or teacher questionnaire due to its context-specific nature. Thus, incorporating a qualitative perspective, like classroom observation, could be beneficial. For instance, a recent video study in Iceland reported frequent occurrences of low-level cognitive activation in mathematics classrooms (Sigurjónsson, 2023). Students often engaged in individual work that focused on procedural fluency, with limited connection to understanding mathematical concepts.

Additionally, the use of teacher rather than student questionnaires to measure cognitive activation may have impacted the findings. Teacher questionnaires may be more susceptible to social desirability bias, a tendency for teachers to answer the questionnaire in a way that will be perceived favorably by others (Muijs, 2006). By using the student questionnaires, this study found that the frequency of conducting experiments, as an indicator of cognitive activation in science, was related to student achievement in a non-linear pattern (inverted U-shape). This finding aligns with previous research (Cairns, 2019; Teig et al., 2018, 2021), and suggests that there may be an optimal level of conducting experiments that leads to the highest achievement in science.

Discrepancies in findings on the relationship between cognitive activation and achievement in mathematics and science may be due to how cognitive activation was measured and whether teacher or student questionnaires were used. Further research is needed to better understand this relationship and identify effective strategies for optimizing cognitive activation in these subjects.

5.8.3 The Relations Between Assessment Practice and Student Achievement

The findings of this study suggest that the associations between homework as part of teacher assessment practice and student achievement were limited. While the frequency of homework was positively related to mathematics achievement in Denmark and Sweden, homework time had a negative relationship with mathematics achievement in Finland. This finding may indicate that completing more homework does not necessarily lead to higher achievement, but rather that the quality and relevance of homework assignments are more important factors to consider (see a review by Fernández-Alonso & Muñiz, 2022). In Finland and Sweden, correcting assignments and providing feedback to students was negatively related to mathematics and science achievement, which is a finding that may warrant further investigation. These findings could indicate reversed causality, where low-achieving students spent more time completing homework and high-achieving students finished their work at school, and teachers placed more emphasis on correcting assignments and giving feedback to the struggling students. In Iceland, a recent study using classroom video data showed that although providing feedback is a common practice in mathematics classrooms, there was limited evidence of feedback being delivered with a clear purpose and of high quality to students (Svanbjörnsdóttir et al., 2023).

Moreover, no clear correlations emerged between the amount of emphasis teachers placed on various assessment strategies and student achievement in either mathematics or science. This may suggest that while assessment strategies are important for monitoring and evaluating student learning, they may not necessarily have a direct impact on student achievement. Furthermore, it is possible that the assessments used by teachers may not be aligned with the content and skills emphasized in the curriculum, and thus may not be as effective in promoting learning (Andrade & Brookhart, 2020; Gardner et al., 2010). Teachers may also need additional support and training in developing and using assessments that are aligned with the curriculum and promote student learning.

An alternative explanation might be linked to the limitations of the items utilized to measure assessment strategies in TIMSS. The items used in TIMSS might not encompass the complete array of assessment practices employed by teachers in the classroom, or they may not fully capture the complexity of these practices.

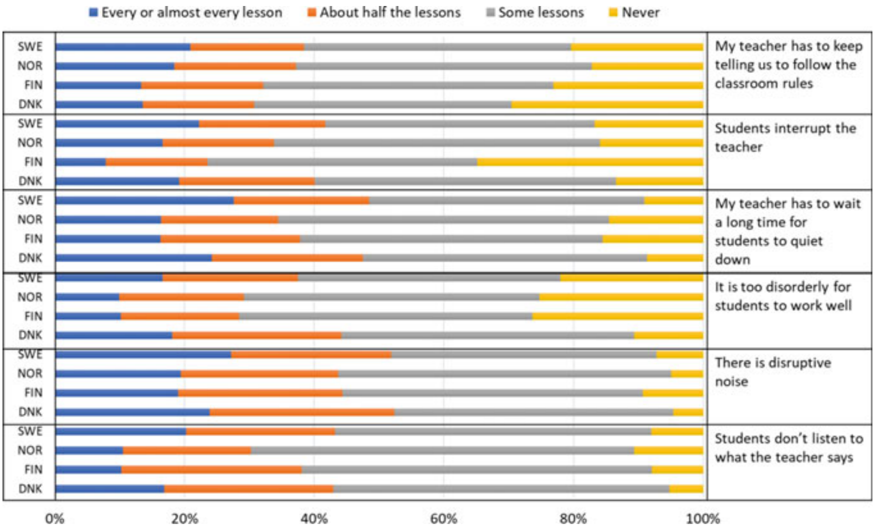
To conclude, this study highlights the need for a more holistic approach to teaching quality and assessment practices in mathematics and science education, one that considers multiple dimensions of teaching quality and emphasizes the need for further analysis of how the composition of the classroom may also impact the relationships between these constructs. With more nuanced understandings of these relationships, policymakers can allocate targeted resources, allowing educators to create a more supportive and equitable learning environment that promotes student achievement and success in mathematics and science.

Appendices

Appendix 1 Response Patterns of Teaching Quality Across Nordic Countries in TIMSS 2019

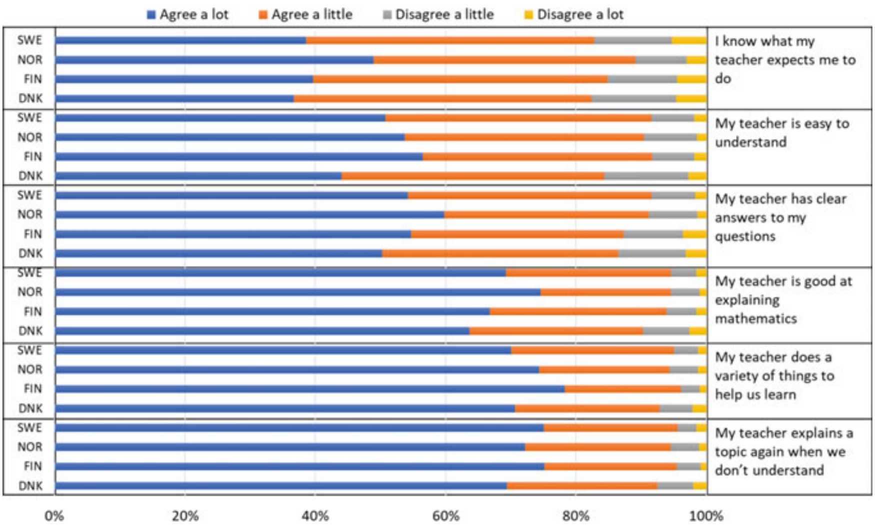
Classroom management (only in mathematics)

Student questionnaire: How often do these things happen in your mathematics lessons?



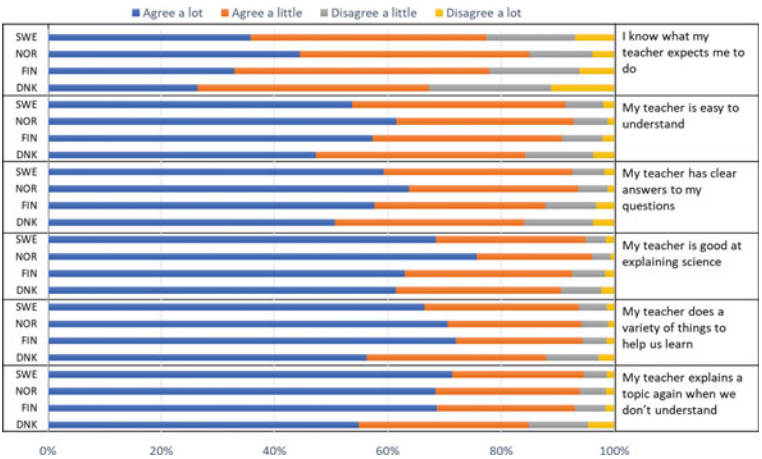
Teacher support and clarity of instruction in mathematics

Student questionnaire: How much do you agree with these statements about your mathematics lessons?



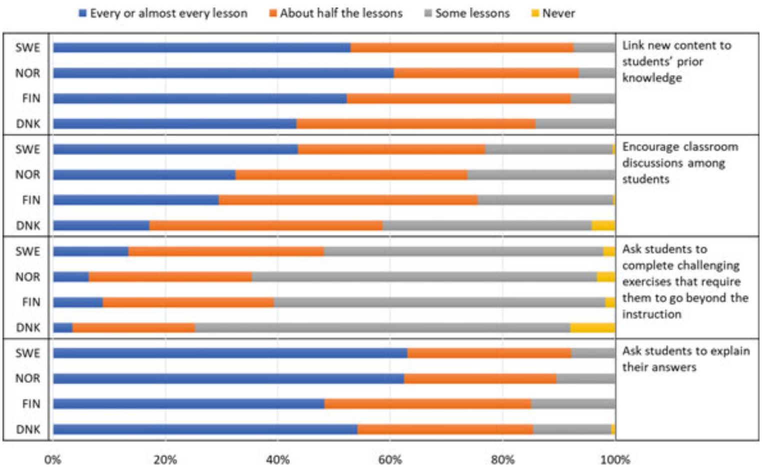
Teacher support and clarity of instruction in science

Student questionnaire: How much do you agree with these statements about your science lessons?



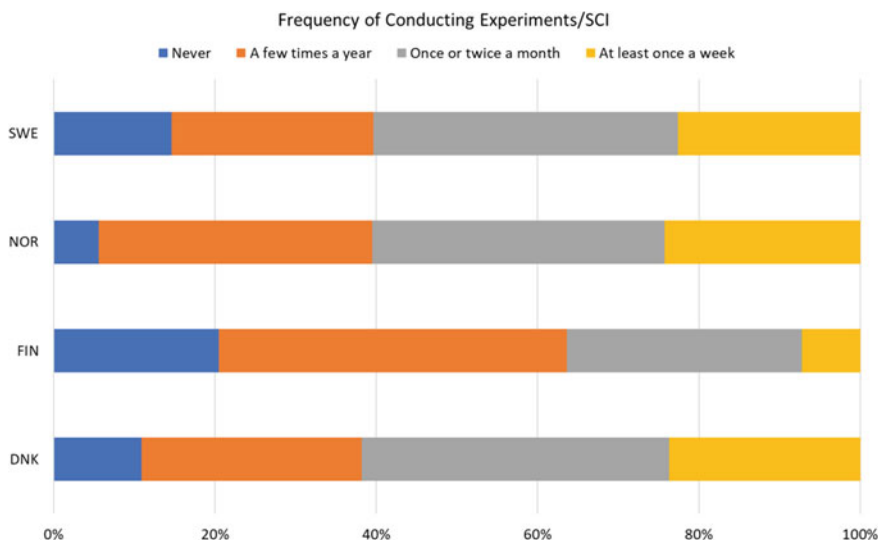
Cognitive activation (general)

Teacher questionnaire: In teaching mathematics/science to the students in this class, how often do you usually ask them to do the following?



Cognitive activation in science

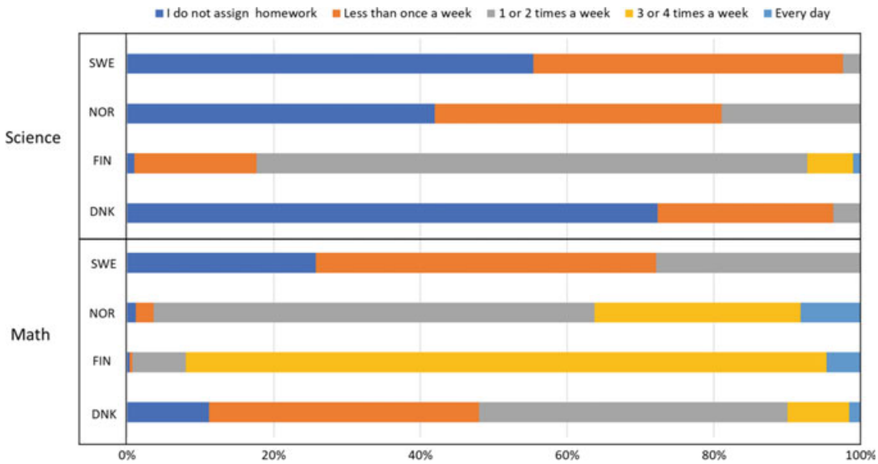
Student questionnaire: In science lessons, how often does your teacher ask you to conduct science experiments?



Appendix 2 Response Patterns of Assessment Practice Across Nordic Countries in TIMSS 2019

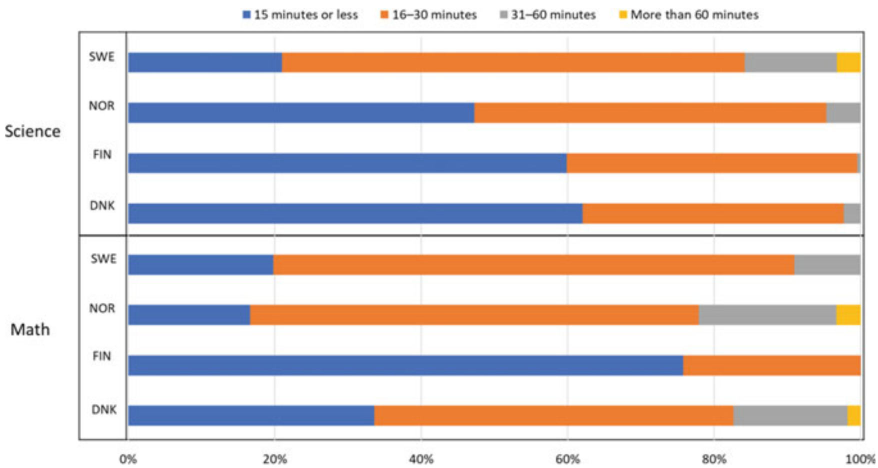
Homework frequency in mathematics and science

Teacher questionnaire: How often do you usually assign mathematics/science homework to the students in this class?



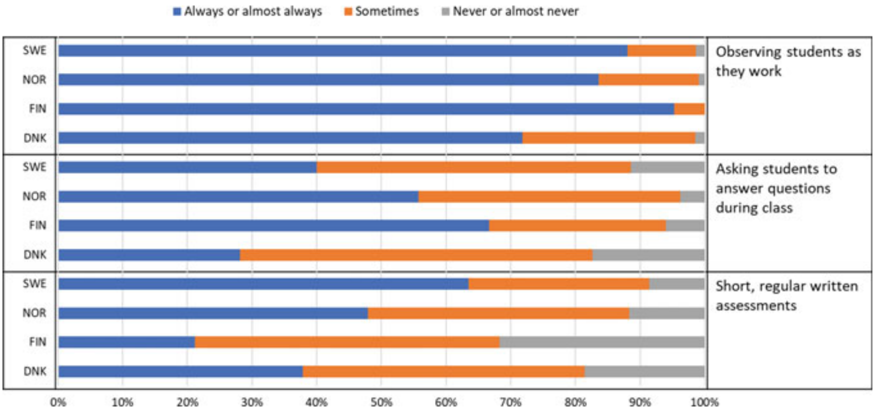
Duration of assigned homework in mathematics and science

Teacher questionnaire: When you assign mathematics homework to the students in this class, about how many minutes do you usually assign?



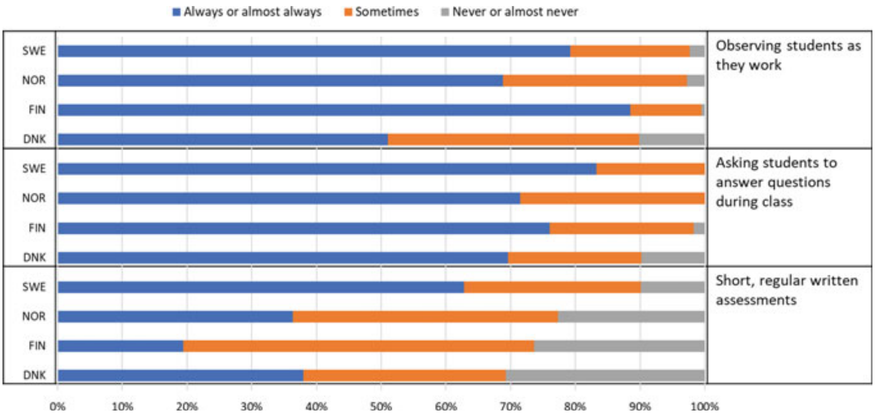
In-class homework discussion in mathematics

Teacher questionnaire: How often do you do the following with the mathematics homework assignments for this class?



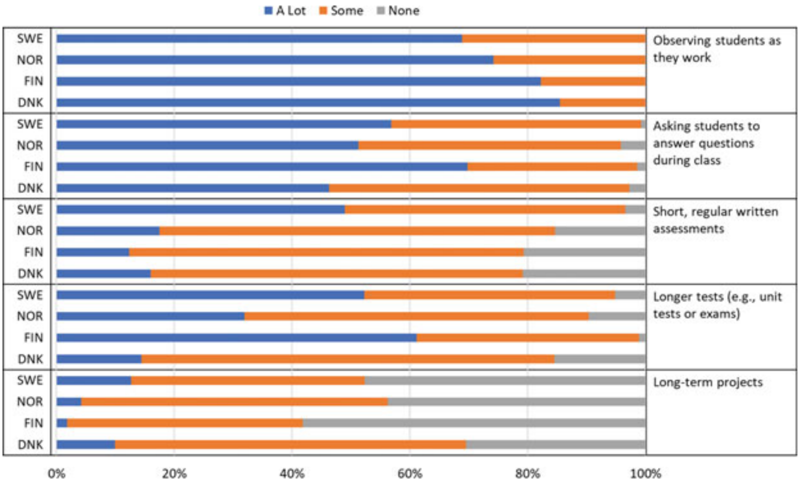
In-class homework discussion in science

Teacher questionnaire: How often do you do the following with the science homework assignments for this class?



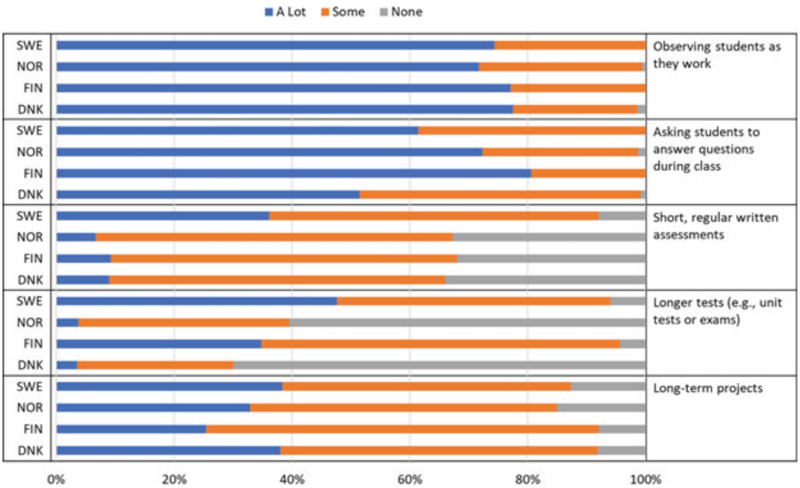
Teachers' emphasis on various assessment strategies in mathematics

Teacher questionnaire: How much emphasis do you place on the following sources to monitor students' progress in mathematics?



Teachers' emphasis on various assessment strategies in science

Teacher questionnaire: How much emphasis do you place on the following sources to monitor students' progress in mathematics?



References

Andersson, C., & Palm, T. (2017). The impact of formative assessment on student achievement: A study of the effects of changes to classroom practice after a comprehensive professional development programme. *Learning and Instruction, 49*, 92–102. https://doi.org/10.1016/j.learninstruc.2016.12.006

Andrade, H. L. (2019). A critical review of research on student self-assessment. *Frontiers in Education*, *4*. https://doi.org/10.3389/feduc.2019.00087

Andrade, H. L., & Brookhart, S. M. (2020). Classroom assessment as the co-regulation of learning. *Assessment in Education: Principles, Policy & Practice, 27*(4), 350–372. https://doi.org/10.1080/0969594x.2019.1571992

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S., Neubrand, M., & Tsai, Y.-M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal, 47*(1), 133–180. https://doi.org/10.3102/0002831209345157

Bellens, K., Van Damme, J., Van Den Noortgate, W., Wendt, H., & Nilsen, T. (2019). Instructional quality: catalyst or pitfall in educational systems' aim for high achievement and equity? An answer based on multilevel SEM analyses of TIMSS 2015 data in Flanders (Belgium), Germany, and Norway. *Large-scale Assessments in Education, 7*. https://doi.org/10.1186/s40536-019-0069-2

Bempechat, J., Li, J., Neier, S. M., Gillis, C. A., & Holloway, S. D. (2011). The homework experience: Perceptions of low-income youth. *Journal of Advanced Academics, 22*(2), 250–278. https://doi.org/10.1177/1932202X1102200204

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability (formerly: Journal of personnel evaluation in education), 21*, 5–31. https://doi.org/10.1007/s11092-008-9068-5

Brown, T. A. (2015). *Confirmatory factor analysis for applied research.* Guilford Publications.

Cairns, D. (2019). Investigating the relationship between instructional practices and science achievement in an inquiry-based learning environment. *International Journal of Science Education, 41*(15), 2113–2135. https://doi.org/10.1080/09500693.2019.1660927

Carlgren, I., Klette, K., Mýrdal, S., Schnack, K., & Simola, H. (2006). Changes in Nordic teaching practices: From individualised teaching to the teaching of individuals. *Scandinavian Journal of Educational Research, 50*(3), 301–326.

Charalambous, C. Y., & Praetorius, A.-K. (2020). Creating a forum for researching teaching and its quality more synergistically. *Studies in Educational Evaluation, 67*. https://doi.org/10.1016/j.stueduc.2020.100894

Clement, N. (2010). Student wellbeing at school: The actualization of values in education. In T. Lovat, R. Toomey, & N. Clement (Eds.), *International research handbook on values education and student wellbeing* (pp. 37–62). Springer Netherlands. https://doi.org/10.1007/978-90-481-8675-4_3

Creemers, B., & Kyriakides, L. (2008). *The dynamics of educational effectiveness: A contribution to policy, practice and theory in contemporary schools.* Routledge.

Creemers, B., Kyriakides, L., & Antoniou, P. (2012). Teacher professional development for improving quality of teaching. *Springer*. https://doi.org/10.1007/978-94-007-5207-8

Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). *Effective teacher professional development*. Learning Policy Institute.

Delahunty, T., Seery, N., & Lynch, R. (2020). Exploring problem conceptualization and performance in STEM problem solving contexts. *Instructional Science, 48*(4), 395–425. https://doi.org/10.1007/s11251-020-09515-4

Fan, H., Xu, J., Cai, Z., He, J., & Fan, X. (2017). Homework and students' achievement in math and science: A 30-year meta-analysis, 1986–2015. *Educational Research Review, 20*, 35–54. https://doi.org/10.1016/j.edurev.2016.11.003

Fernández-Alonso, R., & Muñiz, J. (2020). Homework: Facts and fiction. In T. Nilsen, A. Stancel-Piątak, & J.-E. Gustafsson (Eds.), *International handbook of comparative large-scale studies in education: Perspectives, methods and findings* (pp. 1–31). Springer International Publishing. https://doi.org/10.1007/978-3-030-38298-8_40-1

Fernández-Alonso, R., & Muñiz, J. (2022). Homework: Facts and fiction. In *International handbook of comparative large-scale studies in education* (pp. 1209–1239). Springer International Publishing. https://doi.org/10.1007/978-3-030-88178-8_40

Förtsch, C., Werner, S., Dorfner, T., von Kotzebue, L., & Neuhaus, B. J. (2017). Effects of cognitive activation in biology lessons on students' situational interest and achievement [journal article]. *Research in Science Education, 47*(3), 559–578. https://doi.org/10.1007/s11165-016-9517-y

Gardner, J., Harlen, W., Hayward, L., Stobart, G., & Montgomery, M. (2010). *Developing teacher assessment*. McGraw-Hill Education (UK).

Yang Hansen, K., Gustafsson, J.-E., & Rosen, M. (2014). School performance differences and policy variations in Finland, Norway and Sweden. In K. Yang Hansen, J.-E. Gustafsson, M. Rosen, S. Sulkunen, K. Nissinen, P. Kupari, R. F. Olafsson, J. K. Björnsson, L. S. Grønmo, L. Rønberg, J. Mejding, I. C. Borge, & A. Hole (Eds.), *Northern lights on TIMSS and PIRLS 2011: Differences and similarities in the Noordic countries* (pp. 23–45). Nordic Council of Ministers. https://doi.org/10.6027/TN2014-528

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.

Herppich, S., Praetorius, A.-K., Förster, N., Glogger-Frey, I., Karst, K., Leutner, D., Behrmann, L., Böhmer, M., Ufer, S., Klug, J., Hetmanek, A., Ohle, A., Böhmer, I., Karing, C., Kaiser, J., & Südkamp, A. (2018). Teachers' assessment competence: Integrating knowledge-, process-, and product-oriented approaches into a competence-oriented conceptual model. *Teaching and Teacher Education, 76*, 181–193. https://doi.org/10.1016/j.tate.2017.12.001

Kanjee, A. (2009). Enhancing teacher assessment practices in South African schools: Evaluation of the assessment resource banks. *Education as Change, 13*(1), 73–89. https://doi.org/10.1080/16823200902940599

Klieme, E. (2019). *Teaching quality. Conceptualization, measurement, and findings for European countries.* Keynote lecture presented at the EU "PISA and Beyond Conference", https://doi.org/10.13140/RG.2.2.11701.55529

Klieme, E., & Nilsen, T. (2022). *Teaching quality and student outcomes in TIMSS and PISA* (Edited by T. Nilsen, A. Stancel-Piątak, & J.-E. Gustafsson, pp. 1089–1134). Springer International Publishing. https://doi.org/10.1007/978-3-030-88178-8_37

Klieme, E., Pauli, C., & Reusser, K. (2009). The pythagoras study: Investigating effects of teaching and learning in swiss and german mathematics classrooms. In T. Janik & T. Seidel (Eds.), *The power of video studies in investigating teaching and learning in the classroom* (pp. 137–160). Waxmann.

Krzywacki, H., Pehkonen, L., & Laine, A. (2016). Promoting mathematical thinking in Finnish mathematics education. In H. Niemi, A. Toom, & A. Kallioniemi (Eds.), *Miracles of education* (Vol. 2, pp. 109–123). Sense Publishers. https://doi.org/10.1007/978-94-6300-776-4_8

Lipowsky, F., Rakoczy, K., Pauli, C., Drollinger-Vetter, B., Klieme, E., & Reusser, K. (2009). Quality of geometry instruction and its short-term impact on students' understanding of the pythagorean theorem. *Learning and Instruction, 19*(6), 527–537. https://doi.org/10.1016/j.learninstruc.2008.11.001

Luoto, J. M., Klette, K., & Blikstad-Balas, M. (2022). Patterns of instruction in Finnish and Norwegian lower secondary mathematics classrooms. *Research in Comparative and International Education, 17*(3), 17454999221077848. https://doi.org/10.1177/17454999221077848

Marder, J., Thiel, F., & Göllner, R. (2023). Classroom management and students' mathematics achievement: The role of students' disruptive behavior and teacher classroom management. *Learning and Instruction, 86*, 101746. https://doi.org/10.1016/j.learninstruc.2023.101746

Marsh, H. W., Hau, K.-T., & Grayson, D. (2005). Goodness of fit evaluation in structural equation modeling. In A. Maydeu-Olivares & J. J. McArdle (Eds.), *Contemporary Psychometrics* (pp. 275–340). Lawrence Erlbaum.

Martin, M. O., Mullis, I. V., Foy, P., & Stanco, G. M. (2016). *TIMSS 2015 international results in science*. http://timssandpirls.bc.edu/timss2015/international-results/

Morin, A. J. S., Marsh, H. W., Nagengast, B., & Scalas, L. F. (2014). Doubly latent multilevel analyses of classroom climate: An illustration. *The Journal of Experimental Education, 82*(2), 143–167. https://doi.org/10.1080/00220973.2013.769412

Mostafa, T., Echazarra, A., & Guillou, H. (2018). *The science of teaching science: An exploration of science teaching practices in PISA 2015*. OECD Education Working Papers, No. 188, OECD Publishing. https://doi.org/10.1787/f5bd9e57-en

Muijs, D. (2006). Measuring teacher effectiveness: Some methodological reflections. *Educational Research and Evaluation, 12*(1), 53–74. https://doi.org/10.1080/13803610500392236

Mullis, I. V. S., Martin, M. O., Foy, P., Kelly, D. L., & Fishbein, B. (2020). *TIMSS 2019 international results in mathematics and science*. TIMSS & PIRLS International Study Center, Boston College. https://timssandpirls.bc.edu/timss2019/

Muthen, L. K., & Muthen, B. O. (2022). MPlus (Version 8.5) [Computer software]. Muthen & Muthen. https://www.statmodel.com/index.shtml.

Nilsen, T., Gustafsson, J.-E., & Blömeke, S. (2016). Conceptual framework and methodology of this report. In T. Nilsen & J.-E. Gustafsson (Eds.), *Teacher quality, instructional quality and student outcomes: Relationships across countries, cohorts and time* (pp. 1–19). Springer International Publishing. https://doi.org/10.1007/978-3-319-41252-8_1

Nilsen, T., Scherer, R., & Blömeke, S. (2018). The relation of science teachers' quality and instruction to student motivation and achievement in the 4th and 8th grade: A Nordic perspective In A. Wester (Ed.), *Northern lights on TIMSS and PISA 2018* (pp. 61–90). Nordic Council of Ministers. https://norden.diva-portal.org/smash/get/diva2:1237833/FULLTEXT01.pdf

Palm, T., Andersson, C., Boström, E., & Vingsle, C. (2017). A review of the impact of formative assessment on student achievement in mathematics. *Nordic Studies in Mathematics Education, 22*(3), 25–50.

Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review, 22*, 74–98. https://doi.org/10.1016/j.edurev.2017.08.004

Praetorius, A.-K., Klieme, E., Herbert, B., & Pinger, P. (2018). Generic dimensions of teaching quality: The German framework of three basic dimensions. *ZDM Mathematics Education , 50*(3), 407–426. https://doi.org/10.1007/s11858-018-0918-4

Remmen, K. B., & Iversen, E. (2022). A scoping review of research on school-based outdoor education in the Nordic countries. *Journal of Adventure Education and Outdoor Learning*, 1–19. https://doi.org/10.1080/14729679.2022.2027796

Rohatgi, A., Hatlevik, O. E., & Björnsson, J. K. (2022). Supportive climates and science achievement in the Nordic countries: Lessons learned from the 2015 PISA study. *Large-Scale Assessments in Education, 10*(1), 12. https://doi.org/10.1186/s40536-022-00123-x

Rønning, M. (2011). Who benefits from homework assignments? *Economics of Education Review, 30*(1), 55–64. https://doi.org/10.1016/j.econedurev.2010.07.001

Ryu, E. (2014). Model fit evaluation in multilevel structural equation models. *Frontiers in Psychology, 5*, Article 81. https://doi.org/10.3389/fpsyg.2014.00081

Sass, D. A., & Schmitt, T. A. (2013). Testing measurement and structural invariance. In T. Teo (Ed.), *Handbook of quantitative methods for educational research* (pp. 315–345). SensePublishers. https://doi.org/10.1007/978-94-6209-404-8_15

Scheerens, J. (2016). *Educational effectiveness and ineffectiveness : A critical review of the knowledge base* (1st ed.). Springer. https://doi.org/10.1007/978-94-017-7459-8

Senden, B., Nilsen, T., & Teig, N. (2023). The validity of student ratings of teaching quality: Factorial structure, comparability, and the relation to achievement. *Studies in Educational Evaluation, 78*, 101274. https://doi.org/10.1016/j.stueduc.2023.101274

Sigurjónsson, J. Ö. (2023). *Quality in Icelandic mathematics teaching: Cognitive activation in mathematics lessons in a Nordic context*. University of Iceland.

Sigurgeirsson, I., & Björnsdóttir, A. (2016). Heimanám í íslenskum grunnskólum-Scope and vidhorf for students, parents and teachers [Homework in Icelandic Compulsory schools: Its Amount and Students', Parents' and Teachers' Attitudes]. Netla. https://ojs.hi.is/index.php/netla/article/view/2384

Sortkær, B. (2019). Feedback for everybody? Exploring the relationship between students' perceptions of feedback and students' socioeconomic status. *British Educational Research Journal, 45*(4), 717–735. https://doi.org/10.1002/berj.3522

Svanbjörnsdóttir, B., Zophoníasdóttir, S., & Gísladóttir, B. (2023). Quality of the stated purpose and the use of feedback in Icelandic lower-secondary classrooms results from a video study. *Teaching and Teacher Education, 121*, 103946. https://doi.org/10.1016/j.tate.2022.103946

Teig, N., Bergem, O. K., Nilsen, T., & Senden, B. (2021). Gir utforskende arbeidsmåter i naturfag bedre læringsutbytte? [Does inquiry-based teaching practice in science provide better learning outcomes?]. In T. Nilsen & H. Kaarstein (Eds.), *Med blikket mot naturfag [A view towards science]* (pp. 46–72). Universitetsforlaget.

Teig, N., & Nilsen, T. (2022). Profiles of instructional quality in primary and secondary education: Patterns, predictors, and relations to student achievement and motivation in science. *Studies in Educational Evaluation, 74*, 101170. https://doi.org/10.1016/j.stueduc.2022.101170

Teig, N., Scherer, R., & Nilsen, T. (2018). More isn't always better: The curvilinear relationship between inquiry-based teaching and student achievement in science. *Learning and Instruction, 56*, 20–29. https://doi.org/10.1016/j.learninstruc.2018.02.00

Teig, N., Scherer, R., & Nilsen, T. (2019). I know I can, but do I have the time? The role of teachers' self-efficacy and perceived time constraints in implementing cognitive-activation strategies in science. *Frontiers in psychology, 10*(1697). https://doi.org/10.3389/fpsyg.2019.01697

Teig, N., & Steinmann, I. (2023). Leveraging large-scale assessments for effective and equitable school practices: The case of the nordic countries. *Large-Scale Assessments in Education, 11*(21). https://doi.org/10.1186/s40536-023-00172-w

Wang, S., Rubie-Davies, C. M., & Meissel, K. (2018). A systematic review of the teacher expectation literature over the past 30 years. *Educational Research and Evaluation, 24*(3–5), 124–179. https://doi.org/10.1080/13803611.2018.1548798

Wolff, C. E., Jarodzka, H., & Boshuizen, H. P. A. (2021). Classroom management scripts: A theoretical model contrasting expert and novice teachers' knowledge and awareness of classroom events. *Educational Psychology Review, 33*(1), 131–148. https://doi.org/10.1007/s10648-020-09542-0

Nani Teig is an associate professor at the University of Oslo, Norway. Her research focuses on inquiry-based teaching, scientific reasoning, teaching quality, and academic resilience. She integrates multilevel analyses using data from videos, surveys, assessments, and computer log files. Dr. Teig has received several awards and fellowships, including the Global Education Award, Bruce H. Choppin Dissertation Award, Young CAS Fellow, and UNESCO GEM Fellow.

Jennifer Maria Luoto (born 1988) is a Postdoctoral fellow at the Department of Teacher Education and School Research, University of Oslo. Her research focuses mainly on classroom observations and teaching quality in different classrooms and national contexts.