

Exploring, understanding, and problematizing patterns of  
instructional quality:

A study of instructional quality in Finnish–Swedish and Norwegian  
lower secondary mathematics classrooms

Jennifer Maria Luoto



Thesis submitted for the degree of Philosophiae Doctor

Department of Teacher Education and School Research  
Faculty of Educational Sciences  
University of Oslo  
2020

© Jennifer Maria Luoto, 2021

*Series of dissertations submitted to the  
Faculty of Educational Sciences, University of Oslo*  
No. 327

ISSN 1501-8962

All rights reserved. No part of this publication may be  
reproduced or transmitted, in any form or by any means, without permission.

Cover: Hanne Baadsgaard Utigard.  
Print production: Representralen, University of Oslo.

## Acknowledgements

When I moved to Norway from Finland in 2011, I had no ambitions to pursue a PhD. However, due to the changing circumstances in life, I got the opportunity four years ago when I became part of the LISA team as one of four PhD candidates. Now, the period is coming to its end, and I want to thank all the people who have supported me along the way.

First, I want to thank the researchers behind the LISA study: Project leader Professor Kirsti Klette, senior researcher Astrid Roe, and Professor Marte Blikstad-Balas for letting me take part in this exciting project and for all the work you did making the LISA study possible.

My supervisors Kirsti Klette and Marte Blikstad-Balas—you are two sharp, ambitious, and inspiring women, and I will always be grateful for having had you as my supervisors, guiding me through this sometimes very frustrating PhD journey. Kirsti, many thanks to you for always believing in me and for being so kind and inclusive. Marte, thank you for your optimism and your helpful and constructive advice on how to navigate the PhD time. Thank you both for all the support, discussions, and comments on my work and for the co-authorship in Articles I and III.

Nils Buchholtz, thank you for your support along the way, and especially for your constructive comments during the final evaluation seminar. Guri Nortvedt and Jelena Radišić, thank you for taking the time to read and comment on my work and for sharing your mathematical expertise. Mark White, thank you for your straightforward and constructive comments on my work.

Thank you to my research group, SISCO, and its leaders, for creating a constructive and safe environment where new researchers can develop and share their work. Thank you especially to Lisbeth Brevik, Inga Staal Jensen, and Joke DeWilde, but also to all the other SISCO members, for your careful readings and all the helpful comments on my work throughout these years.

Thank you, PhD coordinator Glenn Ole Hellekjær for your calmness and down-to-earth advice during stressful times. Thank you also PhD coordinator Lisbeth Brevik, for being so helpful during the sudden COVID-19 situation at the end of my PhD period, assisting me in applying for extended time to be able to finish my PhD while combining home–office, home–school, and home–kindergarten demands.

A big thank you to Roar Bakken Stovner for helping me navigate the field of mathematics education and all the cooperation we have had through VIST, while writing a book chapter, and in seminars during these years.

Thank you to my fellow LISA PhDs Camilla Magnusson, Roar Bakken Stovner, and Ida Lodding Gabrielsen for all the critical and engaging discussions, for your team spirit, and for sharing laughs, as well as tears. You are all incredible people, and I am very lucky to call you my friends!

Also, a big thank you to Professor Hilda Borko at Stanford University, who on several occasions took the time to visit the LISA project and give us PhD students clear and constructive advice.

I also thank the Department of Teacher Education and School Research for the opportunity to pursue a PhD. I send a big thank-you to the Center of Excellence Justice for Education (JustEd) for organizing seminars and making it economically possible for me to conduct fieldwork in Finland. I also want to thank all the scholars connected to the Center of Excellence Quality in Nordic Teaching (QUINT) for the exciting and stimulating discussions during conferences and seminars.

Thank you to the teachers and students in Helsinki and Oslo for participating in the LISA study. A special thanks go to the Helsinki teachers who, during my fieldwork, invited me to follow their lessons and made time for interviews and discussions. Thank you also to Bjørn, Astrid, and Kirsti for your company and assistance during filming in Finland.

Thank you Bjørn, Torgeir, Kenneth, and the rest of the ILS-IT crew for your technical support and availability!

I also want to thank my friends Alessandra, Ceren, and Karina, for sharing the good times and the bad times, as well as your pep talks and PhD jokes. Thank you Pia, Malin, and Emma and your families, for our play-dates and girl-nights, and for keeping Finland alive. Thank you to my family and friends in Finland, especially my mother for her big heart as, even though we live in different countries, you flew to Oslo to help take care of my children every time you could. Furthermore, thank you my dear friends Lotta and Karin for always being there. Thank you to my neighbors and friends in Holmlia, for your community spirit and support, who have always made my family and me feel welcome.

Finally, I want to thank my beautiful children Klara and Axel for your love and your smiles always reminding me what is important in life, and say I am sorry that you, at times, had to deal with a much-stressed mother. And thank you Alex for being my rock and keeping it together during these sometimes intense years.

Jennifer Maria Luoto  
Blindern, August, 2020



## Summary of the thesis

What goes on in the mathematics classroom is something that engages educational researchers across the globe. Recently, many have taken an interest not only in investigating instruction in a specific national context, but in comparing how mathematics is taught in different countries. This article-based PhD thesis, situated in the field of cross-national observation research on instructional quality in mathematics education, investigates patterns of instructional quality in lower secondary mathematics classrooms in Helsinki and Oslo contexts in three articles using three different lenses. This thesis is part of the video study Linking Instruction and Student Achievement (LISA).

Article I is a cross-national classroom video study taking a comparative approach by *exploring* patterns of instructional quality in eight Nordic mathematics classrooms in both Helsinki and Oslo, respectively ( $N = 16$ ), using the standardized observation system Protocol for Language Arts Teacher Observation (PLATO) to decompose instructional quality into observable behavior that enable comparative and systematic analyses. The results in Article I demonstrated distinct patterns of instructional quality in the two contexts related to activity formats, presentation of content, and classroom discourse. Instruction in the Helsinki classrooms is characterized by individual seatwork and teacher-led whole-class discussions, where teachers often clearly explain and connect content as well as frame purposes for learning throughout the lessons, but there are few opportunities for students to communicate and/or collaborate with peers. Instruction in the Oslo classrooms is characterized by more variety in terms of activity formats (i.e., whole class instruction, individual seatwork, and group work), and students have more opportunities to engage in content-related discussions, while content is less explicitly explained, connected, and framed in terms of learning purposes. These differences across the two contexts indicate distinct patterns of instructional quality that are possibly shaped by various contextual factors such as mathematics education tradition, curriculum, and educational policies. Even though the sample was too small to generalize to a greater Finnish and Norwegian context, previous research has indicated similar patterns, which strengthen the conclusions about possibly distinct patterns of instruction and instructional quality. This article informs researchers, teacher educators, and policy makers about what goes on in the mathematics classrooms in these two contexts.

To further *understanding* of the limited use of student participation in classroom discourse in the Helsinki context, I conducted a case study, reported in Article II, scrutinizing the instructional rationales of Anna and Bea—two purposefully sampled teachers’ from the Helsinki sample in Article I for their differently enacted discourse practices—drawing on observation and interview data. This article was conducted in light of reform-oriented mathematics education literature emphasizing discourse and talking mathematics as well as recent curricula reform in Finland underscoring that mathematics instruction should provide students with opportunities to engage in content-related discourse. The findings showed that the teachers, who enacted very different classroom discourse patterns—with Anna continually providing students opportunities to discuss and Bea only providing students opportunities to participate in strict teacher-led discussions—rationalized their discourse practices with similar concerns. Bea perceived a tension between engaging students in discourse practices and her concerns for student learning, student well-being, and equity, while Anna embraced reform-oriented views of the benefits of student engaging discourse practices and created activities for giving all her students access to mathematical discussions. This insight into how teachers rationalize their enacted discourse practice in a Finnish context may be useful for teacher educators when developing and promoting mathematical classroom discourse practices as reflected in the new curriculum.

Inspired by the first and the second articles, Article III takes a theoretical and methodological approach *problematizing* how standardized observation measures embed possible biases when scoring instructional quality in different classroom contexts. Using empirical examples from Helsinki and Oslo classrooms, this article discusses how the observation system PLATO conceptualizes, operationalizes, and sequences instructional quality, and it illustrates how contextual factors such as lesson structure may influence scoring in a way that risks misrepresentation of the intended construct. This article stresses that transparency of possible biases embedded in observation systems is crucial for valid interpretations in classroom research, as conclusions about instructional quality otherwise might be misleading. This insight is relevant for increasing knowledge about how different observation systems might work in Nordic mathematics education contexts, as well as for anyone using and/or developing classroom observation systems both within and across national contexts.

Taking all three articles together, this thesis contributes updated knowledge and interpretations on patterns of instructional quality in the Helsinki and Oslo contexts as well as



how such patterns may be understood when viewed through the lenses of standardized observation systems and teachers' perspectives.



## **PART I: Extended abstract**

1 Introduction and rationale .....	1
1.1 Overarching aim and research questions .....	3
1.2 Relationship among the three articles.....	5
1.3 Key concepts.....	5
1.3.1 Instructional quality and patterns of instructional quality .....	5
1.3.2 Observation systems .....	6
1.3.3 Instructional rationale .....	6
1.3.4 Context and context-sensitivity .....	6
1.4 The Finnish and Norwegian educational contexts.....	7
1.4.1 School system and teacher education .....	7
1.4.2 The National Core Curriculum for Basic Education and The Knowledge Promotion ...	7
1.5 Overview of the thesis .....	8
2 Review of literature.....	9
2.1 Mathematics instruction in Finland .....	9
2.2 Mathematics instruction in Norway.....	10
2.3 Cross-national classroom observation studies of mathematics instruction .....	11
2.4 Different conceptualizations of instructional quality.....	14
2.5 Summary of the research review .....	15
3 Theoretical perspectives on instructional quality .....	17
3.1 Process-product tradition .....	17
3.2 Cognitive perspectives.....	18
3.3 Constructivist and socio-constructivist perspectives .....	19
3.4 Sociocultural perspectives .....	20
3.4.1 Learning through discourse .....	21
3.5 Integrating perspectives when studying instructional quality .....	22
3.6 PLATO's conceptualization of instructional quality .....	23
3.7 Current trends and debates in conceptualizing instructional quality .....	26
3.8 Summary of theoretical perspectives in the thesis.....	27

4 Methodology and research design .....	28
4.1 Situating this thesis within the Linking Instruction and Student Achievement study .....	28
4.1.1 Research design of Article I .....	29
4.1.2 Research design of Article II .....	29
4.1.3 Research design of Article III .....	30
4.2 Sampling .....	31
4.3 Data sources .....	32
4.3.1 Video observations .....	33
4.3.2 Non-participant observation and field notes .....	34
4.3.3 Semi-structured interviews .....	34
4.4 Analytical procedures .....	35
4.4.1 Protocol for Language Arts Teacher Observation (PLATO) .....	35
4.4.2 Teacher moves framework .....	37
4.4.3 Analytical procedure in Article III .....	39
4.5 Research credibility .....	39
4.5.1 Reliability of standardized observations .....	39
4.5.2 Validity of standardized cross-national observations .....	42
4.5.3 Reliability and validity of interviews and non-participant observations .....	43
4.5.4 Representativeness and generalization .....	45
4.6 Ethical considerations .....	46
4.6.1 Evaluative frameworks in cross-national research .....	47
4.7 Methodological limitations and strengths .....	47
5 Summary of the articles .....	49
5.1 Article I .....	49
5.2 Article II .....	51
5.3 Article III .....	52
6 Discussion of research contribution .....	54
6.1 Empirical contribution .....	54
6.2 Theoretical contribution .....	57
6.3 Methodological contribution .....	58
6.4 Suggestions for future research .....	59

6.5 Concluding remarks.....	59
References.....	61
Appendices.....	76
Appendix 1: PLATO rubrics.....	76
Appendix 2: Overview of sample in Articles I and III .....	79
Appendix 3: Analytical framework of discourse moves.....	80
Appendix 4: Interview guide .....	81
Appendix 5: Informed consent, Helsinki 2016.....	83
Appendix 6: Informed consent, Helsinki 2018.....	86
Appendix 7: Informed consent, Oslo 2014–2015 .....	91

## **PART II: Articles**

### **Article I**

Luoto, J. M., Klette, K., & Blikstad-Balas, M. (in review). Patterns of instructional quality in Finnish and Norwegian lower secondary mathematics classrooms. In review at *International Journal of Science and Mathematics Education*.

### **Article II**

Luoto, J. M. (2020). Scrutinizing two Finnish teachers' instructional rationales and perceived tensions in enacting student participation in mathematical discourse. *LUMAT: International Journal on Math, Science and Technology Education*, 8(1), 133–161. doi:10.31129/LUMAT.8.1.1329

### **Article III**

Luoto, J. M., Klette, K., & Blikstad-Balas. (in review). Possible biases in observation systems when applied across contexts: Conceptualizing, operationalizing and sequencing instructional quality. In review at *Educational Assessment, Evaluation and Accountability*.



# 1 Introduction and rationale

The quality of teachers' instruction has been found to have significant impact on student learning (Baumert et al., 2010; Klieme, Pauli, & Reusser, 2009; Nye, Konstantopoulos, & Hedges, 2004; Seidel & Shavelson, 2007). That is why assuring the quality of instruction is of great concern internationally, and over the past decades we see a growing interest in observation approaches when trying to understand why and how teachers instructional practices make a difference. Drawing on a large data-corpus with video-recorded mathematics lessons from the Linking Instruction and Student Achievement (LISA) study, this thesis is situated in the field of cross-national observation research on instructional quality in mathematics education. A key ambition of the thesis has been to systematically investigate patterns of instructional quality in lower secondary mathematics classrooms in Helsinki and Oslo, in order to provide new insight and nuance about instruction in these two contexts.

To systematically observe and measure the highly complex concept of instructional quality, and ultimately improve it, researchers have developed standardized *observation systems* (Bell, Dobbelaer, Klette, & Visscher, 2019) drawing on knowledge from previous research about effective teaching practices. Such systems thus attempt to decompose instructional quality into observable entities to facilitate systematic coding. They can be used to study instructional quality within national contexts (e.g., Kane & Staiger, 2012) as well as cross-nationally (see Praetorius, Rogh, Bell, & Klieme, 2019). However, observation systems necessarily prioritize certain features of instruction and exclude others, and are challenged when it comes to the comparison of instruction across different classrooms or national contexts. Ultimately, the interpretation of the data derived from observation systems and the observation systems themselves might profit from complimentary perspectives. Therefore, this thesis investigates patterns of instructional quality in a sample of video-recorded Finnish–Swedish and Norwegian mathematics classrooms using three different lenses: (i) by systematically comparing and exploring patterns of instructional quality using a standardized observation system, (ii) by focusing on understanding two Finnish–Swedish mathematics teachers' own instructional rationales for diverging classroom discourse patterns observed in their lessons, and (iii) by critically assessing possible biases in a standardized system of instructional quality when applied across different classroom contexts.

There are many benefits in using multiple lenses when analyzing instructional quality. First of all, a comparative lens can “make the familiar strange” (Alexander, Broadfoot, & Phillips, 2000) by illuminating practices that otherwise are taken for granted and, in turn,

increase knowledge of one's own practices that may otherwise not be discernable (Stigler, Gallimore, & Hiebert, 2000). Comparative studies on instruction may thus offer important insights into different contexts, as they can illuminate “ways in which teaching is powerfully shaped by contextual factors, including policies and material conditions, institutional norms, and cultural practices and beliefs” (Paine, Blömeke, & Aydarova, 2016, p. 732).

Finland and Norway are particularly interesting contexts to compare, as the Finnish education system has gained special interest among its Nordic neighbors (e.g., Breakspear, 2012), due to their having similar education systems (Blossing, Imsen, & Moos, 2014) yet differences in international comparisons of student achievement. Finnish students have consistently scored higher in Programme for International Student Assessment (PISA) tests by Organisation for Economic Co-operation and Development (OECD; 2004, 2016b)—however less so in recent years (OECD, 2019). Particularly in Norway, the Finnish system has been used as a benchmark for justifying policy suggestions (Østerud, 2016), on everything from homework to free school lunches. The interest in “what happens in Finland” is high, but there has been little comparative work systematically looking into similarities and differences in how teachers across these two contexts actually enact their teaching. Thus, the political use of global rankings highlights a need for empirical and comparative studies to nuance and inform educational debates, in Finland and Norway and beyond, about actual classroom practice (Simola, Kauko, Varjo, Kalalahti, & Sahlström, 2017).

Second, an in-depth lens increases understanding of features of mathematics instruction in a specific context, enabling detailed analyses of how classroom instruction is shaped by contextual factors such as teachers and their students, as well as in the society, which is often neglected in mathematics education research (Chazan, Herbst, & Clark, 2016; Skott, 2019). One way of taking an in-depth approach is to focus on how teachers rationalize their enacted practices, and in one of the articles in this thesis, I focused on how two Finnish mathematics teachers think and rationalize their enacted classroom discourse, captured through video observations, in their particular context.

Third, educational research and theories of teaching and learning is often driven by unchallenged theoretical concepts (Simola et al., 2017). A critical lens can challenge such concepts, and an example of this is the concept of “instructional quality,” which has mostly been influenced by Western research on relationships between learning and teaching (Blömeke, Olsen, & Suhl, 2016). While conceptualizations of instructional quality embedded in observation systems often derive from American and central European contexts (Praetorius & Charalambous, 2018), we know little about the potential conceptual and



methodological biases in observation systems when applied across different classroom contexts. Thus, one of the articles discussed in this thesis focused on possible biases in scoring procedures and conceptualization of instructional quality in standardized observation systems when applied in Finnish and Norwegian mathematics classroom contexts. Awareness of such biases might enable researchers to better interpret assessments of instructional quality in a specific context, as well as improve observation systems so they facilitate comparable data and knowledge aggregation across contexts. Additionally, this might help to build a “Nordic perspective” of instructional quality, particularly relevant for comparing and identifying mathematics instruction within Nordic contexts.

### **1.1 Overarching aim and research questions**

The overarching aim of the thesis is to explore, understand, and problematize patterns of instructional quality in lower secondary mathematics classrooms, drawing on video data from the capital areas of Finland and Norway (i.e., Helsinki and Oslo). The three different lenses—exploring, understanding, and problematizing—each reflect one of the three studies, respectively: *exploring* differences and similarities in patterns of instructional quality across the two contexts in Article I; *understanding* teachers’ rationales and perceived tensions related to their instructional discourse patterns in Article II; and *problematizing* patterns of instructional quality produced when applying a standardized observation manual across Helsinki and Oslo mathematics classrooms in Article III. The overarching research question for the whole study is: *How can patterns of instructional quality be understood through observation systems and teachers’ perspectives?* I conducted the following three sub-studies to meet this aim and answer this question.

*Patterns of instructional quality* in two Nordic contexts, Helsinki and Oslo, is the main topic of Article I, which investigates the following three research questions: *What activity formats do teachers use to engage students? What is the quality of instructional explanations of content, connections to prior knowledge, and setting a purpose for learning?* and *What characterizes discourse features in mathematics classrooms?* This was a descriptive cross-national study designed to answer the overarching research question by exploring patterns of instructional quality drawing on systematically analyzed video observations from 16 lower secondary mathematics classrooms, eight from the Oslo area and eight from the Helsinki area, from three to four consecutive lessons from each classroom ( $N = 47$  lessons). I applied the observation system *Protocol for Language Arts Teacher Observation* (PLATO; Grossman, 2015) as the analytical lens. I focused on two dimensions of instructional quality, *presentation of content* and *discourse features*, as well as *activity format*, that is, how teachers structure

their lesson activities. This article is in review at *International Journal of Science and Mathematics education*.

Luoto, J. M., Klette, K., & Blikstad-Balas, M. (in review) Patterns of instructional quality in Finnish and Norwegian lower secondary mathematics classrooms.

*Teachers' instructional rationales for differently enacted discourse practices* was the topic in Article II, which was selected to shed light on how discourse patterns may be understood from teachers' perspectives. The main research question was *How do two Finnish mathematics teachers with diverging practices perceive and enact student participation in discourse?* For Article II, I gathered and combined data from video observations and teacher interviews. This article drew on the findings from Article I of Helsinki teachers' scant use of student engaging classroom discourse practices—a key feature of instructional quality. The participants were two purposefully sampled Finnish–Swedish mathematics teachers' who had been involved in Article I: Anna, who constantly provided students with opportunities to participate in discourse, and Bea, who had a more typical practice of rarely providing students with such opportunities. The article is published as:

Luoto, J. M. (2020). Scrutinizing two Finnish teachers' instructional rationales and perceived tensions in enacting student participation in mathematical discourse. *LUMAT: International Journal on Math, Science and Technology Education*, 8(1), 133–161. <https://doi.org/10.31129/LUMAT.8.1.1329>

*Possible biases in standardized observation systems when applied across different national and classroom contexts* was the topic for the study in Article III. The main research question investigated in this article was: *What possible biases are embedded in the conceptualization, operationalization, and scoring procedures of a specific observation system attempting to measure instructional quality?* This was a theoretical and methodological discussion approaching the overarching research question from a critical perspective by problematizing the way the PLATO observation system conceptualizes, operationalizes, and sequences instructional quality. The article drew on empirical examples derived from video data of 47 Helsinki and Oslo mathematics lessons and the PLATO scoring patterns these lessons produced. The article is in review at *Educational Assessment, Evaluation and Accountability*.

Luoto, J. M., Klette, K., & Blikstad-Balas, M. (in review). Possible biases in observation systems when applied across contexts: Conceptualizing, operationalizing, and sequencing instructional quality.

This thesis, the abovementioned articles, and all the data are part of the LISA study (see Section 4.1), funded by the Norwegian Research Council (Grant #222620/F10).

## **1.2 Relationship among the three articles**

The three articles are related in many ways. In Article I, I explore and compare three interrelated features of instruction in Helsinki and Oslo contexts: presentation of content, classroom discourse, and activity formats. Based on results in Article I, classroom discourse in Helsinki classrooms became the topic for Article II, as I wanted to understand teachers' perspectives of their scarce use of student engaging discourse. I also purposefully sampled teachers for Article II from Article I, based on the criterion of teachers providing students with very different discourse opportunities. Article II and Article III are connected in the sense that both studies used contextual factors (i.e., teacher rationales and lesson structure) to understanding patterns of instructional quality. The PLATO observation system as an analytical framework connects Articles I and III, as I in Article I apply it descriptively, while in Article III, I discuss it critically. In addition, the theoretical and methodological discussion in Article III draws partly on experiences of using PLATO described in Article I. Sequencing lessons for coding purposes connects all articles. Sequencing is a topic in Article III where we discuss issues with standardized time segments (i.e., 15-minute segments in PLATO), and this discussion builds on reflections made in Article I about applying PLATO, as well as in Article II where I sequence chunks of classroom discourse episodes with no set time frame.

## **1.3 Key concepts**

In this section, I will present the following key concepts of this thesis: instructional quality and patterns of instructional quality, observation systems, and instructional rationale as well as context and context-sensitivity.

### *1.3.1 Instructional quality and patterns of instructional quality*

*Instructional quality* specifically refers to aspects of teachers' instructional practice that mediates students' opportunities to learn, and is thus assumed to be positively related to some kind of student outcome (e.g., Nilsen & Gustafsson, 2016; Schoenfeld, 2016). Several different observation systems have been developed to capture teachers' instructional quality in the mathematics classroom (for an overview, see Praetorius & Charalambous, 2018). However, there is no universal agreement on what constitutes instructional quality (Praetorius et al., 2019), and how we define this construct, and what system we choose, depends on what purpose we have (Ball & Hill, 2009). In this thesis, the purpose of applying the PLATO observation system (see Section 4.4.1) is to capture and possibly differentiate patterns of instructional quality across contexts with an observation system corresponding to key features of instruction highlighted in the literature of effective teaching as well as in mathematics education literature.

*Patterns of instructional quality* refer to patterns of instruction produced by systematic coding, resulting in numeric and textual representations of instructional quality. In Article I, we used selected PLATO elements to identify and compare patterns of instructional quality in Helsinki and Oslo mathematics classrooms. In Article II, I used a framework targeted to capture patterns of teachers' discourse moves (see Section 4.4.2), and classroom discourse is one dimension of quality instruction (see Section 3.4.1). In Article III, we problematized patterns of instructional quality by exemplifying how, in some contexts, the production of such patterns may be a result of potential biases in scoring procedures and sequencing of standardized observation systems.

### *1.3.2 Observation systems*

*Observation systems* (see Bell et al., 2019; Hill, Charalambous, & Kraft, 2012), often termed observation manuals, are used to capture features of instruction through video or live observations. Bell et al. (2019, pp. 4–5) defined observation systems as comprising *scoring tools*, which specify the dimensions of teaching that will be measured, and scales for scoring instruction, such as *rating quality procedures*, specifying rater training, and *sampling specifications*, including the number of observations, length of time, frequency of scoring, and how lessons are sampled. Observation systems/manuals are thus much more than scoring rubrics and used as synonyms in this thesis.

### *1.3.3 Instructional rationale*

In Article II, I used the term *instructional rationale*, referring to teachers' stated views of factors that shape their instructional decisions of enacted classroom discourse practices. The term is based on previous research on factors shaping classroom discourse, as well as grounded in the empirical data of what the teachers perceive as shaping their practice.

### *1.3.4 Context and context-sensitivity*

*Context* is central in all three articles, yet operating on and across different levels. In Article I, context refers to the observed Helsinki and Oslo mathematics classrooms as well as to national contexts in which the classrooms are embedded. In Article II, context refers primarily to the context of the classrooms (i.e., teachers and their students), yet is framed within the specific Finnish context. In Article III, *context-sensitivity* refers to whether features of instructional quality are operationalized in a robust and broad enough way to capture features of instructional quality in Helsinki and Oslo classrooms and the ways teachers in these contexts structure their lessons (Praetorius, Rogh, Bell, & Klieme 2019). Another meaning of context-sensitivity discussed mainly in Section 6.2, and partly in Article III, concerns whether

conceptualizations of instructional quality reflect, for example, policies or teacher perspectives of quality instruction relevant to specific country and/or classroom contexts (see also Section 2.4).

#### **1.4 The Finnish and Norwegian educational contexts**

This thesis focuses on patterns of instructional quality in mathematics classrooms within two Nordic contexts, Finland and Norway. In what follows, I concisely summarize key educational aspects of the school system, teacher education, and the curricula.

##### *1.4.1 School system and teacher education*

Finland and Norway subscribe to the *Nordic model of education*—a model established in the 1960s based on a public non-streamed comprehensive school for all children regardless of social background (Blossing et al., 2014; Lundahl, 2016). In Finland, comprehensive school includes grades 1–9, and children start school the year they turn seven. In Norway, comprehensive school comprises grades 1–10, and children start school the year they turn six. The comprehensive school was created to provide equitable education opportunities for all children regardless of socio-economic background, and in both contexts, students' socioeconomic status has relatively little influence on their achievement (OECD, 2016a).

Another aspect of the school system that recently has become more similar is the organization of teacher education. In Finland, a master's degree has been required of all teachers since 1979; K–6 teachers have a master's degree in education and grade 7–9 teachers in a school subject (e.g., mathematics). In Norway, a master's degree and a certain number of mathematics courses became a requirement for all teachers in 2017,<sup>1</sup> in order to raise the “quality and status” of teachers (Government of Norway, 2017; Munthe & Rogne, 2016).

##### *1.4.2 The National Core Curriculum for Basic Education and The Knowledge Promotion*

Both Finland and Norway have long traditions with national curriculum, which is a core aspect in the abovementioned Nordic model (Blossing et al., 2014). The curricula most relevant for the studies of this thesis are the *National Core Curriculum for Basic Education 2004*, implemented in 2006<sup>2</sup> (Finnish National Board of Education, 2004), the *National Core Curriculum for Basic Education 2014*, implemented in 2016 (Finnish National Agency for Education, 2014), and the *Knowledge Promotion* (Norwegian Ministry of Education and Research, 2013a), implemented in 2006.<sup>3</sup> On a general level, the Finnish curriculum from

---

<sup>1</sup> The Oslo teachers in this study had an older teacher education with fewer required courses (see Appendix 2).

<sup>2</sup> The Finnish classrooms in Articles I and III follow the 2004 curriculum, while the two classrooms in Article II follow the 2014 curriculum.

<sup>3</sup> A new Norwegian curriculum took effect in the fall of 2020.

2014 more specifically than its forerunner promotes twenty-first century skills such as collaboration, networking, and digital literacy (Finnish National Board of Education, 2016; Vahtivuori-Hänninen, Halinen, Niemi, Lavonen, & Lipponen, 2014). The most recent national initiative to improve mathematics teaching in Finland is the ongoing LUMA Suomi<sup>4</sup> project (LUMA Suomi, 2020), consisting of resources targeted to support teachers' implementation of the 2014 curriculum (see also Hemmi, Krzywacki, & Partanen, 2017).

*The Knowledge Promotion* is a competence-based curriculum focusing on learning outcomes and basic skills, while the *National Core Curriculum* of 2004 and 2014 focus on both learning aims and social aspects of learning (Carlgren, Klette, Mýrdal, Schnack, & Simola, 2006; Mølsted & Karseth, 2016; Vahtivuori-Hänninen et al., 2014). While mathematics curricula in Norway have been criticized for encouraging teachers to follow the latest fads in mathematics instruction (Grønmo, 2017), recent governmental initiatives, including those involving the curriculum, have highlighted teachers' use of a *variety* of learning activities, especially collaborative and problem-solving activities, mathematics for everyday life, and engaging all students in mathematical thinking and reasoning (Bergem, 2014; Nortvedt, 2018; Norwegian Ministry of Education and Research, 2013b, 2015).

### **1.5 Overview of the thesis**

This thesis consists of two parts; the extended abstract (Part I) and the articles (Part II). This introduction chapter (Chapter 1) has presented the rationale, the overarching research question and key concepts for this thesis, as well as the Finnish and Norwegian educational systems. Chapter 2 provides an overview of research of mathematics instruction in Finland and Norway as well as of cross-national studies and conceptualizations of instructional quality. Chapter 3 outlines the theoretical foundation of this thesis and Chapter 4 clarifies the thesis' methodology and research design. Chapter 5 is a summary of the three articles, and finally, Chapter 6 discusses the empirical, theoretical and methodological contributions of this thesis in light of the main research question and aim, as well as offers suggestions for future research and concluding remarks.

---

<sup>4</sup> LUMA is an abbreviation of luonnontieteet [science] and matematiikka [math]. <https://www.luma.fi/en/centre>

## 2 Review of literature

In this chapter, I situate this thesis in the field of cross-national observation research on instructional quality in mathematics education, by reviewing other studies related to the main topics of this thesis. The following sections are included: mathematics instruction in Finland (2.1), mathematics instruction in Norway (2.2), cross-national classroom observation studies of mathematics instruction (2.3), different conceptualizations of instructional quality (2.4), and finally a summary of the chapter (2.5).

### **2.1 Mathematics instruction in Finland**

While no recent large-scale observation study has been conducted in Finnish mathematics classrooms, small-scale studies (Andrews, 2013a, 2013b; Andrews, Ryve, Hemmi, & Sayers, 2014), general descriptions of mathematics instruction (e.g., Krzywacki, Pehkonen, & Laine, 2016), studies on teacher self-reports (Kupari, 2003), and interview studies (Kaasila & Pehkonen, 2009; Pehkonen, 2007) have shed light on mathematics instruction in this context. The classroom studies by Andrews et al. (2014) and Andrews (2013a, 2013b) described lessons with a clear content focus, especially in teachers' making connections and selecting appropriate mathematical tasks as well as tools and representations for explaining content. So-called "reform-based practices", such as teachers building on student thinking or facilitating mathematical communication, are, however, absent (e.g., Andrews, 2013a). The absence of such practices is also evident in the way Krzywacki et al. (2016) has characterized a typical Finnish mathematics lesson, starting with a short mental calculation activity, followed by checking homework and the teacher explaining difficult tasks, and then the teacher introducing new topics, thereafter a large portion of individual seatwork in textbooks, with the lesson ending with the teacher assigning new homework. Also, Kupari's study (2003) supported this characterization, as Finnish lower secondary mathematics teachers' have self-reported that students practicing computational skills individually with teacher guidance is the most common activity they use, while whole group sessions including students interacting with each other and group work are less common. Furthermore, homework was reported to be a regular feature of mathematics lessons, with 85% of teachers assigning homework more than three times a week (Kupari, 2003).

Interview studies can illuminate mathematics teachers' rationales for enacted practices. In Pehkonen's (2007) study on Finnish teachers' perception of change in mathematics instruction, teachers stressed the importance of good textbooks for providing structure as well as tasks with different challenge levels. These teachers underscored that

learning mathematics is not supposed to be entertaining, and they perceived change in instruction as risky as it could lead the focus away from content and learning mathematics. Pehkonen (2007) concluded that the teachers perceived stability of instruction as more desirable than change moving towards collaborative and inquiry-based practices. In Kaasila and Pehkonen's (2009) study of Finnish student teachers' perception of good mathematics instruction, the importance of clear objectives for their lessons, their instruction, and the assessment of their teaching was stressed. The student teachers preferred whole-class instruction and individual seatwork over problem-solving activities because it allowed time for teachers to assist students who need more guidance (Kaasila & Pehkonen, 2009).

Taken together, these studies portrayed mathematics instruction in Finland as teacher-centered and content-focused yet procedural and with little attention to communication or interaction, with lessons often structured as individual seatwork. The prominence of individual seatwork has been discussed by several scholars (Carlgren et al., 2006; Kaasila & Pehkonen, 2009; Krzywacki et al., 2016) as a sign of increased individualization of teaching supported by constructivist learning ideals, where knowledge building is seen as a personal process and all students have their own pace and individual needs. Simola (2005) argued that such "traditional" ways of teaching (teacher-led instruction in combination with individual seatwork) is still possible in Finland because teachers believe in their role as authorities, and pupils accept it. However, others have argued that such traditional roles cannot foster skills needed for the twenty-first century and that this is one reason for the emphasis on collaboration and communication in the new curriculum (Hemmi et al., 2017).

## **2.2 Mathematics instruction in Norway**

A number of classroom observation studies have focused on instructional patterns in Norwegian lower secondary mathematics classrooms, most notably an evaluation of a previous curriculum (Reform 97; Alseth, Breiteig, & Brekke, 2003), the PISA+ study (Klette et al., 2008), and the recent LISA study (see Section 4.1). Across the first two studies, there are many commonalities in how mathematics instruction is characterized. For example, the most frequent activity formats were individual seatwork and teacher-centered whole-class instruction, with little allocated time for group work (Alseth et al., 2003; Bergem, 2009, 2016; Klette, Bergem, & Roe, 2016). During teacher-led whole-class instruction, teachers rarely asked open-ended questions inviting students to justify their thinking, while seatwork was more often than not guided by work plans prescribing the work students should complete over a period of time (Alseth et al., 2003; Bergem, 2016). These studies further suggested that mathematics instruction had a procedural rather than conceptual focus as teachers rarely



connected mathematical themes or previous knowledge, and that instruction seldom was driven by communicated learning goals or included cognitively challenging tasks (Alseth et al., 2003; Bergem & Klette, 2010; Klette et al., 2008).

From the more recent LISA study, analyses showed more variation in activity formats, for example, increased use of peer work (Klette, 2020). In addition, Stovner and Klette (in review) focused on Norwegian mathematics teachers' feedback practices, finding that feedback was mostly procedural, vague, or perfunctory. They also found that, while many teachers also provided some feedback likely to help students understand mathematics conceptually, very few did so consistently. The most common feedback situations were those showing students the correct procedures during individual and group work, and very little conceptual feedback was directed to the whole class. Thus, this study indicated a procedural and task-at-hand focus in the mathematics classroom, in accordance with previous research (Alseth et al., 2003; Klette et al., 2008).

There seem to be scarce research on Norwegian mathematics teachers' perspectives of their own instruction in lower secondary school. An exemption is Fauskanger (2016), who studied Norwegian lower secondary mathematics teachers' perspective of "good instruction." From these teachers' perspective, the teacher's task is creating student engagement and positive attitudes towards mathematics, and the teacher's role is to be enthusiastic and positive. Fauskanger (2016) concluded that it seems that the characteristics of the teacher and students were considered more important for good instruction than knowledge of content or how to teach mathematics.

### **2.3 Cross-national classroom observation studies of mathematics instruction**

In the following, I summarize key cross-national observation studies of mathematics instruction, followed by a section about their methodological approaches and challenges when analyzing and coding instruction. To the best of my knowledge, the only cross-national observation study including both Finnish and Norwegian mathematics classrooms is the VIDEOMAT study, which focused on how algebra is introduced across different contexts (Kilhamn & Säljö, 2019). While the sample only included two classrooms in Norway (grades 7 and 8), and three Finnish–Swedish classrooms (grades 6 and 7), the authors concluded that what characterizes algebra introduction in the Norwegian classrooms is the use of teacher-designed examples (not from the textbook), while the Finnish–Swedish teachers highly rely on the textbook and the teacher's guide for presenting algebra.

The TIMSS 1995 Video Study (Stigler & Hiebert, 1999), including American, German, and Japanese classrooms, is one of the first and most famous cross-national

observation studies. By analyzing single lessons from around 100 randomly sampled 8<sup>th</sup> grade mathematics classrooms in each context, this study intended to identify national teaching patterns and presented the idea of “cultural scripts.” The authors suggested that the script in Japan (high-level thinking, problem-solving, direct teaching, and extended seatwork and group work) was one reason for the success in international achievement tests. The follow-up TIMSS 1999 Video Study (Hiebert et al., 2003) compared mathematics instruction in a variety of high-achieving contexts: Australia, Czech Republic, Hong Kong, the Netherlands, Switzerland, the United States, and Japan.<sup>5</sup> They concluded that good mathematics teaching might look very different across contexts as, for instance, real-life examples were almost absent in Japanese classrooms while common in the Netherlands (Hiebert et al., 2003). However, reflecting back on the TIMSS studies, Stigler and Miller (2018) noted three commonalities of effective mathematics instruction for student learning across contexts: productive struggle (students engage in hard intellectual work), explicit connections (students receive support in making explicit connections between problems and concepts), and deliberate practice (students engage in sustained practice over time through a variety of strategies and are supported by feedback), and they suggested that these three commonalities could work as the basis for a framework for further comparative analyses.

The Learner’s Perspective Study (LPS; Clarke, Emanuelsson, Jablonka, & Mok, 2006) focused on mathematics instruction in 8th grade classrooms from 16<sup>6</sup> different national contexts. In contrast to the TIMSS studies, the aim of the LPS study was not to identify national patterns of instruction, but to document teaching and learning practices in classrooms of competent teachers in different contexts across a lesson sequence of 10 lessons. One key analyses of the LPS study concentrated on the same contexts as the TIMSS 1995 study (Stigler & Hiebert, 1999), namely the USA, Germany, and Japan, in order to assess if they could identify similar cultural scripts when studying sequences of lessons instead of single lessons. The LPS study concluded that there was little evidence of consistent cultural scripts across national contexts and that, instead, lesson activities depend on *where* lessons are located within a lesson sequence (Clarke, Mesiti, Jablonka, & Shimizu, 2006).

There have also been some cross-national studies on mathematics instruction at the elementary school level, situated in the school effectiveness field, and I will highlight The

---

<sup>5</sup> While Japan did not participate in TIMSS 1999, the Japanese classroom data from the TIMSS 1995 Video Study was reanalyzed for the TIMSS 1999 Video Study.

<sup>6</sup> Australia, the Czech Republic, Germany, Hong Kong, Shanghai and mainland China, Israel, Japan, Korea, New Zealand, Norway, The Philippines, Singapore, South Africa, Sweden, and the USA.

International School Effectiveness Research Project (ISERP) study and The International Comparative Analysis of Learning and Teaching (ICALT) study, since their goal included comparing instructional quality across contexts. The ISERP study (Creemers, Reynolds, Stringfield, & Teddlie, 2002) focused on nine<sup>7</sup> different countries and included several “school factors” that together make up school effectiveness, including instructional quality. Based on the findings from this study, Creemers et al. (2002) suggested that some aspects of instructional quality can explain variations in student achievement in any context, namely, the quality of teachers’ classroom management, clarity of instruction, quality of questioning, high expectations, and clear lesson structure as well as classroom climate. The ICALT study (van de Grift, 2007) compared instructional quality in the contexts of England, Flanders (Belgium), Lower Saxony (Germany), and the Netherlands. This study concluded that most aspects of instructional quality are similar across these central European contexts, while English teachers stood out on the features “adaption of teaching,” “clear instruction,” and “teaching learning strategies.” The author thus suggested that this indicates a possible English teaching style (van de Grift, 2007), that is, a “cultural script” (Stigler & Hiebert, 1999).

Across the reviewed cross-national studies, excluding VIDEOMAT where instruction was not systematically analyzed across contexts, instruction was analyzed and coded in three different ways: inductive bottom-up/part-to-whole, deductive top-down/whole-to-part, or by iterative abductive approaches shifting between inductive and deductive approaches (Clarke, Emanuelsson, et al., 2006; Erickson, 2006; Praetorius et al., 2019). The TIMSS study included top-down as well as bottom-up codes, inductively developed from video recordings to understand cultural differences (Jacobs, Hollingsworth, & Givvin, 2007; Stigler, Gonzales, Kawanaka, Knoll, & Serrano, 1999). The LPS study applied all three versions in several different studies (Clarke, Emanuelsson et al., 2006). The ISERP study originally applied top-down approaches using instruments developed in the USA; however, these became problematic outside the Anglo-American context as countries define measures of effectiveness differently (Muijs et al., 2018). Therefore, scholars behind the ISERP study designed the International System for Teacher Observation and Feedback (ISTOF), a top-down generic observation manual developed by a research team representing 20 countries (Teddlie, Creemers, Kyriakides, Muijs, & Yu, 2006). Still, Muijs et al. (2018) stressed that the conceptualizations of instructional quality in the ISTOF may have a bias towards student-active ways of working and not sufficiently reflect teacher-centered and direct instruction,

---

<sup>7</sup> Australia, Canada, Hong Kong, Ireland, the Netherlands, Norway, Taiwan, the UK, and the USA.

especially important for learning basic skills. The ICALT study also abductively developed their own instrument (also called ICALT) by combining a review of indicators of instructional quality with the perspectives of researchers from the participating countries (van de Grift, 2007). While studies have tested for measurement invariance in ICALT (e.g., Maulana et al., 2020), to my knowledge, no cross-national studies using ICALT have problematized issues with context-sensitivity in conceptualizations of instructional quality.

#### **2.4 Different conceptualizations of instructional quality**

Cross-national researchers investigating cultural perspectives have emphasized that conceptualizations and definitions of instructional quality depend on context and culture and their embedded views of learning and teaching (Alexander, 2000; Cai, Perry, Wong, & Wang, 2009; Martinez, Taut, & Schaaf, 2016). Alexander (2000) exemplified this argument in his five-country comparative study of teaching and learning (not mathematics specific), stating that actively asking questions is an indication of quality instruction in American classrooms, while quiet and listening students indicate quality instruction in Indian classrooms. Similarly, Cai et al., (2009) showed how beliefs about effective mathematics teaching differ among teachers from the USA, Australia, Mainland China, and Hong Kong. In their study, they identified an East–West dichotomy in beliefs about mathematics reflected in conceptualizations of effective teaching; teachers from the “East” held a structural view of mathematics as an abstract body of connected knowledge, while teachers from the “West” emphasized a functional view of mathematics as a useful tool to solve real-life problems. Yet, they also noted that there were similarities; for example, teachers from all contexts emphasized understanding and accommodating students’ needs as a characteristic of effective teaching. Further, Martinez et al., (2016) argued that cross-national differences in views of instructional quality are manifested in observation manuals. In their study of 16 observation manuals from different parts of the world, they found that, while the manuals were claimed to be designed for the same purposes (e.g., to evaluate teaching), they differed substantially in how they operationalized quality instruction. For example, US frameworks focused on behavior and narrowly defined aspects of general or content-specific instruction, building largely on research within the process-product tradition (see Section 3.1), while Singapore frameworks were developed inductively by examining teachers identified as exceptional, including both what they *do* in the classroom as much as what teachers *are*, covering psychological characteristics such as emotional intelligence (Martinez et al., 2016).

Observation manuals from the same country may also differ substantially in how they conceptualize and operationalize quality instruction. Schoenfeld et al., (2018) analyzed

alignment in scoring across three different observation manuals, all developed for US classrooms. The authors showed how a lesson could score high or low on Framework for Teaching (FFT; Danielson, 2013), a generic observation manual used across subjects, while the same lesson might score the opposite when applying specific mathematics manuals such as Mathematics Quality Instruction (MQI; Hill et al., 2008) and Teaching for Robust Understanding (TRU; Schoenfeld & Teaching for Robust Understanding Project, 2016). Schoenfeld et al., (2018) stated that this is because FFT is a generic manual focusing on general aspects of instruction, such as smoothness of lessons, clarity of instructions, and classroom management, while MQI and TRU are subject-specific manuals focusing on mathematical richness and the opportunities students have to grapple with key mathematical ideas. However, Schoenfeld et al. also argued that mathematics-specific frameworks might produce different results, as MQI scores a lesson high if the teacher frequently addresses connections between mathematical representations, while the same lesson receives a low score with TRU, if rote presentation of content undermines the connections made. In addition, Berlin and Cohen (2020) found that mathematical discourse as measured by a math-specific observation system was more common in classrooms that score high on emotional support as measured by a generic system, indicating that subject specific and generic systems might measure different yet complementary aspects of quality mathematics instruction. Still, Praetorius and Charalambous (2018) systematically compared 12 common observation manuals including mathematics subject-specific, generic, and hybrid manuals, concluding that they largely cover similar aspects of instructional quality.

## **2.5 Summary of the research review**

Together, the reviewed studies shed light on patterns of instructional quality in mathematics classrooms in Finland and Norway as well as on key conceptual and methodological challenges of studying instructional quality in a cross-national perspective. Studies of Finnish mathematics instruction indicated content-focused practices with a large portion of individual seatwork and whole-class discussions dominated by teachers, where textbooks and homework are important components (e.g., Andrews, 2013a, 2013b; Krzywacki et al., 2016). Studies of teacher perspectives showed that teachers may view such instruction as benefitting students' learning as it allows for individual guidance (Kaasila & Pehkonen, 2009; Pehkonen, 2007). Studies from Norwegian mathematics classrooms described instruction as often occurring in teacher-led whole-class discussions and individual seatwork guided by work plans, while often being fragmented in terms of content explanations and learning objectives (e.g., Alseth et al., 2003; Bergem, 2009; Klette et al., 2016). Recent analyses (Klette, 2020), however, have

pointed to an increase in group work *and* mathematical discussions, while research on teacher perspectives suggested that teachers perceive active and engaged students as indicators of good mathematics teaching (Fauskanger, 2016). Finally, the small-scale VIDEOMAT study (Kilhamn & Säljö, 2019) revealed differences in algebra introduction across these contexts, as teachers' constructed examples in Norway while using textbook examples in Finland. Many of the previous studies on instruction in Norway and Finland are, however, from a decade ago, and as stated in Section 1.4, there have been recent initiatives to improve mathematics instruction in both contexts.

The review of cross-national classroom studies revealed that some have suggested that distinctive “cultural scripts” exist (e.g., Stigler & Hiebert, 1999; van de Grift, 2007), while others argued that instruction also varies considerably within contexts (Clarke, Mesiti, et al., 2006). Generally, the studies methodologically applied different approaches to study patterns of instruction across contexts. Some reported a mix of different coding procedures (e.g., LPS study), while others reported top-down coding procedures informed by abductive processes (e.g., ISERP, ICALT). In respect to conceptualization, cross-national classroom *observation* studies have contended that quality instruction across different cultural contexts shares similar traits (e.g., Creemers et al., 2002; Stigler & Miller, 2018), while researchers interested in cultural *perspectives* emphasized how culture influence what is considered instructional quality in different contexts (Alexander, 2000; Cai et al., 2009; Martinez et al., 2016). In addition, conceptualizations may differ both between generic and subject-specific observation manuals as well as among subject-specific manuals (Berlin & Cohen, 2020; Schoenfeld et al., 2018), while generic and subject-specific manuals overall seem to share more similarities than differences (Praetorius & Charalambous, 2018).

Overall, this review illustrated how studies value different things when looking for quality in instruction, and that using different perspectives may be the most useful approach for understanding this phenomenon through classroom observations. As reflected in the reviewed studies in this chapter, empirical studies from Finnish mathematics classrooms are scarce (Simola et al., 2017), as is cross-national research of instruction other than East–West comparisons (Phillips & Schweisfurth, 2014). Thus, there is a call for updated empirical and theoretical studies from contexts sharing some basic similarities, such as the Nordic countries, to further nuance and challenge conceptualizations of instructional quality (Paine et al., 2016; Praetorius & Charalambous, 2018), which in turn may increase the usefulness and relevance of cross-national classroom studies (Xu & Clarke, 2019).

### 3 Theoretical perspectives on instructional quality

The purpose of this chapter is to clarify the theoretical anchorage of this thesis. As I study instructional quality in mathematics classrooms and theories of instruction seldom originate from one single theory of learning, but combine different theories, I will take what Tellings (2001) has described as an eclectic theoretical approach. Tellings underscored how educational research benefits from being eclectic and pragmatic, and often needs to integrate different theories to better understand an educational phenomenon. Taking an eclectic approach to studying instructional quality, this thesis incorporates complementing theoretical views of teaching and learning as embedded in the PLATO manual (Articles I and III) and in analyses of teachers' instructional rationales for discourse moves (Article II). Within mathematics education research, there are many different theoretical perspectives of learning and teaching in continuous development and debate (Cobb, 2007; Lerman, 1996; Stinson & Bullock, 2012). The most notable perspectives that shape the field of instructional quality in general, as well as mathematics education, are embedded in the process-product tradition (1960s–1970s), the cognitive and the constructivist traditions (1980s), the sociocultural tradition (mid-1980s), and the sociopolitical tradition (2000s) (Stinson & Walshaw, 2017). The four traditions relevant for this thesis are the *process-product*, *cognitive*, *constructivist*, and *sociocultural* traditions and their theoretical perspectives of learning and teaching. In this chapter, I briefly summarize them and link the different perspectives to features of instructional quality in mathematics education, and show how different perspectives are integrated in relevant observation manuals. Then, I analyze in detail how these perspectives may be seen as embedded in PLATO, the main framework used this thesis. Finally, I consider current theoretical debates related to instructional quality and try to position my own studies and reflections within this vast theoretical landscape.

#### 3.1 Process-product tradition

The process-product tradition is often considered to have properly established itself in the 1960s, with theories focused on identifying effective teaching practices (*process*) and relating these to student outcomes (*product*) (Stinson & Bullock, 2012). The process-product tradition is grounded mainly in experimental psychology and behaviorist theories, building on a theoretical presumption that it is possible to discern a causal relationship between features of instruction and student performance (Cobb, 2007; Stinson & Walshaw, 2017). Research within the process-product tradition decreased the complexity of teaching into decomposed specific instructional features (Gage & Needels, 1989). During the early years of this

tradition, such features were often low-inference, easily measured and captured aspects of instruction, such as classroom management (Brophy & Good, 1984), engaged time on task (Berliner, 1987), students' success rate on tasks (Fisher et al., 1980), and clear content-related objectives (Brophy, 1986). As cognitive and constructively oriented perspectives developed in the 1980s, the process-product tradition and its perspectives and methods were considered insufficient for understanding teaching and learning processes (Cobb, 1994; Winne, 1987).

Generic observation manuals within the school effectiveness tradition, such as the previously mentioned ICALT (van de Grift, 2007) and ISTOF (Teddlie et al., 2006), especially rely on process-product research to operationalize instructional quality. However, all observation systems to some degree decompose instruction into observable behavior, and some features of instructional quality stemming from this tradition are embedded in most of the currently used observation manuals including PLATO, in analytical categories such as classroom management, clarity of goals, and time on task (Bell et al., 2019; Praetorius & Charalambous, 2018).

### **3.2 Cognitive perspectives**

Scholars within the cognitive tradition, mainly inspired by Piaget, shifted the focus from predicting to understanding mathematics learning and teaching by concentrating on individuals' mental construction and sense making (Kilpatrick, 1992; Stinson & Walshaw, 2017). Cognitive theories about *learning* seek to explain teachers and students' inferred interpretations and understandings in terms of how new and existing knowledge is internally organized in cognitive structures and processes (Cobb, 2007; Shepard, 2000). Cognitive theories about *teaching* often address teachers' beliefs, perspectives, experiences, knowledge, and motivation, and how such cognitive traits relate to instructional practice (e.g., Fives & Gill, 2015; Pajares, 1992; Philipp, 2007). In this thesis, the concept of instructional rationales (see Section 1.3.3 and Article II) would also go under the umbrella of cognitive perspectives of teaching, as I in Article II attempt to describe teachers' perspectives of factors they perceive as shaping their enacted classroom discourse practices.

Scholars have criticized cognitive perspectives as insufficient to inform classroom instruction due to their focus on individual's mathematical reasoning (Cobb, 2007), and for assuming universal mental structures and cognitive processes across historical, cultural, and social settings (Kieran, Forman, & Sfard, 2001). However, others have argued that cognitive perspectives provide significant implications for classroom instruction. Kirschner, Sweller, and Clark (2006) as well as others (e.g., Archer & Hughes, 2011; Hammond & Moore, 2018) situated themselves in a cognitive tradition and made the case for *guided* or *explicit*



*instruction*, which is the term used in the PLATO manual. Explicit instruction is defined as a series of supports or scaffolds teachers can use to guide students through the learning process, including clear connections between previous and new knowledge, clear explanations and learning objectives, and decomposing learning into small steps supported by specific examples and targeted feedback (Archer & Hughes, 2011; Kirschner et al., 2006). Within this perspective, learning is defined as a change in long-term memory, and the aim of instruction is thus to guide learners in cognitively manipulating information in ways that are consistent with a learning goal, and ultimately fit and store new information in already existing mental schemes in the long-term memory (Kirschner et al., 2006, p. 75).

Explicit instruction is operationalized in observation manuals that include elements of scaffolding content understanding, for example, strategy instruction, formative assessment and feedback, accuracy and clarity of explanations, and connecting new and old knowledge. Emphasis on scaffolding content is evident in both mathematics-specific manuals, such as the MQI (Hill et al., 2008) as well as in generic ones, such as ICALT (van de Grift, 2007).

### **3.3 Constructivist and socio-constructivist perspectives**

Like cognitive perspectives, constructivist views of learning and teaching are mainly inspired by Piaget, putting the individual's meaning making at the center (Lerman, 1996). However, constructivism challenges cognitive and behavioristic assumptions that knowledge can be passively transmitted from teacher to learner, insisting that learners actively need to discover, construct, and organize knowledge themselves (Glaserfeld, 1995; Steffe & Gale, 1995). In this view, the teacher's role is to make sense of students' mathematical understandings, and when necessary, reconstruct learners' conceptions (Cobb, 2011; Glaserfeld, 1995). However, it is not the teacher's intervention that influences children's learning—it is how children experience and integrate the interventions into their own conceptual structures (Cobb & Steffe, 1983). Within a constructivist perspective, teachers should engage students in real world inspired problem-solving tasks, as this is assumed to facilitate students' constructive activation and integration (Glaserfeld, 1995). Similar to cognitivist scholars, constructivists have also emphasized that teachers need to facilitate student learning by relating new and old knowledge, yet such connections are derived more implicitly through situations and challenges that encourage students to make the connections (Barnes, 2008), unlike guided/explicit instruction focusing on structure and clarity in teacher explanations (e.g., Kirschner et al., 2006). Consequently, Kirschner et al. (2006) criticized constructivist minimally guided instruction, including discovery, experimental, and inquiry-based learning,

for not making a distinction between experts who can create their own knowledge and novices who need guidance as they do not have sufficient prior knowledge to internally guide themselves, which is especially a concern for struggling students.

The constructivist view is critiqued for portraying mathematics learning as an individual project, ignoring social and cultural contexts (Lerman, 1996; Sfard, 2006). Mathematics education scholars, most significantly Paul Cobb, responded to such critique by merging Piaget's psychological constructivist and Vygotsky's situated sociocultural approach to more fully understand teaching and learning in the mathematics classroom (Cobb, 1994; Cobb & Bowers, 1999; Yackel, Gravemeijer, & Sfard, 2011). This *socio-constructivist* approach grants that any aspect of a classroom can be viewed from either a social or an individual perspective, while stressing that individuals' construction is formed by their participation with other members in the classroom (Barnes, 2008; Cobb, Stephan, McClain, & Gravemeijer, 2001; Cobb & Yackel, 1996). From the socio-constructivist viewpoint, the teacher's role is to facilitate students' active construction of knowledge in interaction with others (see also Section 3.4.1).

An example of an observation manual that positions itself within the constructivist view of teaching, emphasizing students opportunities to construct their own knowledge and question their own understandings, is the generic FfT manual (Danielson, 2013). However, also subject-specific manuals such as PLATO (Grossman, 2015) and MQI (Hill et al., 2008) have a socio-constructivist emphasis evident in analytical categories of how students engage in content-related discourse.

### **3.4 Sociocultural perspectives**

Sociocultural perspectives, building on Vygotskian thinking (e.g., Vygotsky, 1978), underscore how teaching and learning are situated and historically, socially, and culturally evolving practices. The sociocultural approach thus focuses on how the wider school institution and/or society influences the classroom, in contrast to socio-constructivist views stressing that individual thought and sociocultural processes are reflexively related in the classroom (e.g., Cobb & Yackel, 1996). Some strands of sociocultural research have emphasized the role of language and cultural artifacts as mediating learning (Mercer & Hodgkinson, 2008; Säljö, 2000; Solomon & Black, 2008; van Oers, 2001), others have stressed that learning is situated (Jaworski, 2006; Lave & Wenger, 1991), or that learning can be facilitated with concepts such as the zone of proximal development (ZPD; Vygotsky, 1978). In essence, the ZPD is a space created by participants in interaction in which development occurs, and in order to create such spaces for learning, the teacher needs to

scaffold instruction based on the learner's needs (Paris & Winograd, 1990). The ZPD is thus based on the idea that students can do more with the help of others than on their own, and teachers should focus on the potential of learning with the help of scaffolding techniques such as interactive dialogue, modeling, and strategy training (Paris & Winograd, 1990). While the ZPD is often referred to as a sociocultural theory because of its Vygotskian roots, it aligns with socio-constructivist views emphasizing that learning occurs in the interactions among participants in a classroom context, as well as with cognitive perspectives supporting explicit instruction through scaffolding techniques.

Since learning and teaching are considered situated and contextually dependent in the sociocultural perspective, few standardized observation manuals exist that would purely adhere to this tradition. The ones that exist often have a focus on language and interaction analysis (see Jordan & Henderson, 1995), where peer learning and discourse are central (see next section).

#### *3.4.1 Learning through discourse*

Classroom discourse is highlighted both in socio-constructivist and sociocultural theories of learning, and it is central in Article II where I analyze teachers' rationales for their classroom discourse patterns. In a sociocultural perspective, interaction patterns are influenced by historically developed communication in a particular community (van Oers, 2001). Classroom discourse thus has a special role in facilitating understanding instruction in different contexts, and as David Clarke (2013a) has shown, different patterns of classroom discourse in American and Asian classrooms may reflect that different learning theories and pedagogical practices are valued and appreciated. In respect to the discussion of different conceptualizations of instructional quality (see Section 2.4), the emphasis on classroom discourse is especially evident in Anglo-Saxon mathematics research drawing on sociocultural theories (Hiebert & Grouws, 2007; Walshaw & Anthony, 2008), while researchers studying other contexts have questioned whether the influence of discourse on student learning is overestimated (Andrews et al., 2014; Clarke, 2013b; Kim, 2002). While learning through participating in discourse is seen as important in these perspectives, socio-constructivist and sociocultural scholars have also contended that learning through discourse is not a given, stressing that students need guidance in how to participate in order for communication to be productive (Kovalainen & Kumpulainen, 2005; Lerman, 2001; Sfard & Kieran, 2001).

Not surprisingly, it is especially within sociocultural perspectives that teachers discursive instructional moves in the classrooms have been carefully described and analyzed

(e.g., Alexander, 2008; Cazden, 2001; Lampert & Blunk, 1998; Mortimer & Scott, 2003; Nystrand, Gamoran, Kachur, Prendergast, & Gundlach, 1997; Wells, 1999). The instructional feature referred to as “dialogic teaching” (Alexander, 2008) has emerged from such work, characterized by classroom discourse patterns in which the teacher and students together construct the discourse, which is assumed to facilitate the construction of knowledge by inviting multiple perspectives, in contrast to “authoritative teaching” where the teacher controls the narrative (Alexander, 2008). In short, dialogic teaching in mathematics classrooms allows students opportunities to present problems, make conjectures, explain solutions, and provide justifications, and involves teacher moves such as revoicing, uptake of student ideas, and asking for clarification and justification (Anthony & Walshaw, 2009; Franke, Kazemi, & Battey, 2007). Furtak and Shavelson (2009), building on Mortimer and Scott (2003) as well as others, distinguished between dialogic and authoritative teacher moves in their framework to capture teachers’ discourse moves in science classrooms, which I applied to mathematics classrooms in Article II (for more information of this framework, see Appendix 3 and Section 4.4.2).

### **3.5 Integrating perspectives when studying instructional quality**

Tellings (2001, 2012) argued that educational theories may be integrated into theoretical frameworks in different ways. In this thesis, theories of instructional quality are integrated in a manner that Tellings (2012) has described as *horizontal addition*; that is, theories are complementing each other and not merged into a novel theory. This type of integration is useful when theories cover different aspects and together form a more complete picture of the studied phenomena, and conceptualizations and operationalization of instructional quality across observation systems thus often draw on several of the four presented theoretical perspectives of learning and teaching (process-product, constructivist, cognitive, and sociocultural) (see Bell et al., 2019).

According to Tellings (2001), integrating theories is possible only if theories and frameworks share the same foundations, and some scholars, for example Lerman (2019), view constructivist and sociocultural theories as contradictory and incoherent. A foundation for integrating theories on instructional quality could be argued for by considering the level of agreement in the field around key dimensions that constitute instructional quality. At the level of conceptualizing instructional quality, scholars have agreed on four dimensions, representing different theoretical stances, which include *instructional clarity* (clear goals, explicit instruction), *cognitive activation* (cognitive challenge, quality of tasks), *discourse features* (teacher–student interaction, student participation in content talk), and *supportive*

*climate* (managing classrooms, respectful atmosphere) (Klette, 2015; Kunter, Baumert, & Köller, 2007; Nilsen & Gustafsson, 2016). However, there is often a discrepancy at the level of operationalization, as frameworks with similar theoretical foundations do not operationalize the dimensions the same way nor use the same terms for similar concepts (Charalambous & Praetorius, 2018; Klette, in press), and even the same frameworks may be interpreted differently across studies due to insufficiently defined operationalization (Schlesinger & Jentsch, 2016). This means that observation manuals that adhere to the same, often integrated, theoretical perspectives of instructional quality may empirically capture and investigate rather different instructional practices across studies. Because of this inconsistency in terminology across manuals and the vagueness of theoretical origins within manuals, it may be difficult to comprehend the theoretical foundation of an observation manual, as well as to interpret the results derived by observation manuals in relation to theories of learning. Theories of learning are however seldom relationally connected with instructional practice or operationalized at the level teaching (Oser & Baeriswyk 2001; Klette, in press), making them weakly connected with distinct instructional practices. While PLATO, like other observation systems, builds on empirical research of student learning, its conceptualization of teaching and instructional quality also builds on the theoretical traditions of learning (Bell et al., 2019; Praetorius & Charalambous, 2018). Thus, to understand the results that any observational manual produces, it is necessary to be explicit about its conceptual grounding. Therefore, in the following, I analyze how PLATO's conceptualization of instructional quality draws on the presented theoretical perspectives.

### **3.6 PLATO's conceptualization of instructional quality**

The main designer of PLATO, Pamela Grossman, together with colleagues, described the theoretical underpinnings of PLATO as emphasizing: "The importance of rigorous content and intellectually challenging tasks, the centrality of classroom discourse in developing sophisticated understanding of content and disciplinary skills, and the critical role of teachers in providing instructional scaffolding for students to help them succeed" (Grossman, Cohen, Ronfeldt, & Brown, 2014, p. 295). Figure 1 gives an overview of PLATO's four domains and respective elements (see Section 4.4.1 for the methodological aspects of PLATO).

<b>Instructional scaffolding</b>	<ul style="list-style-type: none"> <li>• Modeling (MOD)</li> <li>• Strategy use and instruction (SUI)</li> <li>• Feedback (FEED)</li> <li>• Accommodation for language learning (ALL)</li> </ul>
<b>Disciplinary demand</b>	<ul style="list-style-type: none"> <li>• Intellectual challenge (IC)</li> <li>• Classroom discourse (CD); Uptake of student responses and opportunity for student talk)</li> </ul>
<b>Representation and use of content</b>	<ul style="list-style-type: none"> <li>• Representation of content; Quality of instructional explanations (QIE) and Richness of instructional explanations (RIE)</li> <li>• Connections to prior knowledge (CPK)</li> <li>• Purpose (PURP)</li> </ul>
<b>Behavior management</b>	<ul style="list-style-type: none"> <li>• Behavior management (BM)</li> <li>• Time management (TM)</li> </ul>

**Figure 1.** PLATO's domains and elements<sup>8</sup> (Grossman, 2015).

I will now unpack the four domains—drawing on the previously presented process-product, cognitive, constructivist, and sociocultural perspectives—to illustrate how the different traditions and their perspectives of learning and teaching can be seen as integrated in PLATO.

**Instructional scaffolding.** This domain consists of the instructional elements *Modeling* (MOD), *Strategy Use and Instruction* (SUI), *Feedback* (FEED), and *Accommodation for Language Learning* (ALL). The elements are operationalized to evaluate the extent to which teachers provide specific instructional support to facilitate student learning of content. Such instructional practices reflect cognitive perspectives of guided instruction, where content is decomposed into small steps supported by examples and scaffolds (i.e., SUI, MOD, ALL) (Kirschner et al., 2006) as well as by consistent specific and clear feedback (FEED) (Hattie & Timperley, 2016; Archer & Hughes, 2011). However, it can also be argued that these elements reflect the notion of ZPD, where teachers as more capable others support students in learning through scaffolding (Paris & Winograd, 1990).

**Disciplinary demand.** This domain includes the elements *Intellectual Challenge* (IC) and *Classroom Discourse* (CD) (divided into *Uptake of Student Responses* and *Opportunity for Student Talk*), capturing the extent to which students are engaged in cognitively challenging tasks and activities. IC differentiates between tasks that require rote and recall and those that require high-level thinking, and CD differentiates between tight teacher directed talk and instruction that gives students the opportunity to engage in content-related discussions. As

---

<sup>8</sup> The original version also includes the element *text-based instruction* within the disciplinary demand domain (Grossman, 2015).

examined in Article III, the way Disciplinary demand elements are operationalized assumes that all students commonly experience discussions and tasks. Such conceptualizations reflect socio-constructivist and sociocultural views of learning (Barnes, 2008; Sfard, 2006), emphasizing discussions as a means for co-construction of knowledge. However, IC also aligns with cognitive perspectives, since rigorous and complex tasks are assumed to activate students' thinking in a way that develops deeper understanding than rote tasks (Boston, 2012).

***Representation and use of content.*** This domain, which I refer to in Article I as Presentation of Content, includes the elements *Representation of content* (divided into *Quality of Instructional Explanations* (QIE), *Richness of Instructional Explanations* (RIE)), *Connections to Prior Knowledge* (CPK), and *Purpose* (PURP). Together, these elements capture how teachers present and frame content for students. These elements align with the cognitive perspectives ideas that learning occurs when the teacher delivers clear and connected content (Cobb, 2007). CPK has support among different theoretical approaches; cognitive and constructivist perspectives stress the need for clear connections in order for students to construct and expand their individual knowledge schemas (e.g., Stein, 2007), while sociocultural perspectives encourage the connections of new and old knowledge to create a ZPD (Edwards & Mercer, 1987; Lerman, 2001). PURP, however, which concerns how clearly teachers present and frame learning objectives, is an element originating from the process-product tradition (e.g., Brophy, 1986).

***Behavioral management.*** This domain comprises the elements of *Behavior Management* (BM) and *Time Management* (TM). These elements relate to how efficiently a teacher manages the behavior and the time on task in the classroom. Thus, poor management means lost instructional time or a disorderly classroom, and good management is an orderly classroom where instructional time is maximized with little downtime (Berliner, 1987; Brophy & Good, 1984). These elements reflect features of instructional quality highlighted by the process-product tradition.

Although the developers of PLATO in the previously cited quote (Grossman et al., 2014, p. 295) have indicated that PLATO has an eclectic approach to theory, they do not clarify in what ways theoretical perspectives are integrated or which specific theories are relevant for which domains. The above analyses of PLATO elements suggested that theories of learning may be seen as horizontally integrated (Tellings, 2011), drawing on multiple traditions and theories across domains, but sometimes also within domains. Such integration

is natural when the purpose is to broadly capture different aspects of instructional quality, in other words, instructional practices assumed to facilitate student learning (see Section 1.3.1).

### **3.7 Current trends and debates in conceptualizing instructional quality**

The field of observation research on instructional quality have become more complex and nuanced since the early years dominated by the process-product tradition. Within this field, ongoing debates are taking place among scholars who adhere to generic and context-sensitive conceptualizations, as well as generic and subject-specific conceptualizations of instructional quality. Several scholars have argued for harmonizing the field across observation systems by developing a common and, to some extent, generic framework and vocabulary to gain a more comprehensive picture of teaching (Grossman & McDonald, 2008; Klette & Blikstad-Balas, 2018; Praetorius & Charalambous, 2018). Other researchers, particularly scholars with a strong sociocultural stance, such as Pacheco (2009), are skeptical of common measures and have claimed that instructional quality is a construct that arises out of a complex set of social and institutional interactions—and that this complexity and context-sensitivity needs to be considered in frameworks of instructional quality. The concerns of Pacheco (2009) are mirrored in the literature on cross-cultural perspectives of instruction reviewed in Section 2.4, and also partly in cross-cultural assessment literature, as assessment instruments, including observation systems, applied across different contexts risk biased interpretations unless biases of the conceptual and methodological aspects of the measures are managed (van de Vijver & Tanzer, 2004). In addition, several mathematics education scholars have contended that there are distinct differences between subject-specific and generic conceptualizations of instructional quality. For example, Stein (2007) argued that clear and detailed presentations of content, often occurring in generic descriptions of instruction (including PLATO), are in conflict with reform-oriented mathematics education where students' own exploration of content is emphasized. Also, Herbst and Chazan (2017) called for mathematics-specific observation systems that pay attention to the specific mathematical topic being taught while also considering how mathematics is interactively constructed through interaction among teachers and students. Although this would serve some purposes, like understanding teaching of a specific content in a specific classroom, this approach would make it difficult to capture and differentiate patterns of instructional quality across several classrooms, which is the purpose of using PLATO in this thesis. Moreover, as mentioned in Section 2.4, mathematics-specific manuals differ in their conceptualization of quality instruction (Schoenfeld et al., 2018) while across generic and subject-specific manuals many aspects overlap (Praetorius &



Charalombous, 2018). The point is, looking at the current state of classroom observation systems, it looks like neither generic nor subject-specific manuals have privileged access to capture instructional quality in mathematics instruction, while a combination of systems show promising results (e.g., Berlin & Cohen, 2020). Here, more research is needed to identify discrepancies within subject-specific manuals as compared with discrepancies across generic and subject-specific manuals. While I do not think researchers should individually create their own observation systems for each study and research problem, we need synthesizing and integrative efforts for systems to work better across different contexts and contents. That is why I, in Article III, discuss ways in which PLATO insufficiently capture some elements of instruction in Finnish and Norwegian mathematics classrooms and suggest solutions for how to complement and adapt manuals to better capture instruction in these contexts. In the discussion of this Extended Abstract, I attempt to advance this debate and discuss the need for transparency in theoretical perspectives and possibilities in complementing generic conceptualizations of instructional quality with context-sensitive perspectives targeted to specific contexts.

### **3.8 Summary of theoretical perspectives in the thesis**

In this chapter, I have argued that the theoretical umbrella of this thesis is necessarily eclectic, since a variety of perspectives is essential for classroom research that studies different aspects of instructional quality. In Article I, we descriptively explored patterns of instructional quality related to discourse and presentation of content cross-nationally with the PLATO framework, and we embraced its eclectic approach of instructional quality. Yet, in Article II, I wanted to understand two teachers' differently enacted discourse patterns in the Finnish context. Therefore, I applied the Furtak and Shavelson framework (see Section 4.4.2) targeted for analyses of teachers' enacted discourse moves, and I draw on a combination of socio-constructivist and sociocultural perspectives by highlighting the specific classrooms as well as the country-specific context of Finland as important for understanding these teachers' discourse practices. Finally, in Article III, we problematized how PLATO's conceptualization and operationalization of instructional quality work across different classroom and national contexts. Taken together, these articles and this thesis support the concept that instruction is always embedded in a social and cultural context—while simultaneously adhering to an eclectic approach where several complementing theoretical perspectives are applied to explore, understand, and problematize different features of instructional quality (Tellings, 2012).

## 4 Methodology and research design

The intention with this chapter is to provide complementing methodological information to what is accounted for in the articles. This thesis draws on classroom video data from Finnish and Norwegian contexts supplied with teacher interviews from the Finnish contexts collected through the LISA study design (Klette et al., 2017). The way I select and use data and analytical procedures could be labeled a qualitatively driven classroom video research design. All three articles, to some extent, use both deductive and inductive approaches to understanding data, while the analyses are qualitatively driven and sequential (Hesse-Biber, Rodriguez & Frost, 2015). I start by presenting the aforementioned LISA study and clarify how the methodological choices and the LISA research design relate to the data and analyses in my articles. Then, I present the research design in each of the articles, followed by sampling, data sources, and analysis performed. Next, I discuss research credibility and ethical considerations, and finally, I summarize methodological strengths and limitations of this thesis.

### **4.1 Situating this thesis within the Linking Instruction and Student Achievement study**

This thesis is part of the LISA study,<sup>9</sup> a large-scale video study of instructional quality across multiple classrooms, subjects, and contexts. The LISA data include 47 mathematics and 47 language arts classrooms in Norway, and 8 in both subjects in Finland (Swedish speaking classrooms). A central ambition of the LISA study was to use a variety of data sources to investigate key aspects of instructional quality, namely, standardized video observations, gains in student achievement scores on national tests in reading and numeracy, student surveys on how they perceive the quality of their teaching, and teacher background data. The data collection in Oslo classrooms took place during fall 2014 and spring 2015, while data collection in the Helsinki classrooms occurred in spring 2016. Furthermore, I collected additional data from some of the Helsinki classrooms in 2018 for the analyses of Article II on teacher perspectives (see Section 4.1.2). Hence, when I entered the LISA study during autumn 2016, the data had already been collected and partly coded with the PLATO manual (see Section 4.4.1), which is the LISA study's overall analytical framework for analyzing instructional quality based on video data (see more on LISA design in Klette et al., 2017).

Working within a larger project with readily available data entails *secondary data analysis* (Dalland, 2011; Smith, 2011). For Article I, this meant that I did not need to struggle to get access to the field, as I had access to a number and variety of lessons, which enabled

---

<sup>9</sup> For more information, see <https://www.uv.uio.no/ils/forskning/prosjekter/lisa/>.

sampling of matching classrooms across Finnish and Norwegian contexts (see Section 4.2). Yet, secondary data analysis also offers considerable challenges regarding understanding the context of the classroom, as such analysis does not allow for personal experiences or relations with the participants, and provides limited knowledge of the classroom atmosphere (Andersson & Sørvik, 2013; Dalland, 2011). In addition, Hammersley (1997) argued that researchers using secondary data for purposes other than the original might experience a gap between available data and research focus. This was not an issue for the research purpose in Article I—exploring patterns of instructional quality across a number of lessons in two contexts, nor for the purpose in Article III—investigating potential biases in aspects of standardized observations of instructional quality. However, the existing data were not sufficient for my ambition of understanding teachers’ instructional rationales for classroom discourse practices in Article II, for which I collected new data.

#### *4.1.1 Research design of Article I*

The aim of the first article was to compare patterns of instructional quality in the two contexts of Helsinki and Oslo classrooms. This was a descriptive, cross-national video observation study based on a total of 47 lessons from 8 classrooms in each context, 21 from Helsinki and 26 from Oslo (see Appendix 2). In this article, I quantified features of instruction by providing a numeric description of instructional quality derived by the PLATO manual (e.g., patterns of instructional quality), together with qualitative descriptions of what the numbers represented in terms of teachers’ enacted instructional practice. This was descriptive (Phillips, 2006), as it described the phenomenon (patterns of instructional quality) and relations between instructional features (e.g., activity format and classroom discourse). It was cross-national (Phillips & Schweisfurth, 2014), as it applied a comparative approach to explore patterns of instructional quality in two different national contexts.

#### *4.1.2 Research design of Article II*

The aim of the second article was to gain understanding of teachers’ perspectives, that is, their *instructional rationales* (see Section 1.3.3) of their enacted classroom discourse practices, recorded on video, in the Helsinki context. It was a qualitative case study of how two purposefully sampled Finnish teachers, Anna and Bea, perceived and enacted classroom discourse practices in their respective 9th grade classes. They were sampled from Article I due to their differently enacted discourse moves, Anna standing out by constantly providing her students with opportunities to talk, while Bea was more typical of the Helsinki context as she very rarely provided her students with such opportunities. This could be labeled a

qualitative case study, as it used multiple data sources, (i.e., video-recorded observations and semi-structured interviews), to study the phenomenon of instructional rationales in a real-life context (Yin, 2012), using an interpretive approach emphasizing individuals perceptions and values (Denscombe, 2014).

#### 4.1.3 Research design of Article III

The third article addressed issues in using a standardized observation system in studying patterns of instructional quality across different classrooms contexts, and thus draws on insights and experiences from the two first articles. The aim of Article III was to problematize how conceptualization, operationalization, and sequencing of the lessons in observation systems may embed possible biases that could affect interpretation of results when applied in different classroom contexts, focusing especially on how lesson structure and activity format influence and interact with coding procedures. This was a theoretical and methodological discussion focusing on possible biases in cross-national studies, following van de Vijver & Tanzer (2004), and validity issues, following M. Kane (2013a). The discussion was supported and illustrated by examples of scoring the PLATO elements Intellectual challenge, Uptake of student responses, and Feedback, derived from the same sample as in Article I ( $N = 47$ ).

Table 1 summarizes the research designs of the three articles.

**Table 1.** Overview of research designs

	<i>Main research question(s)</i>	<i>Data sources</i>	<i>Unit of analyses</i>	<i>Research focus</i>	<i>Research methods</i>
<i>Article I</i>	<i>What activity formats do teachers use to engage students? What is the quality of instructional explanations of content, connections to prior knowledge, and setting a purpose for learning? What characterizes discourse features in mathematics classrooms?</i>	Video observations	15-minute segments of lessons ( $N = 71$ from Helsinki and $N = 91$ from Oslo)	Patterns of instructional quality	Descriptive, qualitative, cross-national, comparative video-observations
<i>Article II</i>	<i>How do two Finnish mathematics teachers with diverging practices perceive and enact student participation in discourse?</i>	Non-participant video observations and field notes, teacher interviews	Classroom discourse episodes (whole-class discussions and group work)	Instructional rationale, patterns of classroom discourse	Qualitative case study
<i>Article III</i>	<i>In what way may standardized observation systems of instructional quality produce biases when applied across different contexts?</i>	Video observations, patterns of PLATO scores in the elements of Intellectual challenge, Uptake, Feedback	Patterns of PLATO scores	Conceptualization, operationalization, and sequencing of instructional quality in standardized observation systems	Theoretical and methodological discussion of PLATO scores and rubrics

## 4.2 Sampling

I will now provide insight into the sampling procedures described in all three articles to increase transparency of the research. When comparing aspects of instruction across classrooms and contexts, some common criteria for selecting the classrooms should be chosen to enable fair comparisons (Tytler, Chen, Hackling, & Ramseger, 2019). The 16 classrooms in Article I, also used in Article III, were purposefully sampled (Patton, 2015) to be similar to each other based on criteria of similar mathematical topics and location of school—both in terms of urban/suburban location and socioeconomic status (SES) of the location (see overview of sample in Appendix 2). It is important to note that the labels *Helsinki* and *Oslo* include nearby municipalities (i.e., Helsinki and Oslo *regions*). In both contexts, most children attend their local lower secondary school, and since the SES of specific students is confidential in these contexts, it was estimated based on the location of the schools. The students in the observed classrooms were 13- and 14-year-olds, attending 7th grade in Helsinki and 8th grade in Oslo, the first year of lower secondary school, respectively.<sup>10</sup> The Helsinki context consists of *Swedish-speaking Finnish* classrooms, due to similar student achievement results in Swedish-speaking schools in the Helsinki area compared with Finnish speaking schools (Brunell, 2007; Harju-Luukkanen, Nissinen, Stolt, & Vettenranta, 2014) and for practical reasons of language, so that the Norwegian LISA team could analyze the data. Swedish is a national language in Finland, and students can attend Swedish-speaking schools from preschool all the way to university level. Since little empirical research on classrooms in Finland generally exists (Simola et al., 2017), it is difficult to predict whether instruction differs across the two language groups; hence, this thesis can shed light on whether instruction in Swedish-speaking classrooms reflects the results of the limited previous studies of instruction from the larger Finnish context.

The characteristics of the eight Helsinki classrooms served as criteria for finding a matching sample representing a variety of Oslo classrooms within the greater LISA material. While the possibility to sample matching classrooms and lessons from a large pool of video data was a clear strength for the comparison in Article I, there were also some challenges related to the previously mentioned secondary data analyses described in Section 4.1. For example, the time of the year of filming differed, as Oslo classrooms were filmed at different times during the whole school year, while all Helsinki lessons were filmed within the same three weeks in early spring (February/March). Such differences in the time distribution during

---

<sup>10</sup> In Finland, children start school the year they turn seven, in Norway, the year they turn six.

data collection might have influenced the captured instruction (Praetorius et al., 2019). At the beginning of the school year, the teacher and students might be unfamiliar with each other and act differently than later in the school year, especially since this was their first year of lower secondary school—and there has been some evidence from American classrooms that elements of instructional quality fluctuate over the course of a school year (e.g., Meyer, Cash, & Mashburn, 2011). However, sampling is often a question of priorities, and as mentioned, the main priorities when selecting matching classrooms for Article I and III were mathematical content, SES, and location of schools.

The two Helsinki teachers featured in Article II, Anna and Bea, were purposefully selected as a subsample from Article I, as they had shown different discourse patterns, mirrored in their PLATO scores, for the classroom discourse elements. A limitation of the sample was that Anna taught an advanced group in geometry and Bea taught a basic group in exponents. However, in this article, the purpose was to contrast and characterize instructional rationales of teachers enacting different types of classroom discourse, rather than compare the specific teachers. Thus, in contrast to Article I, the criteria for the same mathematical topic and similarity across student groups became less relevant, as the differently enacted classroom discourse practices was the main sampling criteria and the teachers were chosen based on which teachers were likely to best portray different discourse patterns.

In Article III, “odd patterns” (see Section 4.4.3) were identified and chosen to illustrate possible methodological and theoretical biases embedded in the elements Intellectual challenge and Uptake of student responses. In addition, Feedback was chosen because it exemplified how PLATO’s sequencing of lessons into 15-minute segments could give rise to biased representations of teachers’ feedback practices, a methodological issue we had also briefly addressed in Article 1 related to activity format. These elements were thus not sampled through a systematic procedure; however, they were purposefully chosen to illustrate our experiences with and interest in particular issues in coding with PLATO.

### **4.3 Data sources**

This project consisted of several data sources, as was natural due to the substudies’ different designs. An overview of the data sources is presented in Table 2.

**Table 2.** *Overview of data sources*

Data sources	Observations		Semi-structured interviews
	Video	Non-participant	
Article I	X		
Article II	X	X	X
Article III	X		

I will now present each of the different data sources: video-observations, non-participant observations including field notes, and semi-structured interviews. It is important to note that Article III was not entirely empirical; rather, we used video observations and PLATO scores to derive examples for the theoretical and methodological discussion, which I address in Section 4.4.3.

#### *4.3.1 Video observations*

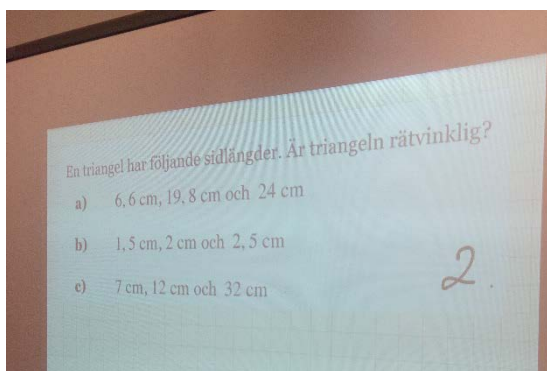
Using video observations for studying classroom instruction has many affordances and is considered vital to enable fine-grained analyses of the complex processes of teaching (Hiebert et al., 2003; Roth, 2013). Especially important for my thesis is the fact that video provides access to culturally distinctive practices which would otherwise be impossible to study and compare (Xu & Clarke, 2019). By viewing and reviewing and by discussing and systematically analyzing a number of videos of mathematics classrooms, I was able to gain an overview of differences and similarities in classroom practices in Helsinki and Oslo contexts. Moreover, video allows researchers to decompose and select specific events of interest (Derry et al., 2010), which was necessary when I conducted detailed analyses of whole-class and group work episodes as well as differences in teachers' use of classroom discourse in Article II.

The video data in this thesis capture “naturally occurring” instruction; that is, teachers were instructed to teach as normally as possible and were not given any specific instructions (Erickson, 2007). Naturally occurring instruction relates to what Clarke and Chan (2019) called “a window metaphor,” where the goal is to capture an undistorted view of the classroom where every attempt is made to minimize the impact of the research on classroom activities. However, even if the video captures instruction as it would naturally occur “through a window,” what researchers actually see is inevitably filtered first through the angle and position of the camera and the audio, and then through analytical frameworks (Clarke & Chan, 2019). Within the LISA design, the cameras had fixed positions in the classroom, one

facing the teacher and one facing the class, while the teacher wore a microphone, and an additional microphone was strategically placed to record student talk. Such a camera and audio setup generally functioned well to capture teacher talk and actions, while it was not always able to capture student talk during group work. The noise from the classroom microphone sometimes resulted in a cacophony of different voices without actually hearing what was said in each group. However, as my focus in this thesis is on the teachers' instruction, I relied mainly on the teacher's microphone. In addition, the findings may be seen as constructed through the analytical lenses applied, and I will address the importance of transparency in the way analytical frameworks portray instruction in classroom observations in the discussion.

#### 4.3.2 *Non-participant observation and field notes*

During fieldwork in Helsinki collecting data for Article II, I was present during the video-recorded lessons as a *non-participant observer*, also called an onlooker (Patton, 2015). This means that I sat at the back of the classroom without interfering with the lesson, taking notes, and managing the video and audio equipment. The notes mainly consisted of pictures and descriptions of student work, activities, and tasks in the lessons. Figure 2 shows one of the tasks assigned in one of Anna's lessons. This was a group task asking whether any of three (a, b, c) triangles were right-angled, and the students could figure this out based on knowing their length. Being present in the classroom, and documenting the different tasks and student work facilitated contextualization of the classroom discourse episodes in the analytical process (see Section 4.4.2). I also visited the two teachers' mathematics lessons with other grades, to gain a broader overview of the teachers' instruction. These additional visits allowed us to discuss



**Figure 2.** Example of task in Anna's classroom.

classroom discourse more broadly than just for a specific class. Overall, the experience of actually being present in the classroom before and after lessons, talking informally to the teachers, gave me a detailed understanding of the context of the classroom, not made available through secondary analyses of videos (see Hammersely, 1997).

#### 4.3.3 *Semi-structured interviews*

Semi-structured interviews with predetermined themes are recommended when the purpose of a study is to gain insight of participants' perspectives of a phenomenon from their everyday



life (Kvale & Brinkmann, 2009). In Article II, semi-structured interviews were crucial, since the phenomenon of instructional rationale directly involves teachers' perspectives of how they perceive their enacted classroom discourse practices (See Interview Guide in Appendix 4).

The interviews with Anna and Bea each lasted for one hour and were conducted in our shared mother tongue, Swedish. I transcribed them in Swedish and later translated sections of them into English. At the start of the interview, I presented the purpose of the interview and the themes I wanted to discuss. During the interview, both I and the teachers referred back to specific events that had occurred in the lessons I had observed. However, the teachers also talked about other events. Such narratives may inform researchers of how the participants understand the phenomena of interest (Kvale & Brinkmann, 2009). For example, Bea's prior experiences with problem-based methods in mathematics as being chaotic and Anna's experiences with teacher education's insufficiency in addressing variation in mathematics lessons provided insights into teachers' perceptions of classroom discourse activities.

#### **4.4 Analytical procedures**

In this section, I will present the two main analytical frameworks of this thesis as well as the analytical procedures described in all three articles.

##### *4.4.1 Protocol for Language Arts Teacher Observation (PLATO)*

PLATO is a standardized observation system (see Section 1.3.2) designed to capture instructional quality, developed by Pamela Grossman at Stanford University (Grossman, 2015). As the name suggests, PLATO is originally designed for English Language Arts, while it has also been used across subjects including mathematics (e.g., Cohen, 2013). Like other standardized observation systems, it offers a "top-down" approach (Praetorius et al., 2019) of predetermined categories for coding and scoring instruction. Several classroom researchers have argued for the benefits of standardized observations: they may minimize personal judgment and maximize valid and systematized analyses (Hammersley, 2010b), they make important features of teaching explicit (Klette, 2015), and they enable comparative analyses (Klette & Blikstad-Balas, 2018). Importantly, for the LISA study and Article I cited in this thesis with 47 lessons divided into 162 segments, PLATO allowed for processing patterns of instructional quality in a large sample of lessons rather quickly (Hardman & Hardman, 2017). PLATO consists of 12<sup>11</sup> elements, divided into four domains, as described in section 3.6. All elements are scored on a 1–4 scale, where 1 means no evidence of a certain practice, 2 means

---

<sup>11</sup> In PLATO 5.0, there is a 13th element, Text-based Instruction; however, this is not scored for mathematics instruction in the LISA study.

vague evidence, 3 means consistent evidence, and 4 means strong and consistent evidence. When coding with PLATO, each element is scored for each 15-min segment. Thus, a 45-min lesson will be sequenced into three segments, with three scoring points for each element.

**Analyzing instructional patterns with PLATO.** For the analyses in Article I, I focused on six elements of the PLATO observation system: the classroom discourse elements *Uptake of student responses* and *Opportunity for student talk*, and the presentation of content elements *Quality of instructional explanation*, *Richness of instructional explanations*, *Connections to prior knowledge*, and *Purpose* (see Appendix 1) to capture and compare instructional quality across Helsinki and Oslo classrooms. I focused on these aspects of instructional quality because, in initial PLATO coding of all 12 elements, they showed the biggest differences across elements and contexts, and because, based on previous research, they suggested possible differences across contexts. In the following, I will describe the analytical procedure in Article I (see Figure 3).



**Figure 3.** Analytical procedure in Article I.

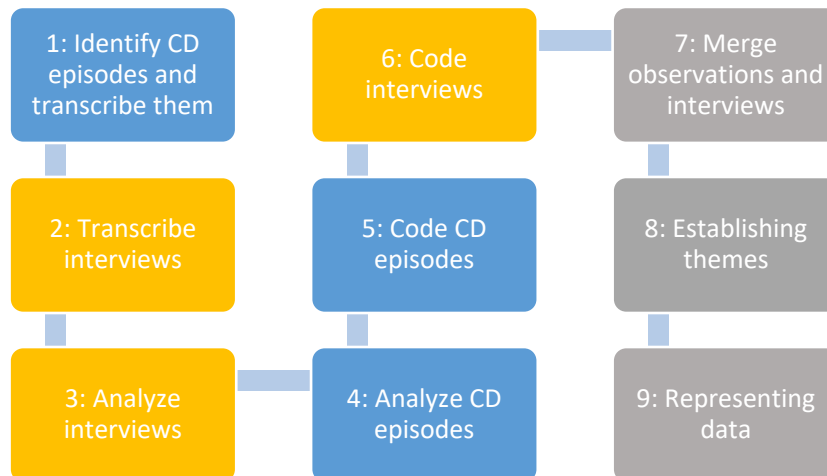
In Step 1, as mentioned, I sampled eight classrooms from the Helsinki area and eight from the Oslo area, based on criteria of mathematical content, location, and SES. The sixteen classrooms amounted to 47 lessons (21 in Helsinki and 26 in Oslo) and 162 15-minute segments. In Step 2, using the video-analysis software InterAct (Mangold, 2019), I PLATO coded and transcribed all Helsinki lessons, while most of the Oslo lessons had already been coded by other LISA coders. Therefore, I checked all previous coding and transcripts, and when necessary, recoded elements (see Section 4.5.1). This process included discussions with other LISA researchers to establish a common understanding of how we applied PLATO in mathematics. In Step 3, I compared the coding across the two contexts. Then, in Step 4, I qualitatively studied transcripts and examined what kind of instructional practices and circumstances typically produced the numeric patterns across and within the two contexts. During these analyses, I began to register limitations in secondary data analysis (see Hammersley, 1997) as well as in standardized observation measures. Thus, in order to more fully understand patterns of instructional quality, I wanted to know how teachers themselves

rationalize their enacted practices, as well as to problematize the limitations of PLATO when capturing instructional quality across different mathematics classrooms and national settings. These issues became the topics of the second and third articles, respectively.

#### *4.4.2 Teacher moves framework*

The *Teacher moves* framework by Furtak & Shavelson (2009) distinguishes between dialogic and authoritative teacher discourse moves (see Appendix 3). This framework builds on a large body of classroom interaction research, and specifically on Mortimer and Scott (2003). It allows for targeted analysis and obtaining an overview of moves that enable and hinder student participation in discourse. It was originally designed for science classrooms; however, it has been used across subjects (e.g., Andersson & Klette, 2016) due to its generic categories. The rationale for shifting frameworks was that in order to capture teachers' classroom discourse in detail, I needed a fine-grained framework compatible as a supplement to the PLATO framework (see Tellings, 2001), in the sense of classroom discourse theories (see Section 3.4.1). In addition, as I focused on the rationales of teachers with contrasting moves, Furtak and Shavelson's framework was well-suited as it allowed me to capture a large spectrum of different discursive teacher moves. As mentioned, PLATO sequences instruction into 15-minute segments, and short but intense discussions do not count as high quality discourse due the quantity–quality criteria, which I reflect on in Article I. The Furtak and Shavelson framework is not standardized requiring uniform sequencing of lessons the way PLATO does, and thus allowed for shifting the unit of analysis from 15-minute segments to specific lesson events (see Clarke et al., 2007), such as discourse episodes. In addition, I adapted the Furtak and Shavelson framework by adding two additional dialogic moves inductively derived from the data that specifically captured teachers' dialogic moves during group work, while the framework originally was developed for whole-class instruction. However, I did not adapt it to individual seatwork, as I operated within the socio-constructivist definitions of classroom discourse as activities that foster joint discussion of mathematics engaging several students (Mercer & Hodgkinson, 2008).

***Analyzing teachers' discourse moves and rationales.*** In the analysis of Article II, I worked with two main data sources (video and interview data) simultaneously to gain an understanding of teachers' rationales (see Figure 4). For the analysis of classroom observations in Article II, focusing on teachers discourse practices, I applied the above mentioned Furtak and Shavelson framework.



**Figure 4.** Analytical procedure in Article II.

In Step 1, the process of video analysis started with identifying and transcribing classroom discourse episodes (i.e., content-discussions during whole-class and group work). The interview analysis process in Step 2 started with transcribing a full word-for-word written version of the questions and answers, including nonverbal events, such as laughs, shrugs, and distracting events (e.g., phone ringing), as such events may influence the interpretation (Rubin & Rubin, 2012). In Steps 3 and 4, I read the interview and video transcripts multiple times to obtain an understanding of key issues and main ideas (Creswell, 2014), while simultaneously consulting the literature to gain an overview of how the themes fit within existing research on teachers' perspectives of classroom discourse. In Step 5, I coded the CD episodes with the analytical framework by Furtak and Shavelson (2009). In Step 6, I coded interviews based on previous research on factors shaping classroom discourse (e.g., belief and context factors), and I added and adjusted categories in response to the data (Cohen, Manion, & Morrison, 2018), thus performing abductive reasoning by combining theory driven categories and inductive reasoning (Patton, 2015). For example, while teachers' perceptions of how students learn have been well studied and were expected to emerge as a theme, the prominence of tensions between discourse and other concerns in the teachers' rationales was unforeseen. In Step 7, I merged observations and interviews by writing descriptions of the classroom discourse events based on the analysis of teacher moves and teachers' reflections about students participating in discourse. This process generated clearer themes of what constitutes teachers' instructional rationales as well as what the two teachers had and did not have in common that shaped their instructional rationale (Step 8). Finally, in Step 9, I chose examples that best represented the themes to be included in the manuscript.

#### *4.4.3 Analytical procedure in Article III*

As mentioned, the third article consists of a theoretical and methodological discussion based on empirical examples highlighting possible biases in how three PLATO elements (Intellectual challenge, Uptake of student responses, and Feedback) conceptualize, operationalize, and sequence instructional quality. When PLATO coding Helsinki and Oslo mathematics lessons ( $N = 47$ ) with all 12 PLATO elements in a pre-analysis for what would become Article I, my co-authors and I identified “odd patterns” in coding these elements. The odd pattern for Intellectual challenge was low inter-rater agreement of 30% in the Helsinki sample, and for Uptake, significant differences between the two contexts (66% on score 1 in Helsinki compared to 35% score 1 in Oslo on the PLATO scale). The Feedback element did not derive from odd patterns in coding, yet we included it as it illustrated issues with lesson sequencing and the quantity–quality criteria of rubrics, as little previous literature has problematized this issue with empirical examples. To analyze possible biases in these elements, we studied the rubrics, the scoring, the transcripts, and scoring justifications along with videos, and within the LISA team, we discussed how classroom factors produced these patterns. This was an iterative process and did not result in any statistical reports on potential biases, yet the illustrative examples enabled a theoretical and methodological discussion of issues of possible biases common in the observation system when applied in different contexts.

### **4.5 Research credibility**

The following sections about research credibility are discussed from reliability, validity, and generalizability perspectives, with the purpose of making the procedures transparent and my reflections about the process as clear as possible.

#### *4.5.1 Reliability of standardized observations*

Reliability is concerned with the degree of accuracy, consistency, and replicability in research (Cohen et al., 2018). In observation studies, standardized observations are considered to have an advantage over unstandardized approaches, as raters are trained to score in an explicit, reliable, and valid way which will lessen the effect of subjective views of quality teaching, and ensure alignment with given criteria of the observation system (Bell et al., 2014; Fischer, Praetorius, & Klieme, 2019; Klette & Blikstad-Balas, 2018). Observation manuals themselves are consequently not inherently reliable or unreliable. The reliability is determined by the ways in which scores are obtained, which is influenced by *training*, *raters*, and *length of observation periods* (Pianta & Hamre, 2016, p. 25).

*Training* of raters, also called observers or coders, is a common way to strengthen reliability, as it guides raters' attention to key features of instruction in line with the view of teaching of the observation manual (Archer et al., 2016; Bell et al., 2014). Together with other LISA-affiliated scholars, I participated in a weeklong intensive PLATO training conducted by the developers of PLATO. The training involved watching and discussing training videos, and scoring instruction according to the PLATO rubrics with the help of clear scoring directions. The training culminated with a certification test where I reached the 80% threshold of agreement with a master observer at Stanford University on all PLATO elements. However, a challenge for applying PLATO to *mathematics* lessons is that the training videos focused on English Language Arts. While the scoring procedure for most of the elements in the PLATO manual is written in a generic way applicable to most subjects, it could be argued that there can still be a greater risk of raters developing idiosyncratic understandings of the rubrics when scoring mathematics, especially in my case since there were no training videos from mathematics classrooms. Therefore, other LISA scholars and I started working on supporting guidelines to establish a "mathematical" understanding of PLATO by exemplifying how PLATO can be translated to the mathematics classroom.<sup>12</sup>

*Raters*, and especially *consistency across raters*, is the most significant reliability threat in observation scores of instructional behavior (Ho & Kane, 2013; Pianta & Hamre, 2016). Consistency is often assessed by evaluating the *inter-rater agreement*, which measures how frequently different coders assign the exact same scoring (Gwet, 2014). Inter-rater agreement of coded observations can be calculated in many ways, most commonly by kappa or percent of absolute agreement. While kappa is recommended to adjust for chance agreement of raters (Gwet, 2014), I chose percentage, which is especially recommended when the sample and number of raters are small (Graham, Milanowski, & Miller, 2012). In classroom observations, calculating the percentage of times the scorings fall within one performance level of one another may also be useful, that is, *the percentage of exact and adjacent agreement* (Graham et al., 2012). For the small Helsinki sample, I thus engaged a fellow LISA researcher in mathematics education to code 15% of the Helsinki segments for inter-rater purposes (see Table 3).

---

<sup>12</sup> The legacy of this work is now continued in the Quality in Nordic Teaching (QUINT) project. See <https://www.uv.uio.no/quint/english/projects/lisa-nordic/>.

**Table 3.** *Percentage of exact and adjacent agreement between coders*

	Purpose	Connection to prior knowledge	Quality of instructional explanations	Richness of instructional explanations	Uptake of student responses	Opportunity for student talk
Exact agreement	5/9 55%	3/9 33%	3/9 33%	5/9 55%	8/9 89%	9/9 100%
Exact and adjacent agreement	7/9 77%	8/9 89%	9/9 100%	9/9 100%	9/9 100%	9/9 100%

The exact agreement for some of the codes displayed in Table 3 may seem very low. Examination of rater disagreement, however, led us to conclude that most divergences resulted from the coding process and practical differences in how the coding was conducted, and not our interpretation of the codes. I had transcribed all videos and used the transcripts to evaluate the scores, while the other researcher coded without transcriptions pausing the video only to note short justifications of each code. When comparing coding, we noted that transcriptions allowed us to detect the framing of wording that is important for several elements in PLATO. For example, for the element CPK and PURP (see Appendix 1), many of the references to prior lessons and lesson goals went unnoticed by the researcher without transcripts. Thus, we concluded that my coding, which relied on the full classroom transcript in line with the scoring system for PLATO, was the more accurate reflection of the PLATO rubric. In this sense, multiple raters can be very challenging as they may have different perceptions of practices and carry biases based on attention to details (Casabianca et al., 2013; Liu, Bell, & Jones, 2019). On my part, an additional bias could be that I had watched all the lessons and was intimately familiar with the teachers' lessons, while the other researcher was randomly assigned segments. Studies have found that scores change depending on how much of a teachers' practice you observe (e.g., Ho & Kane, 2013), and different degrees of data familiarity are thus a potential issue for inter-rater reliability. Therefore, I conducted an additional reliability check by re-watching all the Oslo videos and checking each code (mostly scored by other PLATO raters), adjusting the score according to a detailed coding with PLATO, thus reducing potential rater bias. While this approach may undermine the idea of replicability, it is likely to increase the *internal consistency* of scoring across contexts, which is especially necessary in comparative analyses of classroom practices.

*Length of observations* concerns the observation period—how many segments or lessons are enough for reliable interpretations of teacher practices. While I did not make conclusions about specific teachers' practices looking across teachers' in Article I, I

aggregated teacher scores to display patterns across contexts. Whereas no guidance exists about the optimal number of observations (Martinez et al., 2016), studies often use only one or two lessons as indicators of teachers' instructional quality (Praetorius, Pauli, Reusser, Rakoczy, & Klieme, 2014). Other studies have suggested that four consecutive lessons is enough to provide a sufficient overview of key instructional practices in a classroom (Ball & Hill, 2009; Klette, 2009), and that consecutive lesson sequences are necessary in cross-national studies (Clarke, Mesiti, et al., 2006). Especially for PLATO, a generalizability study of sources of variation in scores by Cor (2011) indicated that five 15-min segments from each teacher are sufficient for obtaining reliable scores per teacher. All the PLATO classroom analyses in this thesis included more than five segments as well as consecutive lessons.

#### *4.5.2 Validity of standardized cross-national observations*

There are many validity issues in systematically capturing instructional quality, and to increase validity of classroom observations, we need critical discussions of our analytical frameworks, especially in cross-national research (Goldman, 2007; Park, 2019; Xu & Clarke, 2019). Thus, in Article III, we engaged in a theoretical and methodological discussion of selected PLATO elements and the validity of the scores they produce, using empirical examples from the Helsinki and Oslo video data. In this thesis, I consider validity using an “argument-based approach” following M. Kane (2013a). This approach suggests that validity is not a property of the research instrument, but rather a property of the proposed claims, that is, the interpretations and the uses of the interpretations. For observation systems, this means that “interpretations and uses of observation scores must make sense and be supported by appropriate evidence to be valid” (Hill, Charalambous, Blazar, et al., 2012, p. 90). The main validity argument in this thesis is that the chosen standardized observation manual, PLATO, could validly capture different aspects of instructional quality in Helsinki and Oslo mathematics classrooms. I recognize that all classroom studies have certain perspectives that risk portraying a reductionist view of instructional quality (Blikstad-Balas, 2017), and that this may be seen as a threat to *construct validity*, since potentially important features of instructional quality may be uncaptured. Yet, drawing on the “argument-based approach” (M. Kane, 2013a), I would argue that this is not necessarily a threat to validity if the limited focus is recognized and the implications are discussed and problematized, so that the interpretations of the results do not go beyond what can be claimed. In this thesis, I am interested in descriptive patterns of instructional quality as it is defined in the PLATO manual, while I refrain from making evaluative claims about, for example, what instructional practice is most beneficial for student learning. If, however, the purpose of using observation manuals is to



make ambitious claims about how specific features of instructional quality relate to student learning, different kinds of validity claims, such as evidence supporting predictive validity, would be required (see Bell et al., 2012; M. Kane, 2013b). The research design in my thesis does not allow for making any claims about predicting students' learning. Instead, I recognize that instructional quality as defined in PLATO considers a particular conceptualization of instructional quality reflecting different perspectives in education literature on what kinds of instruction may be beneficial for learning (see Chapter 3).

*Equivalence* across contexts is an undisputable challenge to validity within any comparative cross-national study (Schweisfurth, 2019). In using a comparative approach in Article I, I created an expectation that what was being compared would be culturally, contextually, structurally and functionally equivalent (Nowak, 1977; Phillips & Schweisfurth, 2014; Schweisfurth, 2019). *Cultural equivalence* concerns whether the same phenomena are valued the same way, in other words, whether instructional quality is viewed the same way across contexts. Using a standardized observation system, Article I operated with a predetermined definition and did not directly consider how instructional quality is viewed across the two contexts of Finland and Norway. However, we problematized standardized definitions by being transparent about the limitations of such definitions, especially when used across national contexts, partly in Article I and mainly in Article III. *Contextual equivalence* concerns whether the objects of study (here, 7th and 8th grade mathematics classrooms) have similar properties. For Article I, this was established through careful sampling (see Section 4.2). *Structural equivalence* concerns whether the schooling in general is structured the same way, whereas *functional equivalence* concerns whether the observed instruction serves the same function across contexts. As presented earlier (see Section 1.4), the Finnish and Norwegian educational systems share many structural and functional similarities regarding the structure, purpose, and values of schooling, as for example, mathematics instruction in both contexts is framed by a national curriculum.

#### 4.5.3 Reliability and validity of interviews and non-participant observations

The identity of the researcher may affect the reliability in interviews, since the researcher may influence the way the interviewee answers and the way the interviewer interprets the answers (Kvale & Brinkmann, 2009; Rubin & Rubin, 2012), and also the way the researcher observes a situation (Hammersley, 2010a). My identity as a Finnish–Swedish PhD student educated within the Swedish-speaking education system in Finland up to teacher education level may have given me some advantages in gaining access to teachers' perspectives and practices. For example, Rubin and Rubin (2012) asserted that interviewees may feel more comfortable

speaking to someone that they consider local, as they are familiar with the social rules. In addition, Hammersley (2010a) argued that, when we observe a scene familiar to us, we have competence in understanding the intentions and reasons for the actions we observe, and the scene of Finnish–Swedish mathematics classrooms was very familiar to me due to my educational background.

Thorough preparation of the interview guide and piloting interviews enhances the reliability of interview data (Cohen et al., 2018; Silverman, 1993). Thus, I prepared for the interviews by constructing and discussing the interview guide with my supervisors and colleagues as well as piloting it with two mathematics teachers to ensure the questions were clear and to become aware of potential issues. In addition, *triangulation* of different data sources can enhance reliability in qualitative research (Creswell, 2014). In Article II, I therefore triangulated findings from interview data and classroom observations in the process of the analysis to justify the themes related to factors that influence teachers’ instructional rationales. While the interview data served as the primary source of teachers’ perspectives, the observations strengthened the interpretations of rationales as in the interviews and in the analysis, I could refer to what teachers actually did in terms of classroom discourse practices.

*Member checking* means that researchers’ bring back their analyses to the participants to validate interpretations (Creswell, 2014). During my fieldwork for Article II, I revisited two of the eight Helsinki teachers, and presented major findings from the comparison in Article I as well as interpretation of their specific instruction in Article II. They concurred with my interpretation of common patterns in Finnish mathematics classrooms in general, as well as in my interpretation of instruction in their specific classrooms. When I asked Bea to reflect on the scarce opportunities for students to participate in classroom discourse in the overall Helsinki sample, she stated: “It doesn’t surprise me, and I hope that we are slowly making progress on this matter. I think it depends on many factors, I think it also lies in our nature to be too bound to books and literature.” Anna drew on her experiences from attending seminars in Sweden when reflecting on the same issue:

We [Finnish] mathematics teachers are quite conservative. In other subjects, they have come much further in using different working methods. . . . I have attended seminars in Sweden where I experienced that they are more focused on working methods than we are in Finland. I have not seen these elements in Finland, and I feel lonely. I think we are in the starting phase. I don’t know why, maybe the curriculum or the teacher education. In Sweden, you become a mathematics teacher. In Finland, you become a mathematician, and you can be the dumbest person ever and not function with people at all. Or, that is how it’s traditionally been.

Thus, the reflections by the teachers functioned as a kind of member checking, and indicated that the findings in Article I and II portray a valid picture of what these participants see as typical mathematics lessons in the Helsinki sample as well as in their classrooms.

*Reactivity* is a reoccurring validity threat in all observation studies, and it concerns to what extent the presence of the camera or/and researcher affects the behavior of the participants. Blikstad-Balas (2017) claimed that reactivity is no bigger validity threat in video studies than in other observation studies, as participants will not likely suddenly act very differently if a camera or just a live observer is present in the classroom. Teachers sometimes told me that, while they were filming for the LISA project, they postponed a session they thought would be uninteresting to film. Teachers involved in my data may have made similar adjustments that I was not aware of. However, this would not be a reactivity threat in terms of capturing instructional quality, since while teachers may change activities, they most likely cannot change their instructional repertoire (Blikstad-Balas, 2017). Thus, while there are issues with reactivity in any classroom research, the observed instruction is not likely to differ from everyday classroom instruction.

#### *4.5.4 Representativeness and generalization*

Generalizability concerns to what extent we can generalize research findings constructed with a certain method and framework in a specific setting (Morgan, 2007). The comparative samples in Articles I, also used in Article III, have specific traits that limit the generalizability to a greater population. The Helsinki sample may be seen as representing Swedish-speaking mathematics classrooms *in this particular area* within and around Helsinki, since there in general are few such schools. They are, however, not representative of *Finnish-speaking* schools in that area, as such schools, for example, tend to be more ethnically diverse. Also, they do not represent Swedish-speaking Finnish schools in general, as the Helsinki region differs from other Swedish speaking parts of the country (Brunell, 2007; Harju-Luukkanen et al., 2014), and Helsinki schools in general differ from other regions in Finland in terms of higher educational background of parents and higher student achievement (Vettenranta et al., 2016). In addition, similar to the case in Finland, Oslo schools tend to have higher student achievement compared with the rest of the country (Norwegian Directorate for Education and Training, 2020). The small sample of two teachers used in Article II is not representable beyond the particular cases, but the detailed descriptions of their instructional patterns and rationales may be recognizable to other teachers, in which case the findings may be generalized to a broader theory of teachers' instructional rationales (see Yin, 2012).

Another aspect of generalizability which is important for classroom studies is *representability of the examples* chosen to illustrate instructional practices of the larger sample (Miller & Zhou, 2007). Video analyses of classroom interactions are usually brief time scales of just a few minutes or a single lesson, while longer time scales of a week or a school year are rare (Lemke, 2007). Short time scales risk *magnification* of details out of proportion which might misrepresent classroom activities as these may be insignificant if a longer time scale had been used (Blikstad-Balas, 2017; Lemke, 2007). This is especially a risk in cross-national research, as researchers without cultural and contextual understanding might emphasize interesting events that do not reflect the larger context (Phillips & Schweisfurth, 2014). In my articles, I used two ways to establish the representativeness of examples. First, I relied on coding the relevant features of the examples (Miller & Zhou, 2007) as I, across the articles, qualitatively analyzed videos and transcripts to determine typical instructional practices behind the quantified scorings, and then selected illustrative examples representing such practices. Second, representativeness may be confirmed by experts on the phenomenon described (Miller & Zhou, 2007), and often these are cultural natives (Xu & Clarke, 2019). Due to my previously mentioned background, I chose the examples from Helsinki classrooms, while a Norwegian mathematics teacher involved in the LISA study was consulted for choosing examples from the Oslo classrooms.

Taken together, while not representative of the larger national contexts, the variation of schools in terms of location and SES, the multiple lessons, and carefully chosen examples increased the likelihood that the described mathematics instruction would be recognizable in schools beyond these specific samples within the context of Finnish–Swedish schools in the Helsinki area and schools in the Oslo area.

#### **4.6 Ethical considerations**

As part of the LISA project, this study was approved by the Norwegian Centre for Research Data (NSD).<sup>13</sup> The two countries have different procedures to obtain permission to film in the classroom. In Norway, researchers can approach individual schools after a study is approved by the NSD, which the LISA study did in 2014–2015. In Finland, classroom research has to be approved at the municipality level first and then by the schools, and this study obtained approval for both data collection periods (2016 and 2018).

All teachers, students, and their parents provided written and informed consent to participate. In line with the guidelines from the National Committee for Research Ethics in the

---

<sup>13</sup> See <https://nsd.no/nsd/english/index.html>.

Social Sciences and the Humanities (NESH<sup>14</sup>; 2016), the informed consent included information about who would be in charge of the project, sponsors of the project, and the purpose of the project (see Appendices 5–7). The information was provided in Norwegian and Swedish in writing, but teachers, students, and their parents could also contact the researchers and ask questions before, during, and after data collection.

The nature of digital video data presents ethical issues about the identification of participants (Derry, 2007). To ensure confidentiality of participants, the LISA project uploaded all data to a secure server at the University of Oslo, which was only available for researchers involved in the project. I will now consider the ethical issue of evaluative frameworks, which is especially relevant in cross-national classroom research.

#### *4.6.1 Evaluative frameworks in cross-national research*

Cross-national comparative research presents many ethical challenges for researchers (see Park, 2019). Some of the most concerning challenges relate to theoretical frameworks and interpretations based on such frameworks, and the evaluative nature of cross-national analyses (Clarke, 2013b; Park, 2019; Xu & Clarke, 2019). That is why many have argued that cross-national research in education should aim at making contributions without being evaluative (e.g., Phillips & Schweisfurth, 2014; Xu & Clarke, 2019), and be humble about the limitations of representing multiple contexts (Park, 2019). Therefore, throughout this thesis, I have challenged context-sensitivity in the conceptualization of instructional quality by emphasizing that PLATO's view of instructional quality has its limitations and adheres to certain theoretical traditions (see Section 3.6), which is the case for all observation systems.

#### **4.7 Methodological limitations and strengths**

I will now list and summarize the main methodological limitations and strengths that I have addressed in this chapter, related to sampling, data collection, analytical procedure, and research credibility.

---

<sup>14</sup> NESH is an abbreviation for Den nasjonale forskningsetiske komité for samfunnsvitenskap, humaniora

**Table 4.** *Overview of methodological strengths and limitations in each article*

	Possible strengths	Possible limitations	Comment on limitations
Sampling and data collection	<p><b>Art 1:</b> Multiple classrooms and lessons; naturally occurring instruction; purposefully sampled classrooms matching in SES, location, and mathematical content</p> <p><b>Art 2:</b> Purposefully sampled teachers; flexible interview guide</p> <p><b>Art 3:</b> Illustrative examples from multiple classrooms; represents common challenges across observation manuals</p>	<p><b>Art 1:</b> Different time of year of filming due to secondary data collection; limited generalizability to a larger context</p> <p><b>Art 2:</b> Low generalizability; different student groups and mathematical contents</p> <p><b>Art 3:</b> Limited representability of examples, as they are based on coding experience for a specific study</p>	<p><b>Art 1:</b> Finnish–Swedish classrooms in Helsinki and only capitol areas limit representability, while multiple and matching classrooms portray possible distinct patterns in these specific contexts</p> <p><b>Art 2:</b> Focus on two illustrative cases, not comparing teachers or generalizing; while contextually specific, it may accumulate theory on instructional rationales</p> <p><b>Art 3:</b> Focus is on problematizing possible biases in selected elements, not a total description of all potential biases</p>
Analytical procedure and research credibility	<p><b>Art 1:</b> Standardized observation manual with qualitative explanations making instructional practices explicit; internal consistency in coding</p> <p><b>Art 2:</b> Member checking; triangulation of data; contextual familiarity; visiting multiple classes with the same teacher</p> <p><b>Art 3:</b> Examining purposefully chosen examples to discuss possible biases with both quantitative and qualitative approaches</p>	<p><b>Art 1:</b> PLATO training not focusing on mathematics; partly low inter-rater agreement</p> <p><b>Art 2:</b> Reactivity; not a standardized observation instrument</p> <p><b>Art 3:</b> No statistical analysis of biases</p>	<p>Art 1: Focus on exploring patterns of instructional quality, as conceptualized in PLATO; low inter-rater agreement an issue in most observation studies; internal consistency necessary for valid comparisons</p> <p><b>Art 2:</b> Not likely that instructional repertoires change due to camera or researcher; focus on flexibly capturing classroom discourse</p> <p><b>Art 3:</b> As this is mainly a theoretical and methodological discussion on our experiences along the way, the priority was discussing and illustrating possible biases with empirical examples</p>

Overall, this thesis investigates the phenomenon of patterns of instructional quality by combining aggregated patterns of instruction with an in-depth analysis of teaching (see Klette et al., 2017; Snell, 2011), together with critical discussions of analytical lenses in classroom observations (Goldman, 2007). This combined approach has hopefully contributed to valid and nuanced interpretations of patterns of instructional quality in mathematics classrooms in Helsinki and Oslo contexts.

## 5 Summary of the articles

In this chapter, I summarize the three articles, focusing on the research questions and main findings. The overarching aim of this thesis is to explore, understand, and problematize patterns of instructional quality drawing on video data from lower secondary mathematics classrooms in Helsinki and Oslo, respectively. In Article I, the focus was on *exploring* patterns of instructional quality in the two contexts. In Article II, the focus was on *understanding* two mathematics teachers' rationale for diverging discourse patterns in the Helsinki context. In Article III, the focus was on *problematizing* standardized measures of instructional quality by assessing possible biases when applying the PLATO observation system in Helsinki and Oslo mathematics classrooms.

### 5.1 Article I

Luoto, J. M., Klette, K., & Blikstad-Balas, M. (in review). Patterns of instructional quality in Finnish and Norwegian lower secondary mathematics classrooms. In review at *International Journal of Science and Mathematics Education*.

The rationale for this article was that while Finland and Norway share many educational characteristics in terms of the Nordic model of education (Blossing et al., 2014); some interesting differences do also exist, including differences in student achievements as measured by for example PISA tests. However, there is little updated, empirical research on instruction in mathematics classrooms in Finland, and comparative analyses of instruction from these two contexts are rare. The aim of this sub-study was thus to explore and compare patterns of instructional quality in these two Nordic contexts, focusing on presentation of content, discourse features, and activity formats. The research questions were as follows: *What activity formats do teachers use to engage students? What is the quality of instructional explanations of content, connections to prior knowledge, and setting a purpose for learning? and What characterizes discourse features in mathematics classrooms?*

In this article, we studied instruction in 16 lower secondary mathematics classrooms, eight from Oslo and eight from Helsinki (see Appendix 2). We systematically decomposed features of instructional quality applying the PLATO observation system, and we analyzed video-recorded lessons ( $N = 47$ ) on the topics of numbers, algebra, and geometry. We coded instruction according to six relevant elements from the PLATO manual related to the domains presentation of content and classroom discourse: *Quality of instructional explanations; Richness of instructional explanations; Connections to prior knowledge; Purpose; Uptake of student responses; and Opportunity for student talk* (see Appendix 1). In line with PLATO

scoring system, we analyzed these features of instruction as well as activity format for each 15-min segment of the lessons.

The findings indicated distinct patterns of instructional quality in the two educational contexts. In the Helsinki classrooms, individual seatwork was the most common activity format, and the teachers explicitly presented and framed content clearly and coherently, while students had very limited opportunities to engage in classroom discourse or collaborate with peers. In the Oslo classrooms, there was more variation in activity formats, yet whole-class instruction dominated, and teachers presented content less coherently, while students had more opportunities to engage in classroom discourse.

Instruction in Helsinki was more often coded at the high end of the scale on three of the presentation of content elements (Quality of instructional explanations, Connections to prior knowledge, and Purpose). We would argue that this is connected to the coherence in the Helsinki mathematics lessons, which often follow a similar routine and lesson structure. Generally, this routine includes reviewing homework, the teacher presenting new content/refreshing old content, individual seatwork with teacher assistance, and finally assigning new homework. In contrast, homework, purpose, and references to prior knowledge and lessons were seldom addressed explicitly in the Oslo classrooms. However, instruction in Oslo was more often coded at the high end of Classroom discourse elements, and whole-class discussions were often dynamic and broken up into short individual and group tasks followed by a plenary discussion, providing students with opportunities to engage in discussions. Still, in both contexts, the whole-class discussion were mostly characterized by Initiation-Response-Evaluation/Feedback (IRE/F) formats (see Cazden, 2001), where the students' contribution to discourse is limited to providing the right answer to teachers' questions. The element of Richness of instructional explanations was coded at the same rather low level in both contexts, indicating that most explanations were focusing on procedures and labels.

In conclusion, some patterns were more prominent in one context than in the other. We discussed how contextual differences such as an explicit national focus on improving mathematics instruction in Norway towards more instructional variety (e.g., Bergem, 2014), may offer some explanations to such differences. We concluded that PLATO was useful for detecting differences in patterns across contexts, while qualitative in-depth analyses of what practices that produced the patterns were necessary in order to accurately describe the differences. We also highlighted issues with sequencing lessons, and thus started the discussion that we continued in Article III, stressing how sequencing and operationalization of instructional quality in standardized observations have implications for results and need to be



critically assessed. Finally, in this article, we offered directions for further research, especially regarding combining different perspectives of instructional quality.

### 5.2 Article II

Luoto, J. M. (2020). Scrutinizing two Finnish teachers' instructional rationales and perceived tensions in enacting student participation in mathematical discourse. *LUMAT: International Journal on Math, Science and Technology Education*, 8(1), 133–161. doi:10.31129/LUMAT.8.1.1329

In this article, I focused on two purposefully sampled mathematics teachers in the Helsinki context in order to understand their instructional rationales for different discourse practices. From Article I, Anna and Bea were sampled because of their contrasting classroom discourse practices; Bea provided few opportunities for her students to participate in discourse, while Anna was a unique case who constantly engaged her students in discourse. The rationale for this sub-study was that student participation in classroom discourse has been highlighted as essential for student learning in contemporary mathematics education literature, and is emphasized in the new Finnish curriculum (Finnish National Agency for Education, 2014). Traditionally, discourse has been less emphasized in mathematics instruction in the Finnish context, yet little research has focused on how teachers perceive student participation in discourse in the Finnish context. The overarching research question was: *How do two Finnish mathematics teachers with diverging practices perceive and enact student participation in discourse?* There were also the following sub-questions: *What instructional moves do the two mathematics teachers use to engage students in classroom discourse, and to what extent are these moves used?*, *What is the instructional rationale for the two mathematics teachers' instructional moves in classroom discourse?*, and *What kind of possible tensions do teachers with different practices perceive as hindering or enabling student participation in discourse?*

Drawing on the enacted discourse moves, analyzed with the Furtak and Shavelson (2009) framework, and on the interviews, the results of this case study showed that the teachers enacted very different discourse practices, while their reasons for their practices reflected rather similar values and concerns. The main concerns for both teachers were students learning needs, equity, and students' emotional well-being—and whether or not the teachers perceived them as in conflict with student participation in classroom discourse. Based on these findings, I argued that it is necessary for teacher education and professional development to address teachers' concerns, especially if more discursive practices in line with curriculum standards are desired. I highlighted Anna's instruction as an example of how

teachers can address perceived tensions among student learning needs, equity, and well-being—while still providing opportunities for content-related discussions. However, Anna’s focus on group work for her advanced students and Bea’s focus on individually assisting her basic students’ underscored how the ideals of “quality classroom discourse” are difficult to apply to different classroom contexts with different students. This issue of assessing instructional quality in different classroom contexts and especially in classrooms where teachers provide individual assistance became important for the discussion in Article III.

### 5.3 Article III

Luoto, J. M., Klette, K., & Blikstad-Balas, M (in review). Possible biases in observation systems when applied across contexts: Conceptualizing, operationalizing and sequencing instructional quality. In review at *Educational Assessment, Evaluation and Accountability*.

This article adopted a conceptual and methodological approach, problematizing patterns of instructional quality by focusing on potential biases with standardized observation systems when applied in different classroom contexts. We used the example of the PLATO observation system to examine how aspects of conceptualization and operationalization, together with sequencing of lessons, embody possible biases that may influence scoring due to how instruction is organized and enacted, using empirical examples from Helsinki and Oslo classrooms. The main research question was as follows: *In what way may standardized observation systems of instructional quality produce biases when applied across different contexts?* There were also three sub-questions: *In what way may the conceptualization of instructional quality produce possible biases when applied in different contexts?*, *In what way may the level of operationalization of instructional quality produce biases when applied in different contexts?*, and *In what way may the sampling specifications of instructional quality produce biases when applied in different contexts?*

To discuss possible biases, we differentiated among construct, method, and item biases based on a theoretical framework for cross-cultural studies (van de Vijver & Leung, 1997; van de Vijver & Tanzer, 2004). We discussed how such possible biases relate to context-sensitivity—the extent that the operationalization of instructional features are broad enough to capture somewhat different instructional practices of the same construct (Praetorius et al., 2019). We used the same sample as in Article I, focusing on “odd patterns,” that is, discrepancies between scores across contexts and coders related to the PLATO elements *Intellectual challenge* and *Uptake of student responses*. In addition, we used the *Feedback* element as an example to illustrate issues with sequencing lessons into shorter segments.

We argued that the way teachers organize lessons and activities have implications for measuring instructional quality with all three PLATO elements, and in our sample, a substantial number of the lessons (approximately 60% in Helsinki and 30% of Oslo) were organized as individual seatwork. The way Intellectual challenge is conceptualized and operationalized assumes common activities and tasks, thus making it difficult to assign a score during individual seatwork where students often work on different tasks and teachers differentiate their help. Uptake of student responses is operationalized as conversations involving several students, thus disregarding individual student–teacher talk during individual seatwork. This means that the quality of such student–teacher conversations is not captured, and in settings such as Helsinki, instruction consequently receives a low score for this element. The last element, Feedback, addresses the problem with sequencing lessons into shorter segments (in PLATO 15-min segments), as high scores require a certain number of feedback instances to cluster within a segment; If the same quality and amount of feedback is isolated in different segments instead of clustered, it results in lower scores. We showed how teachers may sequence lessons differently and noted the tendencies of clustered feedback at the end of the lesson in the Oslo context compared with no clustering in the Helsinki context.

In summary, we concluded that the way Intellectual challenge is conceptualized is not a bias, as it does not systematically misrepresent the intended construct—however, PLATO does not in its current operationalization of Intellectual challenge recognize individual seatwork and differentiated instruction, which is common for mathematics instruction in a Nordic context. We further determined that the way Uptake of student responses is operationalized may be biased, as it does not include individual seatwork situations in which teachers use uptake in discussing content with individual students. Finally, we proposed that sequencing of lessons may introduce a method bias (van de Vijver & Tanzer, 2004) in elements such as Feedback if lessons are systematically structured differently across contexts. However, we discerned that clustered feedback may be more important for student learning, in which case it is not a bias. We concluded that, while contradictions of common conceptualizations and context-sensitivity may be unresolvable, there might be the potential for combining and adapting frameworks for increasing context-sensitivity—thus, we need cultural adjustments in operationalization and flexibility in sequencing procedures of common conceptualizations, to cover different types of lesson structures and activity formats. This insight might lead follow up research to deepen the results of Article I, for example by applying Uptake also to discussions during individual seatwork and examine whether and how this would change discourse patterns of instructional quality in a Finnish context.

## 6 Discussion of research contribution

Drawing on the different lenses of exploring, understanding, and problematizing instructional quality, as outlined in Section 1.2, the three articles in this thesis together address the overarching research question: *How can patterns of instructional quality be understood through observation manuals and teachers' perspectives?* In the following, I discuss the main findings of my articles that provide answers to this question. In doing so, I outline the empirical, theoretical, and methodological contributions of this thesis.

### 6.1 Empirical contribution

In sum, the findings of Article I contribute with updated empirical knowledge about patterns of instructional quality in Helsinki and Oslo contexts drawing on a comparative approach. The analyses indicated that, while there are similarities, some patterns of instructional quality are more common to certain contexts, which aligns with other cross-national classroom studies suggesting distinct contextual patterns of instruction (e.g., Stigler & Hiebert, 1999; Alexander, 2000). The clearest similarity across contexts was a procedural focus in teachers' explanations of mathematical content, and the clearest difference was that instruction and learning was more individualized in the Helsinki sample than in the Oslo sample, as Helsinki teachers spent most of their instructional time guiding students individually. Other researchers have also noted the prominence of individual seatwork in Finnish mathematics classrooms, and called it a “constructivist shift” in teaching practices, moving away from whole-class instruction towards the individual learner (Kaasila & Pehkonen, 2009; Krzywacki et al., 2016). However, students in the Helsinki classrooms work in textbook and not with complex real-world problems, and often with strict procedures to follow. Previous studies from the Norwegian context characterized mathematics instruction as dominated by individual seatwork and whole-class instruction (e.g., Klette et al., 2016). Yet, we found that whole-class instruction and group work—activity formats that favor collective rather than individual learning activities—together were the most common activity formats in 75% of the lesson segments. Whole-class sessions in Oslo were thus more dynamic than previously described (e.g., Alseth et al., 2003), and group work more common than previously observed in the PISA+ study, where students mainly worked individually on work plans (Bergem, 2016).

A distinct characterization of instruction in the Helsinki classroom was a coherency in instruction across lessons, where textbooks often guided the focus on content and procedures, while little attention was paid to mathematical communication and justification. Whereas textbook and content focuses have been reported by previous studies (Andrews, 2013a;

Kilhamn & Säljö, 2019; Kupari, 2003; Pehkonen, 2007), the analyses in Article I further decomposed these patterns, showing how presentation of content—especially in terms of instructional explanations, connections to prior knowledge, and purpose—is nested in coherent instructional routines of steadily proceeding through textbooks and regular homework review. In contrast, the findings from the Oslo context indicated less coherent content presentation in terms of rare instances of teachers’ specifically connecting previous and new content, unspecified learning goals, and rare occasions of homework review. This textbook-centered routine in Helsinki, as well as the findings showing differences in whole-class discussion, with dynamic practices interspersed with short individual and group tasks common in Oslo in contrast with teacher-led instruction observed in Helsinki classrooms, indicates distinct instructional patterns and is a good example of how a comparative approach may “make the familiar strange” (Alexander et al., 2000).

It is interesting that the findings from the Helsinki classrooms relate closely to previous descriptions of mathematics instruction in the Finnish context, as my sample was Finnish–Swedish, suggesting that patterns of instructional quality might be more similar than different across the two language groups in Finland. More empirical research is however needed from Finnish-speaking mathematics classrooms to gain knowledge of possible differences across language groups, and to understand whether and how the new curriculum influences teaching and increases variation in activities (see Hemmi et al., 2017). In contrast, the findings from the Oslo classrooms, together with other LISA studies (Klette, 2020), indicate that some instructional changes, such as a greater variety of activities inviting student engagement, may have occurred since many of the previous Norwegian studies were conducted more than a decade ago. Such changes in practice can be seen in light of *The Knowledge Promotion* and its focus on oral competence and ‘talking mathematics’ (Norwegian Ministry of Education and Research, 2013a), as well as other initiatives highlighting variation in how students engage with content and peers (e.g., Bergem, 2014; Norwegian Ministry of Education and Research, 2015). However, more research on teachers’ perspectives of instructional change is needed to confirm such indications.

The descriptive findings in Article I derive from a specific perspective of instructional quality as conceptualized and operationalized in the PLATO manual. Drawing on the context-specific perspectives of Article II and Article III, we add further understanding to these descriptive patterns. The findings in Article II suggest that teachers’ views of purposeful instruction for their specific context and specific students do not necessarily align with instructional quality as it is conceptualized in observation systems. Anna would score high on

classroom discourse elements in observation systems such as PLATO, while Bea would not; yet, Bea's students who were not academically strong in mathematics may very well benefit from the individual guidance and tightly structured instruction in combination with the explicit instruction that Bea provided (see Archer & Hughes, 2011). Hence, teachers' instructional rationale may be seen as a context-sensitive and complementary perspective to instructional quality as measured by observation systems. In addition, the critical approach taken in Article III, highlight how contextual factors, more specifically lesson structure and activity formats, impact scoring, and that for example Uptake of student responses is not captured during individual seatwork. Consequently, while the findings in Article I describes discourse during whole-class sessions, it does not cover the way teachers use uptake in individual student-teacher conversations. While researchers have pointed out that observation systems in their design do not specify how contextual factors may influence scoring (Cohen & Goldhaber, 2016), the examples in Article III clearly demonstrate how context (i.e., lesson structure) influence results. This will be discussed more in-depth in the below sections 6.2 on theoretical contribution and 6.3 on methodological contribution of the thesis.

In addition, the empirical findings in Article II contribute with knowledge about teachers' instructional rationale for differently enacted discursive practices in lower secondary mathematics classrooms. If the objective is to develop teachers' instruction to include more opportunities for students to engage in discussions, which is highlighted in the current Finnish curriculum (Finnish National Agency for Education, 2014), this knowledge is important for supporting teachers' instructional development. My findings indicated that teachers may enact discourses based on their concerns for their students (i.e., concerns for equity, students' learning needs, and student well-being) and teacher educators could consider such concerns when instructing about productive classroom discourse practices. From previous research, we know that discourse practices in the classroom depend on many different factors, including teacher beliefs about mathematics, about learning mathematics, and about the perceived teacher role (e.g., Brendefur & Frykholm, 2000; Fives & Gill, 2015), as well as teachers' beliefs about students (e.g., Sztajn, 2003). We also know that contextual factors shape discourse, including available curriculum materials and other school factors (e.g., Raymond, 1997), as well as specific mathematical topics and specific students (e.g., Ayalon & Even, 2016). However, while teacher perspectives and values are considered key explanations for cross-national differences in instruction (Alexander, 2000; Cai et al., 2009) such contextual factors are often ignored in mathematics education research (Chazan et al., 2016), and this is also the case for how teachers' values and perspectives affect teaching practices in a Nordic

context (Reichenberg, 2018). Thus, in Article II, I extended previous knowledge of teachers' instructional rationales for enacted classroom discourse as well as contributed empirically by highlighting how in a Finnish context, concerns for equity, student well-being, and student learning may be part of very different teachers' instructional rationales for divergent discourse practices. In this societal context with its ethos of equity in education embedded throughout its education system (Simola et al., 2017), while simultaneously subscribing to teaching methods favoring individual seatwork in mathematics (e.g., Article I; Krzywacki et al., 2016); teachers may especially feel a tension in engaging students in mathematical discourse. Here, more empirical research is needed to continue theorizing instructional rationales in Nordic contexts.

## **6.2 Theoretical contribution**

In this thesis, I make one potential theoretical contribution in stressing that, for valid interpretations of results (M. Kane, 2013); we need transparency about the theoretical perspectives in which the operationalization of instructional features are embedded. The choice of observation systems and how broadly or narrowly they define instructional quality have important implications for results. As I have outlined in Chapter 3 in this Extended Abstract, PLATO, as well as other manuals, typically conceptualize instructional quality by horizontally building on multiple learning theories (Tellings, 2001). If another, mathematics specific observation system would have been used, other patterns may have emerged, however, also such patterns would depend on the specific manual's view of instructional quality (Schoenfeld et al., 2018). In clarifying how the different theories are conceptualized and operationalized, I think we can realize more of the potential in cross-national classroom studies, and distinguish what theoretical traditions are valued in different contexts (see Clarke, 2013a). The findings in Article I, however limited, indicated that cognitive and constructivist perspectives seem to be more prominent in the Helsinki mathematics classrooms focusing on individual seatwork and clear and connected content, while socio-constructivist traditions may be (increasingly so) more prominent in Oslo classrooms focusing on collaborative practices and student engagement. Thus, I show how a decomposition and clarification of conceptual perspectives of instructional quality within and across different domains in an observation system might help researchers to interpret and present results that are clearly conceptually grounded not only in its conceptualization but also in its operationalization of constructs (see Schlesinger & Jentsch, 2016). This may in turn lead to less risk of biased interpretations of results.

### **6.3 Methodological contribution**

In this thesis, I make two methodological contributions. The first is that, in Article III, we identified sources of possible biases in the PLATO observation manual, and pointed out a potential mismatch between conceptualization, operationalization, and sequencing in PLATO and the actual instruction in Helsinki and Oslo mathematics classrooms. This underlines the need for adjustments in how instructional elements are operationalized at the level of rubrics, so that we can capture features of instructional quality regardless of the lesson structure and activity format that is in focus. We illustrated that this is necessary in Article III, and suggested that some of PLATO's elements need further refinement in their operationalization in order to be captured independent of lesson structure (i.e., uptake and intellectual challenge in a differentiated classroom organized as individual seatwork). While other scholars have investigated comparability of constructs across contexts (e.g., Maulana et al., 2020), how the operationalization of constructs may embed biases that influence scoring during different lesson structures has, to the best of my knowledge, not previously been addressed in the current literature on instructional quality. In addition, examples of biases in cross-national research have been primarily derived from survey data (van de Vijver & Leung, 1997; van de Vijver & Tanzer, 2004), while Article III used examples from video observations. Hence, this thesis gives new insights into possible biases in instructional quality in standardized observation systems when applied in lessons structured in different ways, which is an issue amplified in cross-national research where different contexts tend to have distinct lesson structures (e.g., Alexander, 2000). This thesis also contributes to a Nordic perspective of instructional quality in mathematics instruction by providing an initial step in synthesizing how observation systems work in Nordic lower secondary mathematics contexts. This step has highlighted some biases that can be addressed in future steps using other observation instruments. This may be seen as a methodological contribution, as it may guide and inspire other classroom researchers to look for "odd patterns" in their data as well as spur designers of observation systems to address and devise solutions for biases related to lesson structure. In addition, I have shown how complementary perspectives, like observation systems with different grain-sizes (PLATO and Furtak & Shavelson, 2009) as well as in-depth perspectives of teachers and critical approaches to conceptualization in observation systems can limit shortcomings of specific observation systems.

A second methodological contribution is that I extended the Furtak and Shavelson (2009) framework of dialogic and authoritative moves described in Article II by adding two inductively derived dialogic moves suited especially to coding group work sessions, and



extended the manual beyond the whole-class sessions it was originally designed for. Such extensions of the uses of frameworks mirror the ambitions of constructing observation frameworks relying on a generic and specific feature of instruction (e.g., here dialogic teaching), while flexibly operationalized to suit and be relevant to different activity formats and classroom contexts.

#### **6.4 Suggestions for future research**

As implied, future research is especially needed to empirically test combinations of standardized and context-sensitive measures of instructional quality, which I in this thesis have suggested as necessary for driving the field forward and making increasingly useful cross-national comparisons of instruction. Context-sensitive and standardized conceptualizations do not necessarily contradict each other, as for example Stigler and Miller (2018) identified common features of quality instruction (e.g., productive struggle, explicit connections, and deliberate practice) in very different cultural contexts. Yet, more work still needs to be done on how to operationalize such broad concepts so that they can be systematically captured even if they are differently enacted. Continuing the work of others using both generic and subject-specific observation systems on the same data set (e.g., Berlin & Cohen, 2020; Praetorius & Charalambous, 2018), focusing on different types of biases, would also inform the field of what observation systems work best for what purposes and in what contexts. Another suggestion is to further investigate teachers' instructional rationales for enacted classroom instruction in Finland as well as across the Nordic context to better understand how teachers' in these contexts perceive classroom discourse, and whether, and in what ways, concerns for student well-being, learning and equity, shape their enacted classroom discourse practice. Further exploring these issues within the Nordic model of education (Blossing et al., 2014), would also increase understanding of what instructional quality is in the larger Nordic context from a qualitative and comparative point of view.

#### **6.5 Concluding remarks**

In this thesis, I have applied three different lenses to study patterns of instructional quality in lower secondary mathematics classrooms in the contexts of Helsinki and Oslo. Taken together, the empirical findings of this thesis can inform researchers and teacher educators as well as policy makers about current trends in mathematics instruction in a sample of Nordic mathematics classrooms. As I noted in the introduction, such empirical knowledge is especially important as political debates in Norway and beyond often highlight Finland as a benchmark country and as a basis for justifications related to educational reform (Breakspear,

2012, Darling-Hammond et al., 2017). Since such debates often pay little attention to differences in enacted classroom practices, the findings of this thesis, together with other empirical studies, may inform and nuance debates by providing knowledge of actual classroom practices (see Simola et al., 2017). Also, this empirical knowledge may inform teacher educators in both contexts about possible areas for targeting instructional development. In addition, this thesis may especially inform teacher educators in the Finnish context about possible tensions teachers may perceive between engaging students in classroom discourse and their values and concerns for students. This is relevant knowledge for the Finnish context but also other contexts that are adopting new curriculum requirements where collaborative and discursive skills are highlighted as skills needed in today's society (Hemmi et al., 2017).

In line with several scholars in the field, (e.g., Grossman & McDonald, 2008; Klette & Blikstad-Balas, 2018; Pianta & Hamre, 2016; Praetorius & Charalambous, 2018), I have in Article III called for harmonizing classroom observation research by creating a common vocabulary across studies around practices that constitute instructional quality. However, as I have discussed throughout different parts of this thesis, there is contention on what constitutes instructional quality, and how instruction enacted in different classroom contexts influences scoring is often insufficiently addressed in observation system designs, resulting in a possible conceptual mismatch between the intended theoretical construct and the enacted instruction. While harmonizing definitions of instructional quality is important for driving the field forward, in order to improve observation systems we should not stop challenging and problematizing established constructs, nor neglect how contextual aspects such as lesson structure or classroom context influence the enacted and measured instructional quality. Instructional quality is a highly complex construct, and for me, harmonizing the field of instructional quality means to adapt standardized and generic measures and combine them with context- and content-specific measures to gain a more context-sensitive understanding of instructional quality, while using the same language to describe instruction as a means for comparison—a balance also stressed by other cross-national researchers (e.g., Alexander, 2010).

Together, the findings of this thesis may add to the work of developing and refining classroom observation systems of instructional quality—useful and relevant across different classrooms as well as across national contexts—and encourage other researchers interested in cross-national observations to be more transparent and open to problematizing the theoretical perspectives embedded in and across their chosen observation manuals.

## References

- Alexander, R. (2000). *Culture and pedagogy: International comparisons in primary education*. Oxford, England: Blackwell.
- Alexander, R. (2008). *Towards dialogic teaching: Rethinking classroom talk* (3rd ed.). Cambridge, England: Dialogos.
- Alexander, R. (2010). Border Crossings: Towards a comparative pedagogy. *Comparative Education*, 37(4), 507–523. doi:10.1080/03050060120091292
- Alexander, R., Broadfoot, P., & Phillips, D. (2000). *Learning from comparing: New directions in comparative educational research* (Vol. 2). Oxford, England: Symposium Books.
- Alseth, B., Breiteig, T., & Brekke, G. (2003). *Endringer og utvikling ved R97 som bakgrunn for videre planlegging og justering: Matematikkfaget som kasus* [Changes and development in R97 as a background for further planning and adjustment: The case of the mathematics subject]. Notodden, Norway: Telemarksforskning.
- Andersson, E., & Klette, K. (2016). Teachers' use of questions and responses to students' contributions during whole class discussions: Comparing language arts and science classrooms. In K. Klette, O. K. Bergem, & A. Roe (Eds.), *Teaching and learning in lower secondary schools in the era of PISA and TIMSS* (pp. 63–84). Cham, Switzerland: Springer International Publishing.
- Andersson, E., & Sørvik, G. O. (2013). Reality lost? Re-use of qualitative data in classroom video studies. *Forum: Qualitative Social Research*, 14(3). doi:10.17169/fqs-14.3.1941
- Andrews, P. (2013a). Finnish mathematics teaching from a reform perspective: A video-based case-study analysis. *Comparative Education Review*, 57(2), 189–211. doi:10.1086/669124
- Andrews, P. (2013b). What does PISA performance tell us about mathematics teaching quality? Case studies from Finland and Flanders. In H.-D. Meyer & A. Benavot (Eds.), *PISA, power and policy: The emergence of global educational governance* (pp. 99–114). Oxford, England: Symposium Books.
- Andrews, P., Ryve, A., Hemmi, K., & Sayers, J. (2014). PISA, TIMSS and Finnish mathematics teaching: An enigma in search of an explanation. *Educational Studies in Mathematics*, 87(1), 7–26. doi:10.1007/s10649-014-9545-3
- Anthony, G., & Walshaw, M. (2009). *Effective pedagogy in mathematics: Educational practices series, 19*. Retrieved from [http://www.ibe.unesco.org/fileadmin/user\\_upload/Publications/Educational\\_Practices/EdPractices\\_19.pdf](http://www.ibe.unesco.org/fileadmin/user_upload/Publications/Educational_Practices/EdPractices_19.pdf)
- Archer, A., & Hughes, C. (2011). *Explicit instruction: Effective and efficient teaching*. New York, NY: The Guilford Press.
- Archer, J., Cantrell, S., Holtzman, S. L., Joe, J. N., Tocci, C. M., & Wood, J. (2016). *Better feedback for better teaching: A practical guide to improving classroom observations*. Bill and Melinda Gates Foundation. Retrieved from <http://k12education.gatesfoundation.org/download/?Num=2659&filename=BetterFeedbackBetterTeaching.pdf>
- Ayalon, M., & Even, R. (2016). Factors shaping students' opportunities to engage in argumentative activity. *International Journal of Science and Mathematics Education*, 14(3), 575–601. doi:10.1007/s10763-014-9584-3

- Ball, D. L., & Hill, H. (2009). Measuring teacher quality in practice. In D. H. Gitomer (Ed.), *Measurement issues and assessment for teaching quality* (pp. 80–98). Los Angeles, CA: SAGE Publishing.
- Barnes, D. (2008). Exploratory talk for learning. In N. Mercer & S. Hodgkinson (Eds.), *Exploring talk in schools* (pp. 1–16). Los Angeles, CA: SAGE Publishing.
- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., . . . Tsai, Y.-M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, *47*(1), 133–180.  
doi:10.3102/0002831209345157
- Bell, C. A., Dobbelaer, M. J., Klette, K., & Visscher, A. (2019). Qualities of classroom observation systems. *School Effectiveness and School Improvement*, 1–27.  
doi:10.1080/09243453.2018.1539014
- Bell, C. A., Gitomer, D. H., McCaffrey, D. F., Hamre, B. K., Pianta, R. C., & Qi, Y. (2012). An argument approach to observation protocol validity. *Educational Assessment*, *17*(2–3), 62–67. doi:10.1080/10627197.2012.715014
- Bell, C. A., Qi, Y., Croft, A. J., Leusner, D., McCaffrey, D., Gitomer, D. H., & Pianta, R. C. (2014). Improving observational score quality: Challenges in observer thinking. In T. J. Kane, K. A. Kerr, & R. C. Pianta (Eds.), *Designing teacher evaluation systems: New guidance from the Measures of Effective Teaching project* (pp. 50–81). San Francisco, CA: Jossey-Bass.
- Bergem, O. K. (2009). *Individuelle versus kollektive arbeidsformer: En drøfting av aktuelle utfordringer i grunnskolen* [Individual versus collective activity formats: A discussion of prominent challenges in compulsory school]. (Doctoral dissertation, Oslo University). Oslo, Norway: Unipub.
- Bergem, O. K. (2014) Realfag: Relevant engasjerende attraktive lærerike [STEM education: Relevant engaging attractive educational]. In *Rapport fra Ekspertgruppa for realfagene*. Retrieved from [https://www.regjeringen.no/globalassets/upload/kd/vedlegg/rapporter/rapport\\_fra\\_ekspertgruppa\\_for\\_realfagene.pdf](https://www.regjeringen.no/globalassets/upload/kd/vedlegg/rapporter/rapport_fra_ekspertgruppa_for_realfagene.pdf)
- Bergem, O. K. (2016). Usually we are not where the teacher is. In K. Klette, O. K. Bergem, & A. Roe (Eds.), *Teaching and learning in lower secondary schools in the era of PISA and TIMSS* (pp. 47–61). Cham, Switzerland: Springer International Publishing.
- Bergem, O. K., & Klette, K. (2010). Mathematical tasks as catalysts for student talk: Analysing discourse in a Norwegian mathematics classroom. In Y. Shimizu, B. Kaur, R. Huang, & D. Clarke (Eds.), *Mathematical tasks in classrooms around the world* (pp. 35–62). Rotterdam, Netherlands: Sense Publishers.
- Berlin, R., & Cohen, J. (2020). The convergence of emotionally supportive learning environments and college and career ready mathematical engagement in upper elementary classrooms. *AERA Open* *6*(3), 1-20. doi:10.1177/2332858420957612
- Berliner, D. C. (1987). Simple views of effective teaching and a simple theory of classroom instruction. In D. C. Berliner & B. Rosenshine (Eds.), *Talks to teachers* (pp. 93–110). New York, NY: Random House.
- Blikstad-Balas, M. (2017). Key challenges of using video when investigating social practices in education: Contextualization, magnification, and representation. *International Journal of Research & Method in Education*, *40*(5), 511–523.  
doi:10.1080/1743727X.2016.1181162
- Blömeke, S., Olsen, R. V., & Suhl, U. (2016). Relation of student achievement to the quality of their teachers and instructional quality. In T. Nilsen & J.-E. Gustafsson (Eds.), *Teacher quality, instructional quality and student outcomes: Relationships across*

- countries, cohorts and time* (pp. 21–50). Cham, Switzerland: Springer International Publishing.
- Blossing, U., Imsen, G., & Moos, L. (2014). *The Nordic education model: A school for all encounters neo-liberal policy* (Vol. 1). Dordrecht, Netherlands: Springer.
- Boston, M. (2012). Assessing instructional quality in mathematics. *The Elementary School Journal*, *113*(1), 76–104. doi:10.1086/666387
- Breakspear, S. (2012). The policy impact of PISA: An exploration of the normative effects of international benchmarking in school system performance. *OECD Education Working Papers, No. 71*. Paris, France: OECD Publishing.  
<https://doi.org/10.1787/5k9fdfqffr28-en>.
- Brendefur, J., & Frykholm, J. (2000). Promoting mathematical communication in the classroom: Two preservice teachers' conceptions and practices. *Journal of Mathematics Teacher Education*, *3*(2), 125–153. doi:10.1023/A:1009947032694
- Brophy, J. E. (1986). Teacher influences on student achievement. *American Psychologist*, *41*(10), 1069–1077. Retrieved from <https://doi.org/10.1037/0003-066X.41.10.1069>
- Brophy, J. E., & Good, T. L. (1984). *Teacher behavior and student achievement*. East Lansing, MI: Michigan State University.
- Brunell, V. (2007). *Klimat och resultat i den finlandssvenska grundskolan: En fordjupad analys av PISA 2003* [Climate and results in the Swedish-speaking Finnish compulsory school: An in-depth analysis of PISA 2003] Retrieved from [http://epublications.uef.fi/pub/urn\\_isbn\\_978-952-61-0266-5/urn\\_isbn\\_978-952-61-0266-5.pdf](http://epublications.uef.fi/pub/urn_isbn_978-952-61-0266-5/urn_isbn_978-952-61-0266-5.pdf)
- Cai, J., Kaiser, B. Perry, & N.-Y. Wong (2009). *Effective mathematics teaching from teachers' perspectives: National and cross-national studies*. Rotterdam, Netherlands: Sense Publishers.
- 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.  
doi:10.1080/00313830600743357
- Casabianca, J. M., McCaffrey, D. F., Gitomer, D. H., Bell, C. A., Hamre, B. K., & Pianta, R. C. (2013). Effect of observation mode on measures of secondary mathematics teaching. *Educational and Psychological Measurement*, *73*(5), 757–783.  
doi:10.1177/0013164413486987
- Cazden, C. B. (2001). *Classroom discourse: The language of teaching and learning*. (2<sup>nd</sup> ed.). Portsmouth, NH: Heinemann.
- Charalambous, C. Y., & Praetorius, A-K. (2018). Studying mathematics instruction through different lenses: Setting the ground for understanding instructional quality more comprehensively. *ZDM*, *50*(3), 355–366. doi:10.1007/s11858-018-0914-8
- Chazan, D., Herbst, P. G., & Clark, L. M. (2016). Research on the teaching of mathematics: A call to theorize the role of society and schooling in mathematics instruction. In D. H. Gitomer & C. A. Bell (Eds.), *Handbook of research on teaching* (Vol. 5, pp. 1039–1098). Washington, DC: American Educational Research Association.
- Clarke, D. (2013a). Contingent conceptions of accomplished practice: The cultural specificity of discourse in and about the mathematics classroom. *The International Journal on Mathematics Education*, *45*(1), 21–33. doi:10.1007/s11858-012-0452-8
- Clarke, D. (2013b). International comparative research into educational interaction. In K. Tirri & E. Kuusisto (Eds.), *Interaction in educational domains* (pp. 5–22). Rotterdam, Netherlands: Sense Publishers.

- Clarke, D., & Chan, E. (2019). The use of video in classroom research: Window, lens, or mirror. In L. Xu, G. Aranda, W. Widjaja, & D. Clarke (Eds.), *Video-based research in education: Cross-disciplinary perspectives* (pp. 5–18). New York, NY: Routledge.
- Clarke, D., Emanuelsson, J., Jablonka, E., & Mok, I. (2006). *Making connections: Comparing mathematics classrooms around the world* (Vol. 2). Rotterdam, Netherlands: Sense Publishers.
- Clarke, D., Mesiti, C., Jablonka, E., & Shimizu, Y. (2006). Addressing the challenge of legitimate international comparisons: Lesson structure in the USA, Germany and Japan. In D. Clarke, J. Emanuelsson, E. Jablonka, & I. Mok (Eds.), *Making connections: Comparing mathematics classrooms around the world* (pp. 23–46). Rotterdam, Netherlands: Sense Publishers.
- Clarke, D., Mesiti, C., O’Keefe, C., Xu, L., Jablonka, E., & Mok, I. (2007). Addressing the challenge of legitimate international comparisons of classroom practice. *International Journal of Educational Research*, 46(5), 280–293.
- Cobb, P. (1994). Where is the mind? Constructivist and sociocultural perspectives on mathematical development. *Educational Researcher*, 23(7), 13–20.  
doi:10.3102/0013189X023007013
- Cobb, P. (2007). Putting philosophy to work: Coping with multiple theoretical perspectives. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 3–38). Charlotte, NC: Information Age Publishing.
- Cobb, P. (2011). Radical constructivism: Introduction. In E. Yackel, K. Gravemeijer, & A. Sfard (Eds.), *A journey in mathematics education research: Insights from the work of Paul Cobb* (pp. 9–17). Dodrecht, Netherlands: Springer.
- Cobb, P., & Bowers, J. (1999). Cognitive and situated learning perspectives in theory and practice. *Educational Researcher*, 28(2), 4–15. doi:10.2307/1177185
- Cobb, P., & Steffe, L. P. (1983). The constructivist researcher as teacher and model builder. *Journal for Research in Mathematics Education*, 14(2), 83–94. doi:10.2307/748576
- Cobb, P., Stephan, M., McClain, K., & Gravemeijer, K. (2001). Participating in classroom mathematical practices. *Journal of the Learning Sciences*, 10(1–2), 113–163.  
doi:10.1207/S15327809JLS10-1-2\_6
- Cobb, P., & Yackel, E. (1996). Constructivist, emergent, and sociocultural perspectives in the context of developmental research. *Educational Psychologist*, 31(3–4), 175–190.  
doi:10.1080/00461520.1996.9653265
- Cohen, J. (2013). *Practices that cross disciplines?: A closer look at instruction in elementary math and English language arts* (Doctoral dissertation, Stanford University).
- Cohen, J., & Goldhaber, D. (2016). Observations in evaluating teacher performance: Assessing the strengths and weaknesses of classroom observations and value-added measures. In J. A. Grissom & P. Youngs (Eds.), *Improving teacher evaluation systems: Making the most of multiple measures* (pp. 8–21). New York, NY: Teachers College Press.
- Cohen, L., Manion, L., & Morrison, K. R. B. (2018). *Research methods in education* (8th ed.). London, England: Routledge.
- Cor, M. K. (2011). *Investigating the reliability of classroom observation protocols: The case of PLATO*. Retrieved from [http://platorubric.stanford.edu/Cor%20M%20K%20%20\(2011\).pdf](http://platorubric.stanford.edu/Cor%20M%20K%20%20(2011).pdf)
- Creemers, B. P. M., Reynolds, D., Stringfield, S., & Teddlie, C. (2002). *World class schools: International perspectives on school effectiveness*. London, England: Routledge Falmer.
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Los Angeles, CA: SAGE Publishing.

- Dalland, C. P. (2011). Utfordringer ved gjenbruk av andres kvalitative data [Challenges in re-using other researchers' qualitative data]. *Norsk pedagogisk tidsskrift*, 95(6), 449–459.
- Danielson, C. (2013). *The framework for teaching evaluation instrument, 2013 instructionally focused edition*. Retrieved from <https://danielsongroup.org/downloads/2013-framework-teaching-evaluation-instrument>
- Darling-Hammond, L., Burns, D., Campbell, C., Goodwin, A., Hammerness, K., Low, E.-L., . . . Zeichner, K. (2017). *Empowered educators: How high performing systems shape teaching around the world*. San Francisco, CA: Jossey-Bass.
- Denscombe, M. (2014). *The good research guide*. Maidenhead, England: Open University Press.
- Derry, S. J. (2007). Video research in classroom and teacher learning (Standardize that!). In R. Goldman, R. Pea, B. Barron, & S. J. Derry (Eds.), *Video research in the learning sciences* (pp. 305–320). New York, NY: Routledge.
- Derry, S. J., Pea, R. D., Barron, B., Engle, R. A., Erickson, F., Goldman, R., . . . Sherin, B. L. (2010). Conducting video research in the learning sciences: Guidance on selection, analysis, technology, and ethics. *Journal of the Learning Sciences*, 19(1), 3–53. doi:10.1080/10508400903452884
- Edwards, D., & Mercer, N. (1987). *Common knowledge: The development of understanding in the classroom*. London, England: Methuen.
- Erickson, F. (2006). Definition and analysis of data from videotape: Some research procedures and their relations. In J. L. Green, G. Camilli, P. B. Elmore, A. Skukauskaitė, & E. Grace (Eds.), *Handbook of complementary methods in education research*. Mahwah, NJ: Lawrence Erlbaum.
- Erickson, F. (2007). Ways of seeing video: Toward a phenomenology of viewing minimally edited footage. In R. Goldman, R. Pea, B. Barron, & S. J. Derry (Eds.), *Video research in the learning sciences* (pp. 145–155). New York, NY: Routledge.
- Fauskanger, J. (2016). Matematikklæreres oppfatninger om ingrediensene i god matematikkundervisning [Mathematics teachers' perceptions about the ingredients in good mathematics instruction]. *Acta didactica Norge [elektronisk ressurs]*, 10(3), 18–18. doi:10.5617/adno.2560
- Finnish National Agency for Education. (2014). *Grunderna för läroplanen för den grundläggande utbildningen 2014* [The National Core Curriculum for Basic Education]. Retrieved from [http://www.oph.fi/lp2016/grunderna\\_for\\_laroplanen](http://www.oph.fi/lp2016/grunderna_for_laroplanen)
- Finnish National Board of Education. (2004). *Grunderna för läroplanen för den grundläggande utbildningen* [The National Core Curriculum for Basic Education]. Helsinki, Finland: Utbildningsstyrelsen. Retrieved from <http://www02.oph.fi/svenska/ops/grundskola/LPgrundl.pdf>
- Finnish National Board of Education. (2016). *New National Core Curriculum for Basic Education: Focus on school culture and integrative approach*. Retrieved from <https://www.oph.fi/sites/default/files/documents/new-national-core-curriculum-for-basic-education.pdf>
- Fischer, J., Praetorius, A.-K., & Klieme, E. (2019). The impact of linguistic similarity on cross-cultural comparability of students' perceptions of teaching quality. *International Journal of Policy, Practice and Research*, 31(2), 201–220. doi:10.1007/s11092-019-09295-7
- Fisher, C., Berliner, D. C., Filby, N., Marliave, R., Cahen, L., & Dishaw, M. (1980). Teaching behaviors, academic learning time, and student achievement: An overview. In C. Denham & A. Lieberman (Eds.), *Time to learn* (pp. 7–32). Washington, DC: National Institute of Education.

- Fives, H., & Gill, M. G. (2015). *International handbook of research on teachers beliefs* (Vol. 13). New York, NY: Routledge.
- Franke, M. L., Kazemi, E., & Battey, D. (2007). Understanding teaching and classroom practice in mathematics. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 225–256). Charlotte, NC: Information Age Publishing.
- Furtak, E. M., & Shavelson, R. J. (2009). Guidance, conceptual understanding, and student learning: An investigation of inquiry-based teaching in the US. In T. Janik & T. Seidel (Eds.), *The power of video studies in investigating teaching and learning in the classroom* (pp. 181–205). Münster, Germany: Waxmann.
- Gage, N. L., & Needels, M. C. (1989). Process-product research on teaching: A review of criticisms. *The Elementary School Journal*, 89(3), 253–300. doi:10.1086/461577
- Givvin, K., Hiebert, J., Jacobs, J., Hollingsworth, H., & Gallimore, R. (2005). Are there national patterns of teaching? Evidence from the TIMSS 1999 Video Study. *Comparative Education Review*, 49(3), 311–343. doi:10.1086/430260
- Glasersfeld, E. V. (1995). A constructivist approach to teaching. In L. P. Steffe & J. E. Gale (Eds.), *Constructivism in education* (pp. 3–15). Hillsdale, MI: Lawrence Erlbaum.
- Goldman, R. (2007). Video representations and the perspective framework: Epistemology, ethnography, evaluation and ethics. In R. Goldman, R. Pea, B. Barron, & S. J. Derry (Eds.), *Video research in the learning sciences* (pp. 3–39). New York, NY: Routledge.
- Government of Norway. (2017). Promotion of the status and quality of teachers. Retrieved from <https://www.regjeringen.no/en/topics/education/innsikt/larerloftet/id2008159/>
- Graham, M., Milanowski, A., & Miller, J. (2012). *Measuring and promoting inter-rater agreement of teacher and principal performance ratings*. Center for Educator Compensation Reform. Retrieved from <https://files.eric.ed.gov/fulltext/ED532068.pdf>
- Grossman, P. (2015). *Protocol for language arts teaching observations (PLATO 5.0)*. Palo Alto, CA: Stanford University.
- Grossman, P., Cohen, J., Ronfeldt, M., & Brown, L. (2014). The test matters: The relationship between classroom observation scores and teacher value added on multiple types of assessment. *Educational Researcher*, 43(6), 293–303. doi:10.3102/0013189X14544542
- Grossman, P., & McDonald, M. (2008). Back to the future: Directions for research in teaching and teacher education. *American Educational Research Journal*, 45(1), 184–205. doi:10.3102/0002831207312906
- Grønmo, L. S. (2017). Et matematikdidaktisk perspektiv [A mathematics didactics perspective]. In A. Hole & L. S. Grønmo (Eds.), *Prioritering og progresjon i skolematematikken* [Priorities and progression in school mathematics] (pp. 45–61). Oslo: Cappelen Damm Akademisk.
- Gwet, K. L. (2014). *Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among raters* (Vol. 2). Gaithersburg, MD: Advanced Analytics, LLC.
- Hammersley, M. (1997). Qualitative data archiving: Some reflections on its prospects and problems. *Sociology*, 31(1), 131–142. doi:10.1177/0038038597031001010
- Hammersley, M. (2010a). Can we re-use qualitative data via secondary analysis? Notes on some terminological and substantive issues. *Sociological Research Online*, 15(1). doi:10.5153/sro.2076
- Hammersley, M. (2010b). Is social measurement possible, and is it necessary? In G. Walford, E. Tucker, & M. Viswanathan (Eds.), *The SAGE handbook of measurement* (pp. 409–426). London, England: SAGE Publications.



- Hammond, L., & Moore, W. M. (2018). Teachers taking up explicit instruction: The impact of a professional development and directive instructional coaching model. *Australian Journal of Teacher Education*, 43(7), 110. doi:10.14221/ajte.2018v43n7.7
- Hardman, F., & Hardman, J. (2017). Observing and recording classroom processes. In D. Wyse, N. S. Selwyn, E., & L. E. Sutter (Eds.), *The BERA/SAGE handbook of educational research* (Vol. 2, pp. 572–588). London, England: SAGE Publications.
- Harju-Luukkanen, H., Nissinen, K., Stolt, S., & Vettenranta, J. (2014). *PISA 2012: Resultatnivån i de svenskspråkiga skolorna i Finland* [PISA 2012: Result level in Swedish-speaking schools in Finland]. Retrieved from <http://www.kulturfonden.fi/wp-content/uploads/2016/01/PISA-2012.pdf>
- Hattie, John, & Timperley, Helen. (2016). The power of feedback. *Review of Educational Research*, 77(1), 81–112. doi:10.3102/003465430298487
- Hemmi, K., Krzywacki, H., & Partanen, A.-M. (2017). Mathematics curriculum: The case of Finland. In D. R. Thompson, M. A. Huntley, & C. Suurtamm (Eds.), *International perspectives on mathematics curriculum* (pp. 71–102). Charlotte, NC: Information Age Publishing.
- Herbst, P., & Chazan, D. (2017). The role of theory development in increasing subject specificity of research on mathematics teaching. In Cai, J. (Ed.), *Compendium for research in mathematics education* (pp. 102–127). Reston, VA: National Council of Teachers of Mathematics.
- Hesse-Biber, S. N., Rodriguez, D. & Frost, N. A. (2015). A qualitatively driven approach to multimethod and mixed method research. In S. N. Hesse-Biber, & R. B. Johnson (Eds.), *The Oxford handbook of multimethod and mixed methods research inquiry*. New York, NY: Oxford University Press.
- Hiebert, J., Gallimore, R., Garnier, H., Givvin, K. B., Hollingsworth, H., Jacobs, J., . . . Stigler, J. (2003). Teaching mathematics in seven countries: Results from the TIMSS 1999 video study. *Education Statistics Quarterly*, 5(1), 7–15.
- Hiebert, J., & Grouws, A. D. (2007). The effects of classroom mathematics teaching on students' learning. In Lester, F. K. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 371–404). Charlotte, NC: Information Age.
- Hill, H. C., Blunk, M. L., Charalambous, C. Y., Lewis, J. M., Phelps, G. C., Sleep, L., & Ball, D. L. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study. *Cognition and Instruction*, 26(4), 430–511. doi:10.1080/07370000802177235
- Hill, H. C., Charalambous, C. Y., Blazar, D., McGinn, D., Kraft, M. A., Beisiegel, M., . . . Lynch, K. (2012). Validating arguments for observational instruments: Attending to multiple sources of variation. *Educational Assessment: Measurement Issues in the Study of Teaching*, 17(2–3), 88–106. doi:10.1080/10627197.2012.715019
- Hill, H. C., Charalambous, C. Y., & Kraft, M. A. (2012). When rater reliability is not enough: Teacher observation systems and a case for the generalizability study. *Educational Researcher*, 41(2), 56–64. doi:10.3102/0013189X12437203
- Ho, A. D., & Kane, T. J. (2013). *The reliability of classroom observations by school personnel*. Retrieved from <https://files.eric.ed.gov/fulltext/ED540957.pdf>
- Jacobs, J. K., Hollingsworth, H., & Givvin, K. B. (2007). Video-based research made “easy”: Methodological lessons learned from the TIMSS video studies. *Field Methods*, 19(3), 284–299. doi:10.1177/1525822X07302106
- Jaworski, B. (2006). Theory and practice in mathematics teaching development: Critical inquiry as a mode of learning in teaching. *Journal of Mathematics Teacher Education*, 9(2), 187. doi:10.1007/s10857-005-1223-z

- Jordan, B., & Henderson, A. (1995). Interaction analysis: Foundations and practice. *Journal of the Learning Sciences*, 4(1), 39–103. doi:10.1207/s15327809jls0401\_2
- Kaasila, R., & Pehkonen, E. (2009). Effective mathematics teaching in Finland through the eyes of elementary student teachers. In J. K. Cai, G. Kaiser, B. Perry, & N. Wong (Eds.), *Effective mathematics teaching from teachers' perspectives* (pp. 203–216). Rotterdam, Netherlands: Sense Publishers.
- Kane, M. (2013a). The argument-based approach to validation. *School Psychology Review*, 42(4), 448–457. doi:10.1080/02796015.2013.12087465
- Kane, M. (2013b). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50(1), 1–73. doi:10.1111/jedm.12000
- Kane, T. J., & Staiger, D. O. (2012). *Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains*. Bill and Melinda Gates Foundation. Retrieved from <http://k12education.gatesfoundation.org/resource/gathering-feedback-on-teaching-combining-high-quality-observations-with-student-surveys-and-achievement-gains-3/>
- Kieran, C., Forman, E., & Sfard, A. (2001). Learning discourse: Sociocultural approaches to research in mathematics education. *Educational Studies in Mathematics*, 46(1), 1–12. doi:10.1023/A:1014276102421
- Kilhamn, C., & Säljö, R. (2019). *Encountering algebra: A comparative study of classrooms in Finland, Norway, Sweden, and the USA*. Cham, Switzerland: Springer International Publishing.
- Kilpatrick, J. (1992). A history of research in mathematics education. In A. D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 3–38). New York, NY: Macmillan.
- Kim, H. S. (2002). We talk, therefore we think? A cultural analysis of the effect of talking on thinking. *Journal of Personality and Social Psychology*, 83(4), 828–842. doi:10.1037//0022-3514.83.4.828
- Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, 41(2), 75–86. doi:10.1207/s15326985ep4102\_1
- Klette, K. (2009). Challenges in strategies for complexity reduction in video studies. Experiences from the PISA+ study: A video study of teaching and learning in Norway. In T. Janik & T. Seidel (Eds.), *The power of video studies in investigating teaching and learning in the classroom* (pp. 61–83). Münster, Germany: Waxmann.
- Klette, K. (2015). Introduction: Studying interaction and instructional patterns in classrooms. In K. Klette, O. K. Bergem, & A. Roe (Eds.), *Teaching and learning in lower secondary schools in the era of PISA and TIMSS* (pp. 1–16): Cham, Switzerland: Springer International Publishing.
- Klette, K. (2020). Hva vet vi om god undervisning? Rapport fra klasseromsforskningen [What do we know about good instruction? A report from classroom research]. In R. J. Krumsvik & R. Säljö (Eds.), *Praktisk-pedagogisk utdanning: En antologi* [Practice-based teacher education: An anthology] (Vol. 2, pp. 183–214). Bergen, Norway: Fagbokforlaget.
- Klette, K. (in press). Towards programmatic research when studying classroom teaching and learning. In L. Ligozat, A. Rakhkochkine, & J. Almqvist J (Eds.). *Thinking through didactics in a changing world. European perspectives on learning, teaching and the curriculum*. Routledge Education.

- Klette, K., Bergem, O. K., & Roe, A. (2016). *Teaching and learning in lower secondary schools in the era of PISA and TIMSS* (Vol. 12). Cham, Switzerland: Springer International Publishing.
- Klette, K., & Blikstad-Balas, M. (2018). Observation manuals as lenses to classroom teaching: Pitfalls and possibilities. *European Educational Research Journal*, 17(1), 1474904117703228. doi:10.1177/1474904117703228
- Klette, K., Blikstad-Balas, M., & Roe, A. (2017). Linking instruction and student achievement: Research design for a new generation of classroom studies. *Acta Didactica*, 11(3), 19–19. doi:http://dx.doi.org/10.5617/adno.4729
- Klette, K., Lie, S., Ødegaard, M., Anmarkrud, Ø., Arnesen, N., Bergem, O. K., & Roe, A. (2008). *PISA+: Lærings- og undervisningsstrategier i skolen* [PISA+: Learning and teaching strategies in school]. Oslo, Norway: Norges Forskningsråd.
- Klieme, E., Pauli, C., & Reusser, K. (2009). The Pythagoras study: Investigating effects of teaching and learning in Swiss and German mathematics classrooms. In T. Janík & T. Seidel (Eds.), *The power of video studies in investigating teaching and learning in the classroom*. Münster, Germany: Waxmann.
- Kovalainen, M., & Kumpulainen, K. (2005). The discursive practice of participation in an elementary classroom community. *Instructional Science: An International Journal of Learning and Cognition*, 33(3), 213–250. doi:10.1007/s11251-005-2810-1
- 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: The principles and practices of teaching and learning in Finnish schools* (Vol. 2, pp. 109–123). Leiden, Netherlands: Sense Publishers.
- Kunter, M., Baumert, J., & Köller, O. (2007). Effective classroom management and the development of subject-related interest. *Learning and Instruction*, 17(5), 494–509. doi:10.1016/j.learninstruc.2007.09.002
- Kupari, P. (2003). Instructional practices and teachers' beliefs in Finnish mathematics education. *Studies in Educational Evaluation*, 29(3), 243–257. doi:http://dx.doi.org/10.1016/S0191-491X(03)90012-5
- Kvale, S., & Brinkmann, S. (2009). *Interviews: Learning the craft of qualitative research interviewing* (2nd ed.). Los Angeles, CA: SAGE Publishing.
- Lampert, M., & Blunk, M. L. (1998). *Talking mathematics in school: Studies of teaching and learning*. Cambridge, England: Cambridge University Press.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, England: Cambridge University Press.
- Lemke, J. L. (2007). Video epistemology in- and outside the box: Traversing attentional spaces. In R. Goldman, R. Pea, B. Barron, & S. J. Derry (Eds.), *Video research in the learning sciences* (pp. 39–53). New York, NY: Routledge.
- Lerman, S. (1996). Intersubjectivity in mathematics learning: A challenge to the radical constructivist paradigm? *Journal for Research in Mathematics Education*, 27(2), 133–150. doi:10.2307/749597
- Lerman, S. (2001). Cultural, discursive psychology: A sociocultural approach to studying the teaching and learning of mathematics. *Educational Studies in Mathematics*, 46(1), 87–113. doi:10.1023/A:1014031004832
- Lerman, S. (2019). Theoretical aspects of doing research in mathematics education: An argument for coherence. In G. Kaiser & N. Presmeg (Eds.), *Compendium for early career researchers in mathematics education* (pp. 309–324). Cham, Switzerland: Springer International Publishing.

- Liu, S., Bell, C., Jones, N. McCaffley, D. (2019). Classroom observation systems in context: A case for the validation of observation systems. *Educational Assessment, Evaluation and Accountability*, 31(1), 61–95. doi:10.1007/s11092-018-09291-3
- LUMA Suomi. (2020). *Luma Suomi ohjelma: Valtakunnallinen luonnontieteiden ja matematiikan esi- ja perusopetuksen kehittämisohjelma 2014-2019* [Luma Finland: A national development program for science and mathematics pre- and basic education 2014-2019]. Retrieved from <https://suomi.luma.fi/>
- Lundahl, L. (2016). Equality, inclusion and marketization of Nordic education: Introductory notes. *Research in Comparative and International Education*, 11(1), 3–12. doi:10.1177/1745499916631059
- Mangold. (2019). Behavior research with Mangold INTERACT. Retrieved from <https://www.mangold-international.com/>
- Martinez, F., Taut, S., & Schaaf, K. (2016). Classroom observation for evaluating and improving teaching: An international perspective. *Studies in Educational Evaluation*, 49, 15–29. doi:10.1016/j.stueduc.2016.03.002
- Maulana, R., Ané, S. C. H., Helms-Lorenz, M., Ko, J., Chun, S., Shahzad, A., . . . Fadhilah, N. (2020). Observed teaching behaviour in secondary education across six countries: Measurement invariance and indication of cross-national variations. *School Effectiveness and School Improvement*, 1–32. doi:10.1080/09243453.2020.1777170
- Mercer, N., & Hodgkinson, S. (2008). *Exploring talk in schools: Inspired by the work of Douglas Barnes*. London, England: SAGE Publications.
- Meyer, J. P., Cash, A. H., & Mashburn, A. (2011). Occasions and the reliability of classroom observations: Alternative conceptualizations and methods of analysis. *Educational Assessment*, 16(4), 227–243. doi:10.1080/10627197.2011.638884
- Miller, K., & Zhou, X. (2007). Learning from classroom video: What makes it compelling and what makes it hard. In R. Goldman, R. Pea, B. Barron, & S. J. Derry (Eds.), *Video research in the learning sciences* (pp. 321–334). New York, NY: Routledge.
- Mølsted, C. E., & Karseth, B. (2016). National curricula in Norway and Finland: The role of learning outcomes. *European Educational Research Journal*, 15(3), 329–344. doi:10.1177/1474904116639311
- Morgan, D. L. (2007). Paradigms lost and pragmatism regained: Methodological implications of combining qualitative and quantitative methods. *Journal of Mixed Methods Research*, 1(1), 48–76. doi:10.1177/2345678906292462
- Mortimer, E., & Scott, P. (2003). *Meaning making in secondary science classrooms*. Maidenhead, England: McGraw-Hill Education.
- Muijs, D., Reynolds, D., Sammons, P., Kyriakides, L., Creemers, B., & Teddlie, C. (2018). Assessing individual lessons using a generic teacher observation instrument: How useful is the International System for Teacher Observation and Feedback (ISTOF)? *Mathematics Education*, 50(3), 395–406. doi:10.1007/s11858-018-0921-9
- Munthe, E., & Rogne, M. (2016). Norwegian teacher education: Development, steering and current trends. In B. Moon (Ed.), *Do universities have a role in the education and training of teachers? An international analysis of policy and practice* (pp. 35–55). Cambridge, England: Cambridge University Press.
- NESH. (2016). *Guidelines for research ethics in the social sciences, humanities, law and theology*. The Norwegian national committees for research ethics. Retrieved from <https://www.etikkom.no/forskningsetiske-retningslinjer/Samfunnsvitenskap-jus-og-humaniora/>
- Nilsen, T., & Gustafsson, J.-E. (2016). *Teacher quality, instructional quality and student outcomes: Relationships across countries, cohorts and time* (Vol. 2). Cham, Switzerland: Springer International Publishing.

- Nortvedt, G. (2018). Policy impact of PISA on mathematics education: The case of Norway. *A Journal of Education and Development*, 33(3), 427–444. doi:10.1007/s10212-018-0378-9
- Norwegian Directorate for Education and Training. (2020). Nasjonale prøver 8. og 9. trinn – resultater [National tests in 8th and 9th grade - results]. Retrieved from <https://www.udir.no/tall-og-forskning/statistikk/statistikk-grunnskole/nasjonale-prover-8.-og-9.-trinn/>
- Norwegian Ministry of Education and Research. (2013a). *Kunnskapsløftet* [Knowledge promotion]. Oslo, Norway: Udir.
- Norwegian Ministry of Education and Research. (2013b). *Læreplan i matematikk fellesfag* [Curriculum for mathematic education]. Retrieved from <https://www.udir.no/kl06/MAT1-04>
- Norwegian Ministry of Education and Research. (2015). *NOU 2015: 8 The school of the future - Renewal of subjects and competences*. Retrieved from Government.no: <https://www.regjeringen.no/en/dokumenter/nou-2015-8/id2417001/>
- Nowak, S. (1977). The strategy of cross-national survey research for the development of social theory. In A. Szalai & R. Petrella (Eds.), *Cross-national comparative survey research: Theory and practice* (pp. 3–47). Oxford, England: Pergamon Press.
- Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis*, 26(3), 237–257. doi:10.3102/01623737026003237
- Nystrand, M., Gamoran, A., Kachur, R., Prendergast, C., & Gundlach, R. (1997). *Opening dialogue: Understanding the dynamics of language and learning in the English classroom*. New York, NY: Teachers College Press.
- OECD. (2004). *Learning for tomorrow's world: First results from PISA 2003*. Retrieved from <https://www.oecd.org/education/school/programmeforinternationalstudentassessmentpisa/34002216.pdf>
- OECD. (2016a). *PISA 2015 Results (Volume I): Excellence and equity in education*: Paris, France: OECD Publishing.
- OECD. (2016b). *PISA 2015 Results in Focus*. Paris, France: OECD Publishing.
- OECD. (2019). *PISA 2018 Results (Volume I): What students know and can do*. Paris, France: OECD Publishing. Retrieved from <https://doi.org/10.1787/5f07c754-en>.
- Oser, F. K., & Baeriswyk, F. J. (2001). Choreographies of teaching: Bridging instruction to learning. In V. Richardson (Ed.), *Handbook of research on teaching* (Vol. 4, pp. 1031–1065). Washington, DC: American Educational Research Association.
- Østerud, S. (2016). Hva kan norsk skole lære av PISA-vinneren Finland? [What can the Norwegian school learn from the PISA winner Finland?]. *Nordisk Tidsskrift for Pedagogikk og Kritikk*, 2(2), 1–22. doi:10.17585/ntpk.v2.119
- Pacheco, A. (2009). Mapping the terrain of teacher quality. In D. H. Gitomer (Ed.), *Measurement issues and assessment for teaching quality* (pp. 160–178). Los Angeles, CA: SAGE Publishers.
- Paine, L., Blömeke, S., & Aydarova, O. (2016). Teachers and teaching in the context of globalization. In D. H. Gitomer & C. A. Bell (Eds.), *The handbook of research on teaching* (Vol. 5, pp. 717–786). Washington, DC: American Educational Research Association.
- Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, 62(3), 307–332. doi:10.3102/00346543062003307

- Paris, S. G., & Winograd, P. (1990). How metacognition can promote academic learning and instruction. In B. F. Jones & L. Idol (Eds.), *Dimensions of thinking and cognitive instruction* (pp. 15–44). Elmhurst, IL: Lawrence Erlbaum.
- Park, J. (2019). Validity issues in qualitative and quantitative research of cross-national studies. In L. Suter, E. Smith, & B. D. Denman (Eds.), *The SAGE handbook of comparative studies in education* (pp. 164–177). London, England: SAGE Publishers.
- Patton, M. Q. (2015). *Qualitative research & evaluation methods* (4th ed.). Los Angeles, CA: SAGE Publishers.
- Pehkonen, L. (2007). To change or not to change. *Nordic Studies in Mathematics Education*, 12(2), 57–76.
- Philipp, R. A. (2007). Mathematics teachers' beliefs and affect. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 257–315). Charlotte, NC: Information Age.
- Phillips, D. (2006). Comparative Education: Method. *Research in Comparative and International Education*, 1(4), 304–319. doi:10.2304/rcie.2006.1.4.304
- Phillips, D., & Schweisfurth, M. (2014). *Comparative and international education: An introduction to theory, method, and practice* (2nd ed.). London, England: Bloomsbury.
- Pianta, R. C., & Hamre, B. K. (2016). Implementing rigorous observations of teachers: Synchronizing theory with systems for implementation and support. In J. A. Grissom & P. Youngs (Eds.), *Teacher evaluation systems: Making the most of multiple measures* (pp. 22–36). New York, NY: Teachers College Press.
- Pons, A. (2018). What does teaching look like? A new video study. Retrieved from <https://oecdeditoday.com/what-does-teaching-look-like-a-new-video-study/>
- Praetorius, A.-K., & Charalambous, C. Y. (2018). Classroom observation frameworks for studying instructional quality: Looking back and looking forward. *ZDM*, 50(3), 535–553. doi:10.1007/s11858-018-0946-0
- Praetorius, A.-K., Pauli, C., Reusser, K., Rakoczy, K., & Klieme, E. (2014). One lesson is all you need? Stability of instructional quality across lessons. *Learning and Instruction*, 31, 2–12. doi:http://dx.doi.org/10.1016/j.learninstruc.2013.12.002
- Praetorius, A.-K., Rogh, W., Bell, C., & Klieme, E. (2019). Methodological challenges in conducting international research on teaching quality using standardized observations. In L. Suter, E. Smith, & B. D. Denman (Eds.), *The SAGE handbook of comparative studies in education* (pp. 269–288). London, England: SAGE Publishers.
- Raymond, A. M. (1997). Inconsistency between a beginning elementary school teacher's mathematics beliefs and teaching practice. *Journal for Research in Mathematics Education*, 28(5), 550. doi:10.2307/749691
- Reichenberg, O. (2018). Identifying mechanisms of teaching practices: A study in Swedish comprehensive schooling. *Scandinavian Journal of Educational Research*, 62(2), 213–228. doi:10.1080/00313831.2016.1212262
- Roth, K. J. (2013). Using video studies to compare and understand science teaching: Results from the TIMSS video study of 8th grade science teaching. In T. Janik & T. Seidel (Eds.), *The power of video studies in investigating teaching and learning in the classroom* (pp. 22–38). Münster, Germany: Waxmann.
- Rubin, H. J., & Rubin, I. S. (2012). *Qualitative interviewing: The art of hearing data* (Vol. 3). Los Angeles, CA: SAGE Publishers.
- Säljö, R. (2000). *Lärande i praktiken: Ett sociokulturellt perspektiv* [Learning in practice: A sociocultural perspective]. Stockholm, Sweden: Prisma.
- Schlesinger, L., & Jentsch, A. (2016). Theoretical and methodological challenges in measuring instructional quality in mathematics education using classroom observations. *ZDM*, 48(29–22), 29–40. doi:10.1007/s11858-016-0765-0

- Schoenfeld, A. (2016). Making sense of teaching. *Mathematics Education*, 48(1–2), 239–246. doi:10.1007/s11858-016-0762-3
- Schoenfeld, A., Floden, R., Chidiac, F., Gillingham, D., Fink, H., Hu, S., . . . Zarkh, A. (2018). On classroom observations. *Journal for STEM Education Research* 1, 34–59. doi:10.1007/s41979-018-0001-7
- Schoenfeld, A., & Teaching for robust understanding project. (2016). An introduction to the teaching for robust understanding (TRU) framework. Retrieved from [https://www.map.mathshell.org/trumath/intro\\_to\\_tru\\_20161223.pdf](https://www.map.mathshell.org/trumath/intro_to_tru_20161223.pdf)
- Schweisfurth, M. (2019). Qualitative comparative education research: Perennial issues, new approaches and good practice. In L. Suter, E. Smith, & B. D. Denman (Eds.), *The SAGE handbook of comparative studies in education* (pp. 258–268). Los Angeles, CA: SAGE Publications.
- Seidel, T., & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, 77(4), 454–499. doi:10.3102/0034654307310317
- Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational Researcher*, 27(2), 4–13. doi:10.2307/1176193
- Sfard, A. (2006). Participationist discourse on mathematics learning. In J. Maasz & W. Schloeglmann (Eds.), *New mathematics education research and practice* (pp. 153–170). Rotterdam, Netherlands: Sense Publishers.
- Sfard, A., & Kieran, C. (2001). Cognition as communication: Rethinking learning-by-talking through multi-faceted analysis of students’ mathematical interactions. *Mind, Culture, and Activity*, 8(1), 42–76. doi:10.1207/S15327884MCA0801\_04
- Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4–14. doi:10.3102/0013189X029007004
- Silverman, D. (1993). *Interpreting qualitative data: Methods for analysing talk, text and interaction*. London, England: SAGE Publishing.
- Simola, H. (2005). The Finnish miracle of PISA: Historical and sociological remarks on teaching and teacher education. *Comparative Education*, 41(4), 455–470. doi:10.1080/03050060500317810
- Simola, H., Kauko, J., Varjo, J., Kalalahti, M., & Sahlström, F. (2017). *Dynamics in education politic: Understanding and explaining the Finnish case*. London, England: Routledge.
- Skott, J. (2019). Understanding mathematics teaching and learning “in their full complexity.” *Journal of Mathematics Teacher Education*, 22(5), 427–431. doi:10.1007/s10857-019-09446-z
- Smith, E. (2011). Special issue on using secondary data in educational research. *International Journal of Research & Method in Education*, 34(3), 219–221. doi:10.1080/1743727X.2011.615976
- Snell, J. (2011). Interrogating video data: Systematic quantitative analysis versus micro-ethnographic analysis. *International Journal of Social Research Methodology*, 14(3), 253–258. doi:10.1080/13645579.2011.563624
- Solomon, Y., & Black, L. (2008). Talking to learn and learning to talk in the mathematics classroom. In N. Mercer & S. Hodgkinson (Eds.), *Exploring talk in schools: Inspired by the works of Douglas Barnes* (pp. 73–90). London, England: SAGE Publishing.
- Steffe, L. P., & Gale, J. E. (1995). *Constructivism in education*. Hillsdale, MI: Lawrence Erlbaum.
- Stein, M. K. (2007). Teaching and learning mathematics: How instruction can foster the knowing and understanding of number. In J. Brophy, (Ed.) *Subject-specific instructional methods and activities* (pp. 11–144). Bingley, England: Emerald.

- Stigler, J. W., Gallimore, R., & Hiebert, J. (2000). Using video surveys to compare classrooms and teaching across cultures: Examples and lessons from the TIMSS video studies. *Educational Psychologist*, 35(2), 87–100. doi:10.1207/S15326985EP3502\_3
- Stigler, J. W., Gonzales, P., Kawanaka, T., Knoll, S., & Serrano, A. (1999). *The TIMSS videotape classroom study: Methods and findings from an exploratory research project on eighth-grade mathematics instruction in Germany, Japan, and the United States*. Retrieved <https://nces.ed.gov/pubs99/1999074.pdf>
- Stigler, J. W., & Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom*. New York, NY: Free Press.
- Stigler, J. W., & Miller, K. F. (2018). Expertise and expert performance in teaching. In A. M. Williams, A. Kozbelt, K. A. Ericsson, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (2nd ed., pp. 431–452). Cambridge, England: Cambridge University Press.
- Stinson, D., & Bullock, E. C. (2012). Critical postmodern theory in mathematics education research: A praxis of uncertainty. *Educational Studies in Mathematics*, 80, 41–55. doi:10.1007/s10649-012-9386-x
- Stinson, D., & Walshaw, M. (2017). Exploring different theoretical frontiers for different (and uncertain) possibilities in mathematics education research. In J. Cai (Ed.), *Compendium for research in mathematics education* (pp. 128–155). Reston, VA: National Council of Teachers of Mathematics.
- Stovner, R. B., & Klette, K. (in review). Teacher feedback on procedural skills, conceptual understanding, and mathematical practices: A video study in lower secondary mathematics classrooms. Manuscript submitted for publication.
- Sztajn, P. (2003). Adapting reform ideas in different mathematics classrooms: Beliefs beyond mathematics. *Journal of Mathematics Teacher Education*, 6(1), 53–75. doi:10.1023/A:1022171531285
- Teddlie, C., Creemers, B., Kyriakides, L., Muijs, D., & Yu, F. (2006). The international system for teacher observation and feedback: Evolution of an international study of teacher effectiveness constructs. *Educational Research and Evaluation*, 12(6), 561–582. doi:10.1080/13803610600874067
- Tellings, A. (2001). Eclecticism and integration in educational theories: A metatheoretical analysis. *Educational Theory*, 51(3), 277–292. doi:10.1111/j.1741-5446.2001.00277.x
- Tellings, A. (2012). Theories and research in the field of education: An indissoluble union. In The Research Council of Norway, *The role of theory in educational research: Report from the March seminar 2011* (pp. 8–15). Retrieved from <https://www.forskningradet.no/siteassets/publikasjoner/1253979441594.pdf>
- Tytler, R., Chen, H.-L. S., Hackling, M. W., & Ramseger, J. (2019). A video study of quality teaching and learning in three countries. In L. Xu, G. Aranda, W. Widjaja, & D. Clarke (Eds.), *Video-based research in education: Cross-disciplinary perspectives* (pp. 34–49). New York, NY: Routledge.
- Vahtivuori-Hänninen, S., Halinen, I., Niemi, H., Lavonen, J., & Lipponen, L. (2014). A New Finnish National Core Curriculum for Basic Education (2014) and technology as an integrated tool for learning. In H. Niemi, J. Multisilta, L. Lipponen, & M. Vivitsou (Eds.), *Finnish innovations and technologies in schools: A guide towards new ecosystems of learning* (pp. 21–32). Rotterdam, Netherlands: Sense Publishers.
- Van de Grift, W. J. C. M. (2007). Quality of teaching in four European countries: A review of the literature and application of an assessment instrument. *Educational Research*, 49(2), 127–152. doi:10.1080/00131880701369651
- Van de Grift, W. J. C. M. (2014). Measuring teaching quality in several European countries. *School Effectiveness and School Improvement*, 25(3), 295–311. doi:10.1080/09243453.2013.794845



- Van de Vijver, F., & Leung, K. (1997). Methods and data analysis of comparative research. In J. W. Berry, Y. H. Poortinga, & J. Pandey (Eds.), *Handbook of cross-cultural psychology*, (2nd ed., pp. 257–300). Boston, MA: Allyn & Bacon.
- Van de Vijver, F., & Tanzer, N. K. (2004). Bias and equivalence in cross-cultural assessment: An overview. *Revue europeenne de psychologie appliquee*, 54(2), 119–135. doi:10.1016/j.erap.2003.12.004
- Van Oers, B. (2001). Educational forms of initiation in mathematical culture. *Educational Studies in Mathematics*, 46(1), 59–85. doi:10.1023/A:1014031507535
- Vettenranta, J., Välijärvi, J., Ahonen, A., Hautamäki, J., Hiltunen, J., Leino, K., . . . & Vainikainen, M.-P. (2016). *PISA 15 Ensituloksia. Huipulla pudotuksesta huolimatta* [PISA 15 First results. On top despite decline]. Opetus- ja kulttuuriministeriön julkaisu 2016:41. Retrieved from <http://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/79052/okm41.pdf>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Walshaw, M., & Anthony, G. (2008). The teacher's role in classroom discourse: A review of recent research into mathematics classrooms. *Review of Educational Research*, 78(3), 516–551. doi:10.3102/0034654308320292
- Wells, G. (1999). *Dialogic inquiry: Towards a socio-cultural practice and theory of education*. New York, NY: Cambridge University Press.
- Winne, P. H. (1987). Why process-product research cannot explain process-product findings and a proposed remedy: The cognitive mediational paradigm. *Teaching and Teacher Education*, 3(4), 333–356. doi:10.1016/0742-051X(87)90025-4
- Xu, L., & Clarke, D. (2019). Validity and comparability in cross-cultural video studies of classrooms. In L. Xu, G. Aranda, W. Widjaja, & D. Clarke (Eds.), *Video-based research in education: Cross-disciplinary perspectives* (pp. 19–33). New York, NY: Routledge.
- Yackel, E., Gravemeijer, K., & Sfard, A. (2011). *A journey in mathematics education research: Insights from the work of Paul Cobb*. Dordrecht, Netherlands: Springer International Publishing.
- Yin, R. K. (2012). *Applications of case study research* (3rd ed.). Los Angeles, CA, SAGE Publishing.

# Appendices

## Appendix 1: PLATO rubrics

Rubrics for the elements used in Articles I and III: Uptake of student responses, Opportunity for student talk, Quality of instructional explanations, Richness of instructional explanations, Connections to prior academic knowledge, Purpose, Intellectual challenge and Feedback. Based on Grossman (2015), while some clarifications have been adapted for mathematics instruction by the QUINT mathematics network.<sup>15</sup>

### Uptake of student responses

1 Provides almost no evidence	2 Provides limited evidence	3 Provides evidence with some weaknesses	4 Provides consistently strong evidence
<p>Teacher or students rarely if ever respond to students' ideas about mathematical content.</p> <p>Automatic teacher responses that simply acknowledge or echo student contributions (e.g., repetition of "Okay," "Good job," "Thanks") would fall into this category. Teacher accepts answers without asking for clarification or elaboration.</p>	<p>Teacher or students respond briefly to student ideas, and responses do not elaborate or help develop the ideas (e.g., restating without academic language, simple "I agree/disagree" statements that do not specifically reference a previous comment). Alternatively, the teacher may mostly respond to student ideas with automatic responses interspersed with an isolated instance of high-level uptake (e.g., revoicing in academic language: asking for clarification, elaboration, or evidence).</p>	<p>Teacher or student contributions show a balance between brief responses and higher-level uptake (e.g., revoicing in academic language: asking for clarification, elaboration, or evidence). There are multiple instances in which the teacher or students specifically address student ideas.</p>	<p>Teacher and students consistently engage in high-level uptake of students' ideas, responding in ways that expand on student ideas or enable students to further explain, clarify, and specify their thinking.</p>

### Opportunity for student talk

1 Provides almost no evidence	2 Provides limited evidence	3 Provides evidence with some weaknesses	4 Provides consistently strong evidence
<p>There are few or no opportunities for mathematics-related student talk. Teacher lecture, extended introduction (including giving directions) to an assignment or activity, or recitation format lasting less than 5 minutes would fall in this category.</p>	<p>Talk is tightly teacher-directed, but there are occasional opportunities for brief mathematics-related student talk. Examples include recitation formats lasting 5 minutes or longer, or mathematics-related talk (whole group, small group, partner talk) lasting less than 5 minutes.</p>	<p>Teacher provides opportunities for at least 5 minutes of mathematics-related conversation between teacher and students, and/or among students. Some students participate actively by speaking and/or listening, but only 2–3 students are the primary participants. There may still be a substantial amount of teacher direction, and some of the questions that guide the conversation are open-ended. Student-directed discussions that fail to stay on track would also be at this level.</p>	<p>Teacher provides opportunities for at least 5 minutes of mathematics-related conversation between teacher and students and/or among students. The majority of the students participate by speaking/actively listening, and students are responding to each other, even if the teacher is still mediating the conversation. The questions that guide the conversation are mostly open-ended, and the focus of the conversation is clear and stays on track.</p>

<sup>15</sup> For more information, see <https://www.uv.uio.no/quint/english/projects/lisa-nordic/>.

## Quality of instructional explanations

1 Provides almost no evidence	2 Provides limited evidence	3 Provides evidence with some weaknesses	4 Provides consistently strong evidence
Teacher provides no, weak, or incorrect explanations of mathematical concepts that may include incorrect analogies, examples, or explanations.	Teacher provides incomplete or perfunctory examples, analogies, or explanations that only touch surface-level features of mathematical content. The explanations are only partially successful in illuminating a concept.	Teacher provides accurate and clear examples, analogies, or explanations to sufficiently explain mathematical concepts. While the teacher may address misunderstandings, the teacher does not highlight the nuances of concepts, or provide counterexamples to help students distinguish among different features of related ideas.	Teacher provides examples, analogies, or explanations that are accurate and clear. In addition, the teacher addresses student misunderstandings, highlights the nuances of concepts (perhaps through the use of multiple slightly different examples or models), or provides counterexamples to help students distinguish among different features of related ideas.

## Richness of instructional explanations

1 Provides almost no evidence	2 Provides limited evidence	3 Provides evidence with some weaknesses	4 Provides consistently strong evidence
Teacher provides no, weak, or incorrect explanations of mathematical concepts.	The teacher provides superficial representation of mathematical content, focusing on rules, procedures and labels, with little attention to conceptual or deeper understanding.	The teacher's representation of content includes a balance of a focus on rules, procedures, and labels, as well as attention to conceptual or deeper understanding.	The majority of the teacher's instruction focuses on conceptual understanding of mathematical content. The teacher provides instruction that goes beyond the superficial to a focus on interpretation or deeper understanding of the concepts.

## Connections to prior knowledge

1 Provides almost no evidence	2 Provides limited evidence	3 Provides evidence with some weaknesses	4 Provides consistently strong evidence
Teachers or students do not refer to prior lessons nor elicit students' prior/background academic knowledge on a topic.	Teacher or students may refer briefly or superficially to prior lessons and/or attempt to elicit students' prior/background academic knowledge.	Teacher elicits or refers to students' prior/background academic knowledge multiple times on a topic. Connections made between prior knowledge and the day's lesson are clear enough to enable understanding of the material.	Teacher elicits or refers to students' prior/background academic knowledge multiple times on a topic (one or several really clear examples). Connections made between prior knowledge and new mathematical concepts or tasks are clear, explicit, and specifically tied to new material.

## Purpose

1 Provides almost no evidence	2 Provides limited evidence	3 Provides evidence with some weaknesses	4 Provides consistently strong evidence
There is no clear learning goal in the class or the learning goal is not related to the development of mathematical skills or understanding.	There is a learning goal communicated or inferred that is connected to the development of mathematical skills. The goal takes the form of a general topic or activity.	There is a clearly communicated, specific, learning goal that is connected to the development of mathematics skills. The lesson activities align to and target the specific learning goal.	There is a clearly communicated, specific, learning goal that is connected to the development of mathematics skills. The lesson activities align to and target the specific learning goal. There is evidence that students are aware of the purpose. The teacher or students refer back to the purpose during the segment. The teacher makes clear how the lesson will support students' development as learners of mathematics.

## Intellectual challenge

1 Provides almost no evidence	2 Provides limited evidence	3 Provides evidence with some weaknesses	4 Provides consistently strong evidence
Teacher provides activities or assignments that are almost entirely rote or recall.	Teacher provides activities or assignments that are largely rote or recall, but a portion of the segment promotes analysis, interpretation, inferencing, or idea generation.	Teacher provides a mix of activities or assignments: most promote analysis, interpretation, inferencing, or idea generation, and a few are focused on recall or rote tasks.	Teacher provides rigorous activities or assignments that largely promote sophisticated or high-level analytic and inferential thinking, including synthesizing and evaluating information and/or justifying or defending their answers or positions.

## Feedback

1 Provides almost no evidence	2 Provides limited evidence	3 Provides evidence with some weaknesses	4 Provides consistently strong evidence
The teacher does not provide feedback to students.  Feedback is confusing or misleading.	Teacher and/or students provide feedback that is vague, repetitive, perfunctory, or misleading (e.g., "Good job," "Correct," "No"). Suggestions for how to improve student performance are procedural rather than substantive. Teacher questions that imply next steps or suggestions for improvement fall at this level (e.g., "Now you must divide by 100." or "Have you asked your neighbor what they think?").	Teacher and/or students provide some feedback specific to features of students' work or ideas. Feedback is constructive and clear. Suggestions for how to improve work are a mix of the procedural and substantive.	Teacher and/or students frequently and consistently provide specific feedback. Suggestions for how to improve work are largely substantive. It is reasonable to infer that feedback helps students with the activity.

## Appendix 2: Overview of sample in Articles I and III

Region (OSL, HEL)	Teacher (F/M) Math ed.** Years of Experience (YoE)	School location and SES***	Class size*	No. and length of lessons recorded	Mathematical content
Classroom 1 (OSL)	F 1–30 ECTS 2 YoE	Urban Low SES	20	2 × 40 min 1 × 90 min	Geometry; constructing angles and geometric figures
Classroom 2 (OSL)	M 31–60 ECTS 2 YoE	Urban Low SES	17	3 × 60 min	Numbers; even, odd, and prime numbers; repetition of the four arithmetic operations
Classroom 3 (OSL)	F 60–90 ECTS 6 YoE	Urban Low SES	22	4 × 60 min	Geometry; intro to constructing angles
Classroom 4 (OSL)	M Master's degree 25 YoE	Suburb Mixed SES	16	4 × 50 min	Numbers; repetition of base-10 system
Classroom 5 (OSL)	M 60–90 ECTS 3 YoE	Suburb Mixed SES	13	4 × 45 min	Geometry; introduction to types of angles and drawing and constructing angles
Classroom 6 (OSL)	F 60–90 ECTS 4 YoE	Urban High SES	18	2 × 55 min	Algebra; calculating and constructing algebraic expressions, multiplication, and division
Classroom 7 (OSL)	F 60–90 ECTS 17 YoE	Urban High SES	22	2 × 45 min 1 × 90 min	Numbers; order of arithmetic operations; introduction to exponentials
Classroom 8 (OSL)	M (data missing on ECTS and YoE)	Urban High SES	25	3 × 40 min	Algebra; negative numbers and order of arithmetic operations
Classroom 9 (HEL)	F Master's degree 25 YoE	Urban Low SES	14	3 × 40 min	Numbers; areal units, rounding off, exact values, approximations
Classroom 10 (HEL)	F Master's degree 30 YoE	Urban Low SES	16	3 × 45 min	Numbers; areal units, rounding off, exact values, approximations
Classroom 11 (HEL)	F Master's degree 27 YoE	Urban Low SES	11	2 × 65 min	Algebra; variables with the four arithmetic operations
Classroom 12 (HEL)	M Master's degree 12 YoE	Suburb Mixed SES	16	3 × 45 min	Algebra; multiplication and division
Classroom 13 (HEL)	F Master's degree 2 YoE	Suburb Mixed SES	19	2 × 40 min	Numbers; test review; number sequences, patterns
Classroom 14 (HEL)	F Master's degree 7 YoE	Urban High SES	18	2 × 65 min	Geometry; area and circumference
Classroom 15 (HEL)	F 60–90 ECTS 2 YoE	Urban High SES	17	3 × 70 min	Geometry; area and perimeter of different polygons
Classroom 16 (HEL)	F 60–90 ECTS 25 YoE	Urban High SES	15	3 × 45 min	Algebra; multiplication and division of equations

\* Number of students on average taking part in the filmed lesson.

\*\* ECTS refers to the European Credit Transfer and Accumulation System; see

[https://ec.europa.eu/education/resources/european-credit-transfer-accumulation-system\\_en](https://ec.europa.eu/education/resources/european-credit-transfer-accumulation-system_en))

\*\*\* This refers to the socioeconomic status of the general area in which the school is located.

**Appendix 3: Analytical framework of discourse moves**  
*Teaching Moves* (Furtak & Shavelson, 2009, pp. 183–184)

<b>Dialogic Teaching Moves – Teacher and students jointly construct narrative/discussion</b>	
Asking “real” or open questions	Teacher asks a question of a student or entire class to which the answer is not necessarily known or expected by the teacher.
Spontaneous contributions	Students provide unsolicited comments not directly elicited by teacher.
Revoicing/reflecting on student responses	Teacher repeats verbatim what a student has responded without changing or altering the meaning of the statement. This includes when a teacher repeats in a question-style format or asks student to clarify what he or she said or to direct that comment to another student.
Meaning into matter	Teacher uses materials to illustrate or respond to a point or idea raised by student or teacher
Promoting disagreement/leaving lack of consensus	Teacher asks students to share divergent ideas and air differences or encourages them to disagree or not reach consensus.
Providing neutral responses to students	Teacher repeats student responses or provides comments that do not indicate whether student statements are correct or incorrect.
Teacher prompts students to explain to peers*	Teacher prompts students to explain their mathematical ideas, strategies, procedures, or concepts to peers.
Teacher encourages students to talk mathematics together*	Teacher encourages peer talk about mathematical content.
<b>Authoritative Teaching Moves – Teacher controls course of narrative/discussion</b>	
Cued elicitation of students’ contributions	Teacher asks questions while simultaneously providing heavy clues—such as the wording of a question, intonation, pauses, gestures, or demonstrations—to the information required.
Sequence of repeated questions	Teacher asks the same/similar questions repeatedly to seek a particular answer and continues asking the question(s) until answer is provided by students.
Selecting and/or ignoring students’ contributions	Teacher ignores a student’s contribution or selects a particular contribution from a chorus of different ideas stated by students.
Reconstructive paraphrase or recap	Teacher recasts or paraphrases what student has said in a more complete or acceptable form or in preferred terminology, including when the teacher adds to or changes the meaning of what the student has said.
Narrative	Teacher lectures or reviews storyline of unit, lesson, or activity or speaks in an uninterrupted flow to students.
Formulaic phrases	Teacher uses a particular phrase that is easy for students to remember and repeats it over and over again.
Marking significance	Speaking slowly or changing tone so students know that what is being said or what has been said is important.
Promoting/establishing consensus	Teacher encourages students to agree or come to a consensus.
Providing evaluative responses	Teacher clearly indicates, through words or intonation, that a student’s comment is correct or incorrect.

\*Added codes based on dialogic moves in group work situations

## **Appendix 4: Interview guide**

(Translated into English)

### **Introduction:**

How long have you worked as a mathematics teacher?

How did you feel about being filmed while teaching?

Do you think the camera affected your or your students' behavior?

In what way were the observed lessons typical or atypical of your lessons?

### **1. Own teaching practices**

What motivates your way of teaching?

How would you describe your own instructional practice in general?

How would you describe your instructional practices with this group in particular? (targeted questions on basic/advanced groups)

Do you find that the mathematical topic influences your choice of teaching methods?

How would you describe a good mathematics lesson?

What do you think is most rewarding and most challenging about being a mathematics teacher?

In what way, if any, would you say that the new curriculum has influenced your teaching practices?

### **2. How students learn mathematics**

As a mathematics teacher, what is most important for the students take away with them as they leave your lessons?

What is most important for them to learn?

(Follow-up) How do you support your students in learning this?

What instructional methods do you perceive as important for your students to learn mathematics?

How do you assess student learning?

How do you know when students make progress?

What were the learning objectives for the observed lessons?

### **3. Student participation**

What would you say is student participation in your classroom?

How important is it that your students participate?

When visiting your classroom, I observed that . . . (targeted questions)

### **4. Student participation in discourse**

How would you describe classroom discussions in your classroom?

What role do classroom discussions have in your classroom?

What role does whole-class instruction have?

In what ways do you encourage this group of students to participate in classroom discourse?

Does the way you engage students in discourse differ from other classes, and how?

In your classroom, do you have some established rules for how students engage in discussions (for example, who talks and to whom, how to address others' suggestions) and do students need to justify their answers?

When visiting your classroom, I observed that . . . (targeted questions)

### **5. Possibilities/constraints for student participation in CD**

What do you think influences student participation in classroom discourse in this class?

Are there any constraints in engaging your students in classroom discourse? If so, what are they?

How do you deal with students who do not want to participate orally?



## Appendix 5: Informed consent, Helsinki 2016



UNIVERSITETET I OSLO  
DET UTDANNINGSVITENSKAPELIGE FAKULTET

~~Till elever och vårdnadshavare~~

Institutt for lærerutdanning og

Postboks 1099 Blindern  
0317 Oslo

Telefon: +47 22855070

Vi är forskare från Institutt for Lærerutdanning og Skoleforskning (ILS) vid Universitetet i Oslo, som i samarbeide med forskare från Helsingfors universitetet har fått medel från Nordic Centre of Excellence: Justice through education in the Nordic countries (JustEd), Tema 2: "Justice through educational practices – Analysing innovative cultures of teaching and learning in Nordic contexts" för att jämföra undervisningspraktik i norska och finländska klassrum. Inom projektet görs ljud- och videoinspelningar av lärare och elever under lektioner i modersmål och litteratur samt matematik i årskurs 7. Under den sista lektionen ombeds eleverna därtill att fylla i en elevenkät om undervisningen.

Vi önskar spela in ca 180 minuter/3–4 lektioner modersmål och litteratur respektive matematik i årskurs 7 i Mattlidens skola i mars 2016. En kamera kommer att filma klassen som helhet och en kamera kommer att följa läraren.

Med detta brev ber vi om tillstånd för att göra ljud- och videoinspelningar av läraren och klassen som helhet. Vi ber därför om att **både** elever och vårdnadshavare skriver under på det bifogade avtalet.

Projektet kommer att genomföras i enlighet med gällande forskningsetiska riktlinjer. Studien har anmälts hos det norska Personvernombudet for forskning, (NSD) och allt material kommer att behandlas konfidentiellt. Deltagarna kommer inte att kännas igen i projektets publikationer.

Deltagande i projektet är frivilligt och den enskilda elevens val att delta eller inte delta kommer inte att ha konsekvenser i förhållande till undervisning eller skolan.

Med ~~vänliga hälsningar~~

Kirsti Klette (*sign.*)

Astrid Roe (*sign.*)

Anna Slotte (*sign.*)



## Deltagande i forskningsprojektet

---

### Elevens namn

- Ja, jag godkänner medverkan i JustEd-studien.
- Nej, jag önskar att inte delta i JustEd- studien.
- Undertecknade godkänner att utvalda klipp av videoinspelningarna kan användas i undervisnings- och forskningssammanhang vid Universitet i Oslo och Helsingfors universitet.
- Undertecknade godkänner att kopior av elevarbeten samlas in.

---

### Ort och datum

---

### Elevens underskrift

---

### Vårdnadshavares underskrift

**Allt deltagande i projektet är frivilligt och du kan när som helst dra dig ur studien, utan att ange någon orsak. Ifall du drar dig ur blir alla upplysningar anonymiserade.**

Har du några frågor om studien, vänligen ta kontakt med projektledare

Kirsti Klette ([kirsti.klette@ils.uio.no](mailto:kirsti.klette@ils.uio.no))

Astrid Roe ([astrid.roe@ils.uio.no](mailto:astrid.roe@ils.uio.no)) eller

Eller Anna Slotte ([anna.slotte@helsinki.fi](mailto:anna.slotte@helsinki.fi))



UNIVERSITETET I OSLO  
DET UTDANNINGSVITENSKAPELIGE FAKULTET

**Vänligen returnera svarsdelen till din lärare senast den .....**

### **Kort om projektets forskningsetiska riktlinjer**

Forskningsprojektet genomförs i enlighet med lagar för personskydd och de forskningsetiska riktlinjer som gäller i Norge och Finland. Alla personupplysningar kommer att behandlas konfidentiellt och deltagarna kommer inte att kunna kännas igen i projektets publikationer. Datamaterialet lagras i enlighet med nationella regler för säker lagring av detta slags data i Norge och Finland.

Datamaterialet kommer att ingå i forskningssamarbete mellan Universitet i Oslo och Helsingfors universitet. Doktorander och magisterstudenter vid de två universiteten kommer därtill att ha möjligheter att analysera delar av materialet.

Mer information om JustEd finns här:

[http://www.nordforsk.org/no/programmer/prosjekter/nordic-centre-of-excellence-justice-through-education-in-the-nordic-countries?set\\_language=no](http://www.nordforsk.org/no/programmer/prosjekter/nordic-centre-of-excellence-justice-through-education-in-the-nordic-countries?set_language=no)

Tack för hjälpen!

Kirsti Klette (sign.)

Astrid Roe (sign.)

Anna Slotte (sign.)

## Appendix 6: Informed consent, Helsinki 2018



UNIVERSITETET I OSLO  
DET UTDANNINGSVITENSKAPELIGE FAKULTET

Rektor vid XX skola

Institutt for lærerutdanning

Postboks 1099 Blindern  
0317 Oslo

Telefon: 22 85 50 70  
Telefaks: 22 85 44 09

### LISA – Linking Instruction and Student Achievement

Vi är en forskargrupp vid Institutt for Lærerutdanning og Skoleforskning (ILS) vid Universitetet i Oslo, som har fått medel från Norges forskningsråd att genomföra en studie var vi undersöker undervisningsmönster i nordiska klassrum. Våren 2016 var er skola med på detta forskningsprojekt genom videostudien LISA, då vi filmade lektioner i matematik och modersmål under ca en veckas tid. Tack vare er deltagelse har vi kommit vidare i projektet och kunnat göra analyser av undervisningsmönster på tvärs av klassrum.

Vi inom LISA önskar nu följa upp några klasser och lärare i Finland genom att igen filma lektioner i matematik under en veckas tid, gärna med samma lärare och elever (nu 9nde klassister) som för 2 år sedan. Denna gång önskar vi även intervjua lärare (och möjligtvis elever) om hur de upplever matematikundervisningen.

Vi ber härmed om tillåtelse att företa ljud- och videinspelning av lärare och klassen som helhet. Det skulle gälla en veckas lektioner i matematik i 9one klass vid \_\_\_\_\_ skola i slutet av januari/början av februari 2018 (efter avtal med lärare). En kamera kommer att filma klassen som helhet och en kamera följer läraren.

Projektet genomförs enligt gällande lagstiftning för integritets- och forskningsetiska riktlinjer. Studien har anmälts till Personvernombudet för forskning, Norsk samfunnsvitenskapelige datatjeneste (NSD), och all personlig information behandlas konfidentiellt. Deltagarna kommer inte att kunna bli igenkända i publikationer från projektet.



UNIVERSITETET I OSLO  
DET UTDANNINGSVITENSKAPELIGE FAKULTET

Deltagande i projektet är frivilligt och det kommer inte att få konsekvenser i förhållande till arbetsgivaren eller arbetsförhållandet om den enskilda läraren väljer att delta eller inte.

Med vänlig hälsning,

Kirsti Klette (*sign.*)  
Professor, ILS

Astrid Roe (*sign.*)  
Seniorforsker, ILS

Ole Kristian Bergem (*sign.*)  
Forsker, ILS

Jennifer Luoto  
PhD Student, ILS



UNIVERSITETET I OSLO  
DET UTDANNINGSVITENSKAPELIGE FAKULTET



Till elever och vårdnadshavare vid  
XXXX skola

Institutt for lærerutdanning og

Postboks 1099 Blindern  
0317 Oslo

Telefon: +47 22855070

Vi är forskare från Institutt for Lærerutdanning og Skoleforskning (ILS) ved   
Universitetet i Oslo, som i samarbeide med forskare från Helsingfors  
universitetet har fått medel från Nordic Centre of Excellence: Justice  
through education in the Nordic countries (JustEd), Tema 2: "Justice  
through educational practices – Analysing innovative cultures of teaching  
and learning in Nordic contexts" för att jämföra undervisningspraktik i  
norska och finländska klassrum. Inom projektet görs ljud- och  
videoinspelningar av lärare och elever under lektioner i modersmål och  
litteratur samt matematik i årskurs 9.

Vi önskar spela in ca 180 minuter/3 lektioner modersmål och litteratur  
respektive matematik i årskurs 9 i XXX skole i januari/februari 2018. En  
kamera kommer att filma klassen som helhet och en kamera kommer att  
följa läraren.

Med detta brev ber vi om tillstånd för att göra ljud- och videoinspelningar av  
läraren och klassen som helhet. Vi ber därför om att **både** elever och  
vårdnadshavare skriver under på det bifogade avtalet.

Projektet kommer att genomföras i enlighet med gällande forskningsetiska  
riktlinjer. Studien har anmälts hos det norska Personvernombudet for  
forskning, (NSD) och allt material kommer att behandlas konfidentiellt.  
Deltagarna kommer inte att kännas igen i projektets publikationer.

Deltagande i projektet är frivilligt och den enskilda elevens val att delta eller  
inte delta kommer inte att ha konsekvenser i förhållande till undervisning  
eller skolan.

Med vänliga hälsningar,

Kirsti Klette (*sign.*)

Jennifer Luoto(*sign.*)

Anna Slotte (*sign.*)



## Deltagande i forskningsprosjektet LISA

---

### Lärares navn

- Ja, jag deltar i LISA- studien.
- Nej, jag deltar inte i LISA- studien.
- Underteckand godtar att videodata används till undervisning/kursing av lärare/ lärarstudenter vid UiO

---

Dato

Plats

---

Lärares underskrift

**All deltagan i projektet är frivilligt och du kan når som helst återkalla ditt samtykke utan oppgven orsak. Om du drar tillbaka kommer all information att bli anonymiserad.**

Har du frågor om studien, vänligast ta kontakt med projektledar Professor Kirsti Klette ([kirsti.klette@ils.uio.no](mailto:kirsti.klette@ils.uio.no)), telefon 22 85 52 85/ 93 44 66 08)

Har du frågor om praktiska ting angående filmning i din klass, ta kontakt med doktorand Jennifer Luoto ([j.m.luoto@ils.uio.no](mailto:j.m.luoto@ils.uio.no)), telefon (0047)41001368

**Vennligst returner svarslippen til rektor innen** 



## UNIVERSITETET I OSLO

DET UTDANNINGSVITENSKAPELIGE FAKULTET

### Deltakelse i forskningsprosjektet LISA

---

#### Elevens navn

- Ja, jeg godtar å være med i LISA- studien.
- Nei, jeg vil ikke delta i LISA- studien.
- Undertegnede godtar at videodataene brukes til undervisning/kursing av lærere/ lærerstudenter ved UiO
- Undertegnede godtar at kopi av elevarbeider samles inn

---

Dato

Sted

---

Elevenes underskrift

---

Foresattes underskrift

**All deltakelse i prosjektet er frivillig, og du kan når som helst trekke ditt samtykke uten noen grunn. Dersom du trekker deg, vil alle opplysninger bli anonymisert.**

Dersom du har noen spørsmål til studien, vennligst ta kontakt med prosjektleder Professor Kirsti Klette ([kirsti.klette@ils.uio.no](mailto:kirsti.klette@ils.uio.no)), telefon 22 85 52 85/ 93 44 66 08)

**Vennligst returner svarslippen til læreren innen.....**

2



## Appendix 7: Informed consent, Oslo 2014–2015



UNIVERSITETET I OSLO  
DET UTDANNINGSVITENSKAPELIGE FAKULTET

**Rektor ved XX skole**

**Institutt for lærerutdanning**

Postboks 1099 Blindern  
0317 Oslo

Telefon: 22 85 50 70  
Telefaks: 22 85 44 09

### **LISA – Linking Instruction and Student Achievement**

Vi er en forskergruppe ved Institutt for Lærerutdanning og Skoleforskning (ILS) ved Universitetet i Oslo, som har fått midler fra Norges forskningsråd til å gjennomføre en studie hvor vi undersøker sammenhenger mellom undervisning og elevprestasjoner i regning og lesing. Prosjektet vil innhente opplysninger fra nasjonale prøver, elevsynspunkter på undervisning (spørreskjema), samt lyd- og videoopptak av lærere og elever i norsk og matematikk på ungdomstrinnet.

I forbindelse med dette ønsker vi å gjøre videoopptak av fire timer i norsk og matematikk i 8. klasse ved Sofienberg skole høsten 2014. Det vil være ett kamera som filmer klassen som helhet og ett kamera som følger læreren.

Vi ber med dette om tillatelse til å foreta lyd- og videoopptak av lærer og klassen som helhet.

Prosjektet vil bli gjennomført i henhold til gjeldende lovverk for personvern og forskningsetiske retningslinjer. Studien er meldt til Personvernombudet for forskning, Norsk samfunnsvitenskapelige datatjeneste (NSD), og alle personopplysninger vil bli behandlet konfidensielt. Deltakerne vil ikke kunne bli gjenkjent i publikasjoner fra prosjektet.

Deltagelse i prosjektet er frivillig og det vil ikke ha konsekvenser i forhold til arbeidsgiver eller arbeidsforhold om den enkelte lærer velger å delta eller ikke.

Med vennlig hilsen

Kirsti Klette (*sign.*)  
Professor, ILS

Astrid Roe (*sign.*)  
Seniorforsker, ILS

Ole Kristian Bergem (*sign.*)  
Forsker, ILS



**Eksempel på samtykkeerklæringer lærere:**

**Deltakelse i forskningsprosjektet LISA**

---

**Lærers navn**

- Ja, jeg godtar å være med i LISA- studien.
- Nei, jeg vil ikke delta i LISA- studien.
- Undertegnede godtar at videodataene brukes til undervisning/kursing av lærere/ lærerstudenter ved UiO

---

Dato

Sted

---

Lærers underskrift

**All deltakelse i prosjektet er frivillig og du kan når som helst trekke ditt samtykke uten noen grunn. Dersom du trekker deg, vil alle opplysninger bli anonymisert.**

Dersom du har noen spørsmål til studien, vennligst ta kontakt med prosjektleder Professor Kirsti Klette ([kirsti.klette@ils.uio.no](mailto:kirsti.klette@ils.uio.no)), telefon 22 85 52 85/ 93 44 66 08)

**Vennligst returner svarslippen til rektor innen ~~.....~~**



### Kort om LISA -prosjektet

LISA -prosjektet ("Linking Instruction and Student Achievement") har fått midler av Norges Forskningsråd til å gjennomføre en større videostudie hvor vi undersøker sammenhenger mellom undervisning og elevprestasjoner i regning og lesing på ungdomstrinnet. Prosjektet vil innhente opplysninger fra nasjonale prøver på 8. og 9.trinn, elevsynspunkter på undervisning (spørreskjema), samt lyd- og videoopptak av lærere og elever i norsk og matematikk på ungdomstrinnet.

Prosjektet vil bli gjennomført i henhold til gjeldende lovverk for personvern og forskningsetiske retningslinjer og er meldt til Personvernombudet for forskning, Norsk samfunnsvitenskapelige datatjeneste (NSD). Alle personopplysninger vil bli behandlet konfidensielt og deltakerne vil ikke kunne bli gjenkjent i publikasjoner fra prosjektet. Prosjektet har finansiering fram til 2017 men vi regner å analysere disse dataene over en 10-årsperiode.

Datamaterialet vil bli lagret i sikker forskningsserver ved UiO/USIT i 15 år, og nøkkelen som kobler videofilmer med de faktiske skoler/ klasserommene vi har filmet (koblingsnøkkel), vil bli oppbevart hos NSD. Det kan være aktuelt å komme tilbake til enkelte klasser for en mulig oppfølgingsstudie.

Utvalgte deler av datamaterialet vil kunne inngå i forskningssamarbeid med henholdsvis Stanford, Helsinki og Harvard universitet. [Dr.grads stipendiater](#) og Master studenter ved Universitetet i Oslo/Institutt for Lærerutdanning og Skoleforskning (ILS) vil videre kunne få muligheter til å analysere deler av materialet.

Deltagelse i prosjektet er frivillig og det vil ikke ha konsekvenser i forhold til arbeidsgiver eller arbeidsforhold om den enkelte lærer velger å delta eller ikke.

Takk for hjelpen!

Kirsti Klette (sign.)

Astrid Roe (sign.)

Ole Kristian Bergem (sign.)



### Deltagande i forskningsprojektet

|

---

#### Elevers namn

- Ja, jag godkänner medverkan i JustEd-studien.
- Nej, jag önskar att inte delta i JustEd- studien.
- Undertecknade godkänner att utvalda klipp av videoinspelningarna kan användas i undervisnings- och forskningssammanhang vid Universitet i Oslo och Helsingfors universitet.
- Undertecknade godkänner att kopior av elevarbeten samlas in.

---

Ort och datum

---

Elevers underskrift

---

Vårdnadshavares underskrift

**Allt deltagande i projektet är frivilligt och du kan när som helst dra dig ur studien, utan att ange någon orsak. Ifall du drar dig ur blir alla upplysningar anonymiserade.**

Har du några frågor om studien, vänligen ta kontakt med projektledare

Kirsti Klette ([kirsti.klette@ils.uio.no](mailto:kirsti.klette@ils.uio.no)) eller

Jennifer Luoto ([j.m.luoto@ils.uio.no](mailto:j.m.luoto@ils.uio.no))

Eller Anna Slotte ([anna.slotte@helsinki.fi](mailto:anna.slotte@helsinki.fi))

**Vänligen returnera svarsdelen till din lärare senast den .....**

## **PART II: The Articles**









# Scrutinizing two Finnish teachers' instructional rationales and perceived tensions in enacting student participation in mathematical discourse

Jennifer Luoto

Department of Teacher Education and School Research, Oslo University, Norway

This study employs interviews and observations to investigate instructional rationales of two purposefully sampled teachers with divergent classroom discourse practices in Swedish-speaking Finnish lower secondary mathematics classrooms. Studies on classroom discourse often point to beliefs and contextual factors shaping teachers' discourse practices. Less is known about how tensions perceived by teachers can influence the instructional rationale in a context such as Finland, known for traditional and teacher-centered mathematics instruction. The findings of this study suggest that these Finnish teachers' instructional rationales for differently enacted classroom-discourse practices are grounded in similar concerns of student needs, related to student learning, well-being, and equity. One of the teachers perceived tension between these concerns and mathematics education literature's ideals of classroom discourse and avoided engaging students in discussions other than in tightly teacher-led format. The other embraced the idea of discourse as facilitating learning and created methods for giving all students equal access to the perceived benefits of mathematical discussions. The identified tensions of student learning, well-being, and equity can be used as guiding principles in developing teachers' discourse practices in professional development in Finland and beyond.

Keywords: classroom discourse, student participation, teachers' instructional rationale

## Article Details

LUMAT General Issue  
Vol 8 No 1 (2020), 133–161

Received 14 April 2020  
Accepted 11 June 2020  
Published 15 June 2020

Pages: 29  
References: 75

Correspondence:  
[j.m.luoto@ils.uio.no](mailto:j.m.luoto@ils.uio.no)

[https://doi.org/10.31129/  
LUMAT.8.1.1329](https://doi.org/10.31129/LUMAT.8.1.1329)

## 1 Introduction

Student verbal participation in classroom discourse e.g., talking mathematics by sharing thoughts and justifying reasoning, is widely recognized as mediating mathematics thinking and learning (Lampert & Blunk, 1998; Kieran, Forman, & Sfard, 2001; Franke, Kazemi & Battey, 2007; Organisation for Economic Co-operation and Development [OECD], 2016a) and positively affecting motivation (Kiemer, Gröschner, Pehmer, & Seidel, 2015). These ideas of learning mathematics through participating in mathematics discourse are often referred to as sociocultural and Western ideas (e.g., Xu & Clarke, 2019). They were emphasized in American (e.g., National Council of Teachers of Mathematics, 1989) and some European curricular contexts (see Gravemeijer, Bruin-Muurling, Kraemer, & Van Stiphout, 2016) as part of a “paradigm shift” away from traditional, teacher-centered approaches and toward



“reform-oriented” instruction focusing on student engagement and inquiry-based learning (Ellis & Berry, 2005). This shift has been less prominent in the Finnish context, where instructional practices at the lower secondary level are characterized by teacher-centered instruction and individual seatwork, with scarce opportunities for students to participate in mathematical discussions (e.g., Klette et al., 2018; Taajamo, Puhakka, & Välijärvi, 2014). In addition, mathematical argumentation has not been a part of the traditional Finnish school mathematics education (Kaasila, Pehkonen, & Hellinen, 2010), and teachers are viewed as well-established authorities on content knowledge (Pehkonen, Ahtee, & Lavonen, 2007). Thus, perhaps not surprisingly, participation in mathematics discourse has traditionally not been emphasized in national curricula. However, the latest national curriculum (Finnish National Agency for Education, 2014, pp. 438-441) promotes mathematics instruction that develops students’ ability to communicate, interact, and cooperate through presenting and discussing solutions and working in groups as well as individually. Furthermore, the previously high PISA scores—which, in a way, have protected the status quo of traditional instructional practices (see Simola et al., 2017)—are now in decline, while Finnish mathematics educators report a decrease in interest and skills in mathematics in lower secondary schools (Portaankorva-Koivisto, Eronen, Kupiainen, & Hannula, 2018). It is thus timely to study teachers’ instructional rationales and potential tensions that might prevent teachers from prompting discourse among students in a Finnish context. This is important insight for teacher education, as targeting potential tensions that might constrain teachers from discursive practices is needed to develop instruction in line with the curriculum, which also may elevate students’ motivation for mathematics (Kiemer et al., 2015). The goal of the present study is therefore to investigate two Finnish teachers’ instructional rationale for their differently enacted classroom discourse practices and identify perceived tensions related to enabling discourse among students in lower secondary mathematics classrooms.

## 2 Classroom discourse

Discourse practices in mathematics classrooms are considered contextually bound and collectively developed patterned ways of communicating (e.g., O’Connor, 1998; Xu & Clarke, 2013). Yet, classroom interaction research has been able to categorize some generic teacher moves shaping student participation in classroom discourse (e.g., Alexander, 2006; Cazden, 1988; Solomon & Black, 2008). This study uses the

categorization of authoritative and dialogic teacher moves by Furtak and Shavelson (2009), building on Mortimer and Scott (2003), to distinguish between teacher moves in which students engage in co-construction of discussions and moves in which the teacher constructs the discussion.

## 2.1 Authoritative teacher moves

Authoritative teacher moves imply information transmissions from teacher to students and are the most common moves in mathematics classrooms (Alexander, 2006). A common pattern associated with authoritative teacher moves is questioning in the pattern called Initiation-Response-Evaluation/Feedback (IRE/IRF) (Cazden, 1988), where the teacher calls for single responses from students, interspersed within longer sections of teacher talk, and student answers often receive short evaluative responses. Other authoritative teacher moves are *repeating formulaic phrases* and *marking significance* to help students remember information (Furtak & Shavelson, 2009), and *instruction/exposition*, in which the teacher controls the narrative of information, activities, facts, principles, and procedure (Alexander, 2006). In addition, *repeated questions* and *cued elicitation of student contributions* are considered authoritative teacher questions, as they lead students to the right answer, also known as a “funneling pattern” (Wood, 1998). A final example of an authoritative move is if teachers *promote consensus* and *select particular student contributions* as being correct (Furtak & Shavelson, 2009), thus puncturing discussions of misconceptions or alternative solutions. All these listed moves are authoritative (Mortimer and Scott, 2003), as such moves facilitate teacher control over the discourse while not inviting students to contribute to shaping the discourse or knowledge construction.

## 2.2 Dialogic teacher moves

Dialogic teaching moves promote discussions and give students opportunities to participate in the construction of knowledge and discourse (Ball & Bass, 2000). Dialogic teacher moves thus enable what Fennema et al. (1996) call “productive mathematical discourse” that supports inquiry-based learning where students actively grapple with mathematical problems (Artigue & Blomhøj, 2013). Such teacher moves are *open and “real” questions*, in which the teacher does not necessarily know the answer, as well as providing *neutral responses to student ideas* (Furtak & Shavelson, 2009). Dialogic moves are further in line with a “focusing pattern” (Wood,

1998), in which teachers *prompt students to explain* their mathematical ideas. Explaining helps students grasp principles, construct rules for solving problems, and become aware of misunderstandings or lack of understanding as well as develop new understandings (Ingram, Andrews, & Pitt, 2019). Teachers may *re-voice or elaborate on student explanations* by using materials to further illustrate ideas or *ask for justifications* to probe student thinking and direct student contributions to become mathematical (Franke et al., 2009; Walshaw & Anthony, 2008). Taken together, the dialogic teacher moves thus invite students to shape the discussions and their understanding of content (see Mortimer & Scott, 2003).

### 2.3 Balancing teacher moves

The authoritative/dialogic dichotomy presented above is useful for describing discourse patterns within classrooms but less useful for judging discourse quality (Drageset, 2015). Both types of moves have their place in mathematics classrooms. Authoritative moves, such as IRE-patterned questions, may be effective in discussions of previously learned content (Temple & Doerr, 2012), while dialogic moves are beneficial for grappling with new mathematical concepts (Fennema et al., 1996). However, teachers socialize students into ways of thinking and reasoning about mathematics through discourse (O'Connor, 1998), and if teachers use only authoritative moves and never engage students in challenging discourse, students may miss opportunities to develop mathematical reasoning (Cobb & Bowers, 1999). Several scholars thus recommend that teachers balance authoritative and dialogic moves so that students can both explore ideas and learn relevant content (Boerst, Sleep, Ball, & Hyman, 2011; Scott, Mortimer, & Aguiar, 2006).

It is contested whether participation in discourse is equally important for all students. For example, studies show that students may learn just as much by vocal or silent participation in discourse (O'Connor, Michaels, Chapin, & Harbaugh, 2017), and that participation in discourse is not necessarily beneficial for students with learning disabilities (e.g., Gersten et al., 2009). It is also questioned what type of activity format is most beneficial for student participation in discourse. Traditional whole-class instruction is considered inequitable, as it engages only volunteering students (Emanuelsson & Sahlström, 2008). While in group work, some group partners are more engaged in discussions than others; hence not all students have the same opportunities to engage in content discussions (Bergem & Klette, 2010; Webb, Nemer, Chizhik, & Sugrue, 1998). To establish norms and expectations for social

behavior in the content-focused discourse, teachers need to pay attention to both social (eliciting contributions from different students) and analytical scaffolding (prompting students to explain reasoning) (Kovalainen & Kumpulainen, 2005). Consequently, just as teachers need to balance authoritative and dialogic moves, they also need a broad repertoire of techniques for orchestrating classroom discussions that function as productive learning situations for all students (Sfard, 2003; Bergem & Klette, 2016). Moreover, as the following review suggests, there are several different factors that may shape teachers' instructional decisions about classroom discourse practices.

### 3 Teachers' instructional rationale for enacted discourse practices

*Instructional rationale* in this study refers to how teachers rationalize their instruction in the complex and situated environment of mathematics classrooms (Confrey, 2017). Similarly to Jeppe Skott's (2001) concept of *school mathematics images*, instructional rationale is concerned with teachers' idiosyncratic and subjective accounts of their mathematics teaching. Instructional rationale is thus limited to teachers' explicit, avowed, and uttered views of their enacted practices (Fives & Gill, 2015), in contrast to *teacher beliefs*, which refer to psychologically held understandings, premises, or propositions about the world that are thought to be true (Pajares, 1992; Philipp, 2007). From the literature, we know that beliefs (e.g., Atweh, Cooper, and Bleicher, 1998; Brendefur & Frykholm, 2000; Reichenberg, 2018; Sztajn, 2003; Spillane, 2002; Skott, 2001; Pehkonen, 2007) as well as contextual factors (e.g., Ayalon & Even, 2016; Herbel-Eisenmann, Lubienski, and Id-Deen, 2006; Davis et al., 2019; Raymond, 1997) explicitly and implicitly shape classroom discourse practices. For example, Brendefur and Frykholm (2000) found that beliefs about *mathematics* and *the role of the teacher* influence the instructional rationales of teachers' enacted discourse practices in the classroom. The instructional rationale of a teacher with teacher-centered instruction was shaped by beliefs of mathematics as fixed and knowledge as transmissible—believing that learning occurred when students watched examples and listened to explanations. The instructional rationale of another teacher with reform-oriented practices, including group work, was shaped by beliefs that mathematics should be an active endeavor and that mathematics communication facilitated learning and students' construction of knowledge. In a study by Reichenberg (2018), a mathematics teacher rationalized about his preference for

individual seatwork over discussion-based activities. This teacher stressed that individual work was important for developing higher-order skills and logical thinking, which this teacher considered as non-verbal skills, while he perceived discursive practices in whole-class teaching as mainly promoting verbal skills and lower-order thinking.

Sztajn (2003) and Spillane (2002) in their respective studies demonstrate that teachers' instructional rationales may be related to beliefs about the *needs* of students with different socioeconomic status (SES); low-SES students are believed to need teacher-centered direct instruction of basic skills, while high-SES students need to be challenged intellectually—for example, through discourse. Similarly, Atweh et al. (1998) suggest that beliefs about other student needs—depending on gender, abilities, and their futures—shape the instructional rationale of teachers. A teacher who saw his male students as high achievers and future mathematicians stressed student independence and self-control of learning, while a teacher who perceived his female students as middle achievers with a future in tertiary studies preferred direct instruction (Athew et al., 1998). In a study by Skott (2001), the teacher enacted different discourse practices depending on beliefs about the *main concern* for particular students—when the concern was building student confidence, interactions with students were more direct than when the main priority was mathematical learning.

The instructional rationales for discourse practices may also be shaped by tensions and constraints related to *contextual factors*. In Raymond's study (1997), *a large group size, lack of time and resources, and standardized tests* were perceived as constraining a teacher from prompting students to engage in discussions. Similarly, Davis et al. (2019) show how a teacher who generally embraced reform-based teaching, perceived tension between reform-based teaching and *accountability systems*, such as curricula, resources, and expectations from parents and the school. Moreover, Herbel-Eisenmann et al. (2006) found that *parents' demands, curriculum materials, and students' own preferences* were factors a teacher perceived as constraining reform-oriented teaching approaches. Also, more *specific situational factors* influence classroom discourse. Ayalon and Even (2016) show that a specific mathematical topic, the specific teacher, and the characteristics of a specific class shaped students' opportunities for diverging into argumentative discussions, stressing that the mathematical topic and the students themselves shape classroom discussions.

In the Finnish context, empirical research from classrooms is scarce (Simola, 2017) and only a few studies shed light on teachers' instructional rationales of mathematics teachers' discourse practices. For example, in Pehkonen's (2007) interview study on Finnish mathematics teachers' beliefs, teachers implemented teacher-centered methods and the use of textbooks, viewing this as a safe method for delivering content. Kupari (2003), drawing on Trends in International Mathematics and Science Study (TIMSS) survey data, identified how two diverging groups of Finnish mathematics teachers' beliefs reflected their reported practices: the group holding constructivist beliefs embracing understanding as essential for learning were more likely to engage their students in group work than the teachers holding traditional transmissive beliefs. More research is scarcely needed to nuance how such different beliefs may be enacted in classroom practice and instructional decisions in a Finnish context.

In summary, the reviewed studies point to several different factors teachers may perceive as shaping students' participation in mathematics discussions. This study contributes to the field of mathematics education by identifying rationales and possible tensions two teachers with different discursive practices perceive in engaging students in discourse. Situated in a Finnish context, where classroom discourse is not traditionally a part of mathematics education (Kaasila et al., 2010), this study may also nuance the discussion about ideal practices in classroom discourse, as research from different national contexts can contribute to the field by "challenging the relevance of culturally specific evaluative concepts" (Hemmi & Ryve, 2015, p. 504; Skott, 2019). Knowledge of how teachers rationalize their different classroom discourse practices in a Finnish context may thus inform teacher training and professional development on issues that need to be addressed in order to develop teachers' repertoire of enacted discourse practices. The following overarching research question guided the analysis: *How do two Finnish mathematics teachers with diverging practices perceive and enact student participation in discourse?* In order to approach this question, three sub-questions were posed:

1. What instructional moves do the two mathematics teachers use to engage students in classroom discourse, and to what extent are these moves used?
2. What is the instructional rationale for the two mathematics teachers' instructional moves in classroom discourse?
3. What kind of possible tensions do teachers with different practices perceive as hindering or enabling student participation in discourse?

## 4 Methods

### 4.1 Participants

The participating teachers are Anna and Bea (pseudonyms), sampled from the LISA video study focusing on instructional practices in Nordic lower secondary classrooms (see Klette, Blikstad-Balas, & Roe, 2017). These teachers were purposefully sampled (Patton, 2015), since they displayed contrasting and illustrative patterns of different classroom discourse practices in another study involving eight Swedish-speaking Finnish mathematics classrooms (Luoto et al., *in rev*). Anna was sampled due to her atypical practice, in which she constantly engaged her students in discussions in various ways. Bea represents a more typical practice, providing few opportunities for students to discuss mathematics. Thus, they represent different types of classroom discourse practices. In this study, I focus on their ninth grade<sup>1</sup> classes in 2018, when the students are 15 years old. Both teach in schools located in urban, high-SES areas around Helsinki. Anna teaches an “advanced” class, and Bea teaches a “basic” class, but they follow the same curriculum. This kind of tracking was officially discontinued in compulsory education in Finland in the mid-1980s (Pekkarinen & Uusitalo, 2012, p. 132), as it was considered inequitable. However, the national curriculum allows temporary grouping as a means for differentiation (Finnish National Agency for Education, 2014), and over 50% of Finnish principals report some form of ability-based grouping for ninth graders (OECD, 2016b).

### 4.2 Video observations

Three consecutive mathematics lessons from each teacher were video recorded. Two cameras were strategically placed in each classroom, one facing the teacher and one the entire classroom. The teacher wore one microphone, while the other captured student talk. The author was present in the classroom during the filmed lessons, in the role of “observer as participant”— an outsider watching the lesson without intervening (Bernard, 2011). The field notes consisted of pictures of student work and descriptions of tasks and other instructional materials.

---

<sup>1</sup> The 9<sup>th</sup> grade is the final year of compulsory school in Finland.



### 4.3 Interviews

The interviews were semi-structured (Harding, 2013), with mostly open-ended questions on five preselected themes: teachers' perceptions of their own teaching, how students learn mathematics, student participation in general, student participation in discourse in their classroom, and what teachers saw as possibilities and constraints for student participation in discourse (Table 1). The themes in the interview guide were built on the reviewed previous research on beliefs and contextual factors shaping classroom discussions, to broadly include possible factors shaping teachers' instructional rationales. The interview guide was also refined after piloting the interview with two mathematics teachers, to clarify questions that were unclear.

**Table 1.** Overview of interview themes.

<b>Theme</b>	<b>1. Own teaching practices</b>	<b>2. How students learn mathematics</b>	<b>3. Student participation</b>	<b>4. Student participation in discourse</b>	<b>5. Possibilities/constraints for student participation in CD</b>
<b>Example question</b>	<i>How would you describe your own instruction?</i>	<i>What instructional methods do you perceive as important for your students to learn mathematics?</i>	<i>What is student participation in your classroom?</i>	<i>In what ways do you encourage this group of students to participate in classroom discourse?</i>	<i>Are there any constraints in engaging your students in classroom discourse? If so, what are they?</i>
<b>Purpose</b>	To gain an overview of how the teachers perceive d their instruction in the classroom	Investigate whether and how teachers shape their instructional practice with a specific view of learning mathematics (Brendefur and Frykholm, 2000; Kupari, 2003; Reichenberg, 2018)	Investigate how teachers perceive student participation in general	Investigate discursive practices the teachers perceived they enacted and why they enacted it for that particular class (Ayalon & Even, 2016; Atweh et al., 1998; Spillane et al, 2001)	Investigate possible constraints teachers perceive as hindering them from engaging students in discourse (Herbel-Eisenmann et al., 2006; Skott, 2001, Raymond, 1997; Davis et al., 2019)

The interviews were focused (Cohen, Manion, & Morrison, 2011), targeting the teachers' subjective responses to a situation (their instruction) in which they were involved. In line with focused interviews, some questions were tailored to the observed practice. For example, Anna was questioned about the rationale for her group-work practices, and Bea was questioned about the consistent use of teacher-led whole-class sessions. In general, the questions were posed in the same order to both teachers, while still allowing them to pursue topics important to them (Silverman, 1993). The interviews were audio-recorded, lasted approximately one hour, and took place immediately after the last observed lesson so that the teachers would remember the lessons, thus limiting recall bias. Both interviews were transcribed verbatim.

#### 4.4 Application and adaptations of the analytical framework

Furtak and Shavelson's (2009) framework of dialogic and authoritative teacher moves (Table 2), building on a body of previous research (Cazden, 1988; Lemke, 1990; Mortimer & Scott, 2003; Scott, 1998, and others), served as an analytical lens to facilitate a detailed presentation of teacher moves that enable or constrain student participation. It has previously been applied in other video studies in different subjects (see, for example, Andersson & Klette, 2016). The framework was applied on *classroom discourse episodes* (e.g., instances of mathematics discussion in whole class or among peers). This excludes individual teacher-student talk, which is not considered to constitute a joint discussion and understanding of mathematics (Mercer & Hodgkinson, 2008). Teacher utterances during discourse episodes were coded as *authoritative*, *dialogic*, *blended*, or *not applicable*. The blended code was applied when a teacher enacted both dialogic and authoritative moves within a single utterance, such as when Anna, in the below example, both controls the narrative by constructing the guidelines and purpose of the group activity (authoritative) and prompts students to discuss mathematics (dialogic).

“We will do this task together in groups so you can test what you remember and so I can check that you understand. Discuss within the group. I don't want the person who thinks he or she knows best to respond immediately. Check with each other that everybody knows.” (Anna)

Some teacher utterances did not fall into any category and were coded *not applicable*, such as non-content-related questions and comments. These utterances are not included in the results.

**Table 2.** Teaching moves (Furtak & Shavelson, 2009, pp. 183-184)

Dialogic Teaching Moves – Teacher and students jointly construct narrative/discussion	
Asking “real” or open questions.	Teacher asks a question of a student or entire class to which the answer is not necessarily known or expected by the teacher.
Spontaneous contributions.	Students provide unsolicited comments not directly elicited by teacher.
Revoicing/reflecting on student responses.	Teacher repeats verbatim what a student has responded without changing or altering the meaning of the statement. Includes when a teacher repeats in a question-style format or asks student to clarify what he or she said or to direct that comment to another student.
Meaning into matter.	Teacher uses materials to illustrate or respond to a point or idea raised by student or teacher.
Promoting disagreement / leaving lack of consensus.	Teacher asks students to share divergent ideas and air differences or encourages them to disagree or not reach consensus.
Providing neutral responses to students.	Teacher repeats student responses or provides comments that do not indicate whether student statements are correct or incorrect.
Teacher prompts students to explain to peers.	Teacher prompts students to explain their mathematical ideas, strategies, procedures, or concepts to peers.
Teacher encourages students to talk mathematics together.	Teacher encourages peer talk about mathematical content.
Authoritative Teaching Moves - Teacher controls course of narrative/discussion	
Cued elicitation of students’ contributions.	Teacher asks questions while simultaneously providing heavy clues — such as the wording of a question, intonation, pauses, gestures, or demonstrations—to the information required.
Sequence of repeated questions.	Teacher asks the same/similar questions repeatedly to seek a particular answer and continues asking the question(s) until answer is provided by students.
Selecting and/or ignoring students’ Contributions.	Teacher ignores a student’s contribution or selects a particular contribution from a chorus of different ideas stated by students.
Reconstructive paraphrase or recap.	Teacher recasts or paraphrases what student has said in a more complete or acceptable form or in preferred terminology, including when the teacher adds to or changes the meaning of what the student has said.
Narrative.	Teacher lectures or reviews storyline of unit, lesson, or activity or speaks in an uninterrupted flow to students
Formulaic phrases.	Teacher uses a particular phrase that is easy for students to remember and repeats it over and over again
Marking significance.	Speaking slowly or changing tone so students know that what is being said or what has been said is important
Promoting/establishing a consensus.	Teacher encourages students to agree or come to a consensus.
Providing evaluative responses.	Teacher clearly indicates, through words or intonation, that a student’s comment is correct or incorrect.

Two additional codes were developed to capture teacher moves specific to peer work: *Teacher prompts students to explain to peers* and *teacher encourages students to talk mathematics together* (added as dialogic codes in Table 2). While these can be interpreted as authoritative moves since the teacher controls the activity, they are labeled dialogic here as they prompt student explanations and joint construction of knowledge, which are key indicators of dialogic teaching (e.g., Alexander, 2006). In Figure 1, application of the framework is illustrated in a short excerpt from a lesson about triangles using the software GeoGebra<sup>2</sup>, in which Anna instructs a pair of students to “change two of the points of the triangle while maintaining the same area.”

SPEAKER	CODE	D/A
Student: That's easy		
Teacher: "Oh? How do you think then?"	<i>Asking real questions</i>	Dialogic
S: "You just change it, so you keep the shape but at different points."		
T: "You mean exactly the same shape? Then let's say you cannot have the same shape. Change two points and try to think if there is a systematic way of doing it, do you know anything that could help you about triangles?"	<i>Cued elicitation</i>	Authoritative
S: "The formula for the areal $b \cdot h / 2$ [clicks on the computer]. So, I can change it like this!"		
T: "Mm, now you can pretend you are a teacher and tell your partner what you figured out."	<i>Teacher prompts students to explain to peers</i>	Dialogic

Figure 1. Example of coding.

As illustrated above, teacher utterances were coded on the sentence level, and this example shows how dialogic and authoritative moves may be intertwined in teacher-student interactions.

<sup>2</sup> <https://www.geogebra.org/about>

## 4.5 Phases of analysis

The analysis was performed in four phases. In the first phase, drawing on video observations and field notes, all lessons were viewed several times, transcribed, and mined for identifiable discourse episodes. The focus in Anna's lessons was on triangles (e.g., constructing and calculating angles), and in Bea's, the focus was on exponent rules (e.g., how to simplify and multiply exponents). While the topic of the lessons may encourage different discourse practices, I study these lessons as examples representing different teaching approaches to discourse, and not as a comparison on these two particular teachers (see [Section 6.3](#)).

In the second phase, the teacher utterances in classroom discourse episodes were coded using the framework by Furtak and Shavelson (2009) ([Table 2](#)), and their frequency counted. These analyses answer the first research sub-question: *What instructional moves do the two mathematics teachers use to engage students in classroom discourse, and to what extent?*

In the third phase, the interviews were transcribed and analyzed in order to answer the second and third sub-questions: *What is the instructional rationale for the two mathematics teachers' instructional moves in classroom discourse? And What kind of possible tensions do teachers with different practices perceive as hindering or enabling student participation in discourse?* Two themes were extracted in an iterative process guided by the literature and influenced by the interview guide and the data: *perceptions of student participation* and *perceived factors shaping student participation in classroom discourse*. Together, these themes shaped the understanding of the teachers' instructional rationale and possible tensions in engaging students in classroom discourse.

## 5 Findings

Six episodes were identified as classroom discourse episodes: two group-work episodes (10 and 60 minutes) and one whole-class episode (three minutes) in Anna's lessons, and three whole-class episodes, each lasting just under 20 minutes, in Bea's lessons. In the following, the different episodes and discursive moves are described (see detailed results in the [Appendix](#)), followed by interview findings of the teachers' instructional rationales.

## 5.1 Anna's classroom discourse practice

Anna engaged her students in classroom discourse mainly through assigning group work of complex tasks. In Anna's Episode 1, students work in pairs using GeoGebra with triangle tasks. The episode contains 82 dialogic moves, 61 authoritative moves, and nine blended moves. This episode especially provoked the dialogic moves *asking real/open questions* (N=27) and *spontaneous contributions from students* (N=26) commenting on content or asking Anna questions such as "To construct a perpendicular line—was it like this?" The most common authoritative move by far was *narrative* (N=52), manifested in Anna controlling the narrative by guiding and managing group work ("Now I want you to focus on this task").

During group work, Anna frames the rules for participation, illustrated in the following excerpt (lines 3-4) from Episode 2, when she checks in on a peer discussion, requiring all students to be involved in the mathematical discussions. She challenges her students in line with a focusing pattern (Wood, 1998) (lines 6-11), prompting them to explain their mathematical ideas. The task at hand is to figure out whether any of a set of triangles are right triangles.

- 1 Anna: Maja, you tell me what your group has done.
- 2 Maja: I didn't have a calculator. I couldn't hear what they said.
- 3 Anna: Now you [to the group] need to share so that Maja also hears what
- 4 you are doing.
- 5 Lotta: We just take  $a^2 + b^2 = c^2$
- 6 Anna: Yes, and what is that?
- 7 Lotta: I don't know . . . I don't remember
- 8 Anna: Do you remember, Jani?
- 9 Jani: I don't know.
- 10 Anna: Maja?
- 11 Anna: Why can we use this? Why does it work? I let you think about that.

This example illustrates how Anna balances authoritative and dialogic moves, as she controls the students' discussion, yet uses dialogic moves encouraging students to continue exploring mathematics in their discussions by asking for justifications and prompting students to explain their ideas (Franke et al., 2009).

In Anna's Episode 2, three to four students work in groups on triangle tasks, equipped with a whiteboard, which they use to show their process and solution. In this episode, there is a balance of dialogic (N=21) and authoritative moves (N=20); a few moves are blended (N=4). This episode also provoked *asking real/open questions* (N=9) and *spontaneous contributions from students* (N=5), while the most common authoritative moves were *narrative* (N=8) and *providing evaluative responses*

(N=8). During both group-work episodes, there were a combined 15 instances of the peer-work codes *prompting students to explain to peers* and *encouraging students to talk mathematics*.

Anna's Episode 3 is a short whole-class episode summarizing peer work on triangles. In contrast to Anna's first two episodes, this is characterized by authoritative teacher moves (N=5).

- 1 Anna: Okay, let's freeze here. All groups have realized that we need to use
- 2 the Pythagoras' theorem in some way. Some didn't remember its name, but
- 3 you all knew it. But what is difficult is to know why we use Pythagoras'. I
- 4 heard at least two groups who could tell why. So, Mia, you can tell me since
- 5 you knew why do we use Pythagoras' theorem?
- 6 Mia: Because it only works on a right triangle to find the hypotenuse.
- 7 Anna: So the requirement for Pythagoras' theorem is that the sum of all
- 8 the squared lengths is the hypotenuse squared—this formula. In this case it
- 9 is  $a^2 + b^2 = c^2$ . If you know the length of two sides, you can find out the
- 10 length of the third side, but the whole point here is that this only works in
- 11 a right triangle, and that is why you can use it to test whether this triangle is
- 12 a right-angle one.

Anna sums up why the Pythagorean theorem is needed for solving this task by selecting a student contribution she emphasizes as correct, providing an evaluative response (lines 4-5), then paraphrasing what the student said, and lecturing (*narrative*) on why the Pythagorean theorem works to test whether a triangle is a right triangle (lines 7-12). Such authoritative moves help bring the lesson forward and give all students a chance to recall why a particular method worked (Temple & Doerr, 2012).

## 5.2 Anna's instructional rationale

Anna is in her fourth year of teaching. She teaches both mathematics and science and actively participates in professional development programs. In the interview, Anna uses the term *inquiry-based* to describe her teaching. She states that she wanted to move away from patterns “where you just review theory and procedures, and students perform the same procedures individually.” She found this “traditional way” lacking in respect to student learning: “I wanted to find a new way of teaching, a way where students would learn more.” According to Anna, her teacher education did not provide tools for teaching mathematics in a way other than the traditional, but she found a like-minded mentor and a network of study friends with whom she shares tasks, ideas, and experiences. Parents have questioned her methods, but she perceives that the

school leadership and the new curriculum support her way of teaching: “I realized that the people behind the curriculum think the same way as me.” The combination of having a network, a mentor, and support in the curriculum and school leadership appears to have given her a sense of having a professional knowledge base and security to continue developing student-engaging and inquiry-based teaching.

**Perceptions of student participation.** For Anna, student participation in classroom discourse means students engaging in peer discussions around whiteboards, initiated by questions she poses, or students replying individually on digital devices. Anna states that peer work and student engagement in discussions are necessary for teaching inquiry-based and complex problems and that discussions “make them think.” But she states that students also must learn how to work productively in groups, as simply placing them into groups does not automatically enhance learning. In the observed lessons, Anna frames student discussions in multiple instances (N=15) by prompting them to explain to their peers (e.g., checking whether all students in the group follow the discussion) or encouraging them to discuss mathematics (e.g., focusing discussions toward justification of solutions instead of simply providing solutions).

**Perceived factors shaping student participation in classroom discourse.** Anna mentions both school-based and student-related factors as constraining student participation in classroom discourse. The key school-based factor was the necessity to maintain the same pace as all other ninth-grade classes because they have the same tests, preventing her from longer explorations of a topic, which is similar to curriculum constraints reported by Herbel-Eisenmann et al. (2006). Student-related constraints were social factors, such as balancing equity while simultaneously paying attention to students’ well-being and sense of security. Anna perceived the traditional method of students raising hands in a whole-class setting as “only activating the ablest ones.” She states that the inquiry-based approach demands active students, which provokes insecurity in some students not used to working on tasks without prescribed procedures: “Some students do not feel safe in my way of teaching; they miss the traditional way.” To tackle their insecurity, she explicitly credits such students’ performance in front of the class and provides mathematical challenges on all levels so that even the most skilled students sometimes struggle, thus normalizing incorrect answers. Nevertheless, Anna states that some students must be “left alone,” as they are so uncomfortable speaking spontaneously in class. Hence, even though Anna embraces the idea that students learn through participating in



discussions, there seems to be a tension between that and another more pressing concern of certain students' well-being.

### 5.3 Bea's classroom discourse practice

The following example illustrates how classroom discourse in all three of Bea's episodes consisted of long, uninterrupted flows of teacher lecture (lines 10-21), punctuated by short student contributions in IRE format (Cazden, 1988) (lines 7 and 9), with a focus on rules and procedures. Bea reviews a task she has noticed several of her students struggling with. The task is to solve  $3\frac{1}{3}^{-2}$  and it is written on the board.

- 1 Bea: First, I want to remove the 3, so I multiply 3 with this part of the
- 2 fraction,  $3*3$ , which is  $9 + 1$ . I write it as  $\frac{10^{-2}}{3}$ . Can you see this? Then I
- 3 look at my rules. I think it was our rule 8; look in your notebooks. If I have
- 4 a negative exponent, what should I do with the nominator and denominator
- 5 to make it plus, positive? What shall I do with it? Fredrik, what should I do
- 6 with the 10 and the 3?
- 7 Fredrik: We should solve them.
- 8 Bea: No, we don't solve them. What did you do, Allan?
- 9 Allan: You change their positions.
- 10 Bea: We change their positions. That was our last rule. It is in your books,
- 11 and we also wrote it down. If I have  $\frac{a^{-n}}{b}$ , to get rid of the negative here, I
- 12 can absolutely not put it in front of here with a minus—like put the minus in
- 13 front of the fraction and then the parenthesis, and then it is good. No, to
- 14 remove the negative exponent, I change the positions. So b, the
- 15 denominator, will be up in the nominator, and the old nominator will be
- 16 the denominator, and then I change from minus to plus. So 3 here, and 10
- 17 down here, and the parenthesis is from -2 to +2; do you follow? So the
- 18 next rule, I write here 8 since it is our rule number 8. Then I use rule
- 19 number 7 to remove the parenthesis. What shall I do when I have a
- 20 parenthesis with a nominator and denominator squared? How do I remove
- 21 it?

In Bea's Episode 1, she reviews exponent tasks and elicits student answers on these tasks in whole-class format. The discursive moves Bea uses include mostly different authoritative moves (87%), especially *narrative* (N=25), *providing evaluative responses* (N=20), *cued elicitation* (N=11), and *sequences of repeated questions* (N=11). Dialogic moves used were students providing *spontaneous contributions* (N=8), *providing neutral responses* to students (N=2), and *asking real/open questions* (N=1).

In Episode 2, almost all moves are authoritative (97%). Bea reviews a list of exponent rules and occasionally engages students by asking questions in the form of

cued elicitation, as exemplified in the following excerpt (lines 1-2) when Bea gives Ludde a heavy clue of the right mathematical operation to apply when dividing  $8^{13}$  and  $8^{11}$ .

- 1 Bea: When we have division between  $8^{13}$  and  $8^{11}$ , it is the same base. What
- 2 do you think it will be here? If it is not addition, it could be . . . Ludde?
- 3 Ludde: Erhm . . . subtraction?

In Episode 3, Bea reviews homework tasks in whole-class format, again guided by mostly authoritative moves (77%), and the most common ones were *narrative* (N=14), *providing evaluative responses* (N=6), and *cued elicitation of students' contributions* (N=5). The dialogic moves (17%) consisted of *spontaneous student contributions* (N=5) and *re-voicing/reflecting on student responses* (N=3).

The following table summarizes the different moves Anna and Bea enacted in their discourse episodes.

**Table 3.** Overview of teaching moves.

	Authoritative moves	Dialogic moves	Blended moves
Anna Ep 1	40%	54%	6%
Anna Ep 2	44%	45%	9%
Anna Ep 3	100%	0%	0%
Bea Ep 1	87%	13%	0%
Bea Ep 2	97%	3%	0%
Bea Ep 3	77%	17%	6%

#### 5.4 Bea's instructional rationale

Bea has been teaching for 30 years in different grades. She has a double degree in mathematics and special education and actively participates in professional development courses and workshops. While colleagues inspire her, she states that the new curriculum has not changed her instruction. Bea describes her instruction in her basic group as different from in a mixed or advanced group: "I explain more and use more everyday language so that they won't get lost." She states that she spends less time reviewing theory in advanced groups, who then have more time for seatwork on difficult tasks.

**Perceptions of student participation.** For Bea, student participation in discourse means students listening and answering her questions. In Bea's class, verbal participation is voluntary, which she ensures by letting "everyone who raises their

hands gets to answer.” Nevertheless, she appreciates students’ questions: “I like when there is discussion among me and the students, when they ask things, not only me answering my own questions.” Students’ spontaneous questions and comments are also the most common dialogic move observed across Bea’s episodes (N=6). However, the majority of student utterances were short replies given when Bea tried to elicit the right answers to procedural tasks (see above example). Yet Bea also states that students giving the wrong answer is helpful, as she then can try to detangle misunderstandings. In Bea’s view, the teacher reviewing content followed by individual seatwork is the most common instructional pattern for her and her colleagues: “I have been a teacher for many years, and I help out in many classrooms, and this is what we all do.”

**Perceived factors shaping student participation in classroom discourse.** Bea remarks on student-related factors as constraining student participation in classroom discourse. In her view, pressing participation in discourse would be detrimental for her students’ well-being, as some students have strong negative feelings about mathematics and may have other problems that pressure them. She has agreed with some students to never ask them anything when they are unprepared. For Bea, her most important job as a teacher is to “see my students and let them know that I care,” and that is why she prefers to guide and discuss with students individually during seatwork. Another student-related concern is her view of the learning needs of her “basic” students: “Mathematics is a lot about structure and students who have issues concentrating need strict guidance on how to apply rules to not get lost.” This resembles results in other studies, where teachers who perceive students as struggling academically or having low future aspirations in mathematics “need” basic mathematics (e.g., Sztajn, 2003; Atweh et al., 1998). Bea thus doubts the learning value of peer discussions in her ninth-grade classroom: “I’m not sure what kind of mathematics these students could discuss. They would discuss everything else but mathematics.” Further, she views discussions of complex problems as disadvantageous for struggling students: “I tried it once. It was chaos. Only the high-achieving students understood.” These statements imply what other studies have highlighted before (e.g., Brendefur & Frykholm, 2000), which is that beliefs about how students learn and what mathematics is shape instruction, as Bea perceives that these students learn best by listening and that engaging in discourse would be a waste of time. Nevertheless, Bea reflects that the future of mathematics instruction will be different: “I think it will be that you start with a phenomenon or a problem, and then

you build it up from there. I could never do it with my ninth-graders because everybody has to learn. I would have to guide every one of them. But I think it is the future and a huge challenge for teachers.” Thus, while Bea appears to perceive a tension between student participation in discourse and the needs of struggling students, she also recognizes that mathematics teaching and the role of the mathematics teacher is changing in an inquiry-based and discourse-rich direction.

## 6 Discussion

As agreement about the benefits of student participation in mathematics discourse grows (e.g., OECD, 2016a; Walshaw & Anthony, 2008), national curricula in Finland and beyond are starting to promote such instructional practices. This study scrutinized two teachers’ instructional rationales and perceived tensions related to student participation in discourse with the combined analytical foci of teacher perspectives and instructional moves. Findings indicate that the teachers use different discursive moves to engage students. By balancing dialogic and authoritative teacher moves, Anna exemplifies instruction that provides opportunities for all students to participate in what may be called productive mathematical discourse (Fennema et al., 1996). Bea’s authoritative moves exemplify instruction where classroom discourse is limited, and student participation means giving short answers in IRE format (Cazden, 1988) and answering cue-elicited questions (Wood, 1998). While their discourse practices varied vastly, their instructional rationales reflected similar concerns. The following discussion will focus on three main concerns reflected in their rationales—student learning needs, equity, and student emotional well-being—and how teachers with different discourse practices may perceive them as in agreement or in tension with engaging students in classroom discourse.

### 6.1 Instructional rationale for student participation in discourse

The rationales for the diverging discourse practices seem to be shaped by and grounded in similar values and concerns of student needs, emphasizing student learning, student emotional well-being, and equity. Anna’s and Bea’s instructional rationales reflected different views of what it means to learn mathematics and what kind of instruction addresses their students’ learning needs, a difference often demonstrated in research on beliefs and practices (see Kupari, 2003), including research on different classroom discourse practices (e.g., Brendefur & Frykholm,

2000; Sztajn, 2003). Anna's views of learning mathematics were in agreement with the strand of mathematics education research and reform curricula, emphasizing that *all students* should learn how to think and construct knowledge by discussing (e.g., Lampert & Blunk, 1998). Bea held more traditional views of learning mathematics and viewed peer discussions as fruitless for *struggling students*, as she perceived that they needed strict procedural guidance implemented with traditional teacher-centered instruction, similar to the study by Atweh and colleagues (1998). The teachers' different perceptions of student learning needs were also reflected in how they mentioned *equity* as a motivation for their enacted discourse practices, and they differed in how they sought to facilitate equitable practices. Equitable practice for Anna was activating all students through group work, while for Bea, it was giving all students structure and rules through teacher-centered instruction as well as individual guidance.

While the teachers reflected on different views of student learning needs and how to enact equitable practice, they shared concerns about *insecure and shy students* never participating in any kind of classroom discourse. They both suggested that mathematics anxiety and out-of-school issues impaired student engagement in discussions, and challenging such students verbally would conflict with attending to student emotional well-being. They had different ways to engage the most insecure students: Anna gave explicit public recognition to insecure students and attempted to normalize wrong answers by asking all students challenging questions, and Bea discussed mathematics privately during individual seatwork, as she perceived that this was how she could attend to unique student needs.

## 6.2 Overcoming tensions in engaging students in discourse

Teachers such as Anna and Bea socialize students to participate in mathematical discussions in very different ways, likely resulting in very different communication skills (O'Connor, 1998). Anna seems to have embraced the idea of communicative learning for all students, while Bea, though recognizing it as the future of mathematics education, does not seem convinced that such instruction is beneficial for her basic-level students. Drawing on the literature's ideals of mathematics discourse, the discourse practice represented by Anna, balancing authoritative and dialogic moves, are preferred over the discourse practice represented by Bea, of mainly authoritative moves (e.g., Scott et al., 2006; Boerst et al., 2011). Bea's practice may even be seen as problematic, as participation in discourse is considered to improve learning (e.g.,

OECD, 2016a), as well as motivation (Kiemer et al., 2015). However, Bea's rationale for not engaging low-achieving students in group discussions receives support in research suggesting that peer work does not necessarily benefit the learning of struggling students (Bergem & Klette, 2016; Gersten et al., 2009). In diverse classrooms, students have different instructional needs, and some teachers, such as Bea in this study, perceive a tension between talking mathematics and student needs. This finding implies a need for more nuance into the discussion that a high degree of dialogic teacher moves and active students in classroom discourse is a goal independent of student characteristics and classroom context, as assessments of classroom discourse should not neglect how contextual factors shape instruction (Skott, 2019). Instead of focusing solely on the beneficial learning opportunities in "talking mathematics," perhaps tensions between dominant discourses in mathematics education literature and local teachers' concerns—such as student learning needs, student well-being, and equity—could be addressed and recognized in teacher education when focusing on practices that enable "productive mathematical discourse" (Fennema et al., 1996). In addition, the different discourse practices that these teachers represent in the classroom, in combination with their different rationales, might be applicable to the rationales of other teachers with similar patterns of practices. To build more knowledge on this topic, I suggest that future research also focuses on how different styles of teaching relates to instructional rationales. Moreover, research on tensions and teachers' concerns and more good examples of instructional practices balancing discursive moves while attending to different students' needs may assist teachers in developing instructional repertoires that allow all students the opportunity to experience learning mathematics while also developing skills to participate in mathematics discourse (see Sfard, 2003). Anna's instruction—supported by the new curriculum, a mentor, a network of colleagues, and school leadership—may give indications of how teachers' classroom discourse practices can address some of the tensions *and* develop equitable norms for participation. For example, Anna's framing of peer work by scaffolding discourse (Kovalainen & Kumpulainen, 2005) socially (e.g., checking for equal participation in groups) as well as analytically (e.g., prompting students to explain solutions) shows potential for developing productive norms for student participation in content-related discussions. Such knowledge of how to scaffold discourse is especially important to address in in-service and pre-service teachers in the Finnish context, since the traditional instructional patterns in mathematics education (e.g., Kaasila et al., 2010; Taajamo et

al., 2014) may not be sufficient to give equitable opportunities for students to develop mathematical thinking and communicative skills or address the decline in student motivation and achievement in mathematics (Portaankorva-Koivisto et al., 2018).

### 6.3 Limitations of the study

Three aspects of this study in particular limit the conclusions that can be drawn; sample size, differences in mathematical content, and ability groups. First, small samples have received criticism for providing understandings that are so context-specific that they cannot generate any generalizable knowledge (e.g., Richardson, 1996). However, such small case studies highlighting different aspects of teacher rationales build a theory on factors shaping classroom discourse, as shown in the review part (see Chapter 3) of this paper. The short period of three lessons may also be seen as a limitation — however the purpose of this study was not to map out the instructional repertoire of these specific teachers, but to demonstrate different discursive practices. Second and third, the different mathematical content taught (Ayalon & Even, 2016) and the different ability levels of the students (Atweh et al., 1998) are likely to shape classroom discourse. Regardless of these differences, it is by contrasting the instructional rationales of teachers with differing discourse practices that we can learn about perceived tensions and how teachers deal with them, which in turn may inform teacher educators of issues that are important to address in teacher education and professional development.

## 7 Concluding remarks

The significance of the study lies in its approach to studying the instructional rationale behind different kinds of classroom discourse practices in a Finnish context, which facilitates understanding of possible tensions and perspectives associated with classroom discourse practices. This study has shown that teachers' instructional rationales for differently enacted classroom discourse practices may be motivated by concerns for student well-being, learning, and equity, which some teachers perceive as in tension and contradictive to mathematics education literature's ideals of classroom discourse. This study thus provides nuance for contemporary ideals for mathematics classroom discourse by highlighting how teachers with similar values perceive tensions and find solutions for developing discourse practices, which is an insight that could inform teacher educators in a Finnish context and beyond.

## Acknowledgements

This study is part of the Linking Instruction to Student Achievement (LISA) study, funded by the Norwegian Research Council (#222620/F10). I would like to thank my research group SISCO, the rest of the LISA team, and my colleagues Guri and Jelena, for your constructive comments and support.

## References

- Alexander, R. J. (2006). *Towards dialogic teaching: Rethinking classroom talk* (3rd ed.). Cambridge, UK: Dialogos.
- Andersson, E., & Klette, K. (2016). Teachers' use of questions and responses to students' contributions during whole class discussions: Comparing language arts and science classrooms. In K. Klette, O. K. Bergem, & A. Roe (Eds.), *Teaching and learning in lower secondary schools in the era of PISA and TIMSS* (pp. 63-84). Cham, Netherlands: Springer.
- Artigue, M., & Blomhøj, M. (2013). Conceptualizing inquiry-based education in mathematics. *ZDM*, 45(6), 797–810.
- Atweh, B., Cooper, T. J., & Bleicher, R. E. (1998). The construction of the social context of mathematics classrooms: A sociolinguistic analysis. *Journal for Research in Mathematics Education*, 29(1), 63–82. <https://doi.org/10.2307/749718>
- Ayalon, M., & Even, R. (2016). Factors shaping students' opportunities to engage in argumentative activity. *International Journal of Science and Mathematics Education*, 14(3), 575–601. <https://doi.org/10.1007/s10763-014-9584-3>
- Ball, D., & Bass, H. (2000). Making believe: The collective construction of public mathematical knowledge in the elementary classroom. In D. Phillips (Ed.), *Yearbook of the national society for the study of education, constructivism in education* (pp. 193-224). Chicago, US: University of Chicago Press.
- Bergem, O. K., & Klette, K. (2010). Mathematical tasks as catalysts for student talk: Analysing discourse in a Norwegian mathematics classroom. In Y. Shimizu, B. Kaur, R. Huang, & D. Clarke (Eds.), *Mathematical tasks in classrooms around the world* (pp. 35-62). Rotterdam, Netherlands: Sense Publishers.
- Bergem, O. K., & Klette, K. (2016). Conversations as learning tools in mathematics: What do pupils actually learn? In K. Klette, O. K. Bergem, & A. Roe (Eds.), *Teaching and learning in lower secondary schools in the era of PISA and TIMSS*. Cham, Netherlands: Springer International Publishing.
- Bernard, H. R. (2011). *Research methods in anthropology: Qualitative and quantitative approaches* (5th ed.). Lanham, USA: AltaMira Press.
- Boerst, T., Sleep, L., Ball, D., & Hyman, B. (2011). Preparing teachers to lead mathematics discussions. *Teachers College Record*, 113, 2844–2877.
- Brendefur, J., & Frykholm, J. (2000). Promoting mathematical communication in the classroom: Two pre-service teachers' conceptions and practices. *Journal of Mathematics Teacher Education*, 3(2), 125–153. <https://doi.org/10.1023/A:1009947032694>
- Cazden, C. B. (1988). *Classroom discourse: The language of teaching and learning*. Portsmouth, USA: Heinemann.



- Cobb, P., & Bowers, J. (1999). Cognitive and situated learning perspectives in theory and practice. *Educational Researcher*, 28(2), 4–15. <https://doi.org/10.2307/1177185>
- Cohen, L., Manion, L., & Morrison, K. (2011). *Research methods in education*. London, UK: Routledge.
- Confrey, J. (2017). Research: To inform, deform, or reform? In J. Cai (Ed.), *Compendium for research in mathematics education*. Reston, US: National Council of Teachers of Mathematics.
- Davis, B., Towers, J., Chapman, O., Drefs, M., & Friesen, S. (2019). Exploring the relationship between mathematics teachers' implicit associations and their enacted practices. *Journal of Mathematics Teacher Education*. <https://doi.org/10.1007/s10857-019-09430-7>
- Drageset, O. G. (2015). Different types of student comments in the mathematics classroom. *The Journal of Mathematical Behavior*, 38, 29–40. <https://doi.org/10.1016/j.jmathb.2015.01.003>
- Ellis, M. W., & Berry, R. Q., (2005). The paradigm shift in mathematics education: Explanations and implications of reforming conceptions of teaching and learning. *Mathematics Educator*, 15(1), 7–17.
- Emanuelsson, J., & Sahlström, F. (2008). The price of participation: Teacher control versus student participation in classroom interaction. *Scandinavian Journal of Educational Research*, 52(2), 205–223. <https://doi.org/10.1080/00313830801915853>
- Fennema, E., et al. (1996). A longitudinal study of learning to use children's thinking in mathematics instruction. *Journal for Research in Mathematics Education* 27(4): 403–434.
- Finnish National Agency for Education. (2014). *Grunderna för läroplanen för den grundläggande utbildningen* [National core curriculum for basic education]. [http://www.oph.fi/lp2016/grunderna\\_for\\_laroplanen](http://www.oph.fi/lp2016/grunderna_for_laroplanen)
- Fives, H., & Gill, M. G. (2015). *International handbook of research on teachers' beliefs* (Vol. 13). New York, US: Routledge.
- Franke, M. L., Kazemi, E., & Battey, D. (2007). Understanding teaching and classroom practice in mathematics. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 225-256). Charlotte, USA: Information Age Publishing.
- Franke, M. L., Webb, N. M., Chan, A. G., Ing, M., Freund, D., & Battey, D. (2009). Teacher questioning to elicit students' mathematical thinking in elementary school classrooms. *Journal of Teacher Education*, 60(4), 380–392. <https://doi.org/10.1177/0022487109339906>
- Furtak, E. M., & Shavelson, R. J. (2009). Guidance, conceptual understanding, and student learning: An investigation of inquiry-based teaching in the US. In T. Janik & T. Seidel (Eds.), *The power of video studies in investigating teaching and learning in the classroom* (pp. 181-205). Münster: Waxmann.
- Gersten, R., Chard, D. J., Jayanthi, M., Baker, S. K., Morphy, P., & Flojo, J. (2009). Mathematics instruction for students with learning disabilities: A meta-analysis of instructional components. *Review of Educational Research*, 79(3), 1202–1242. <https://doi.org/10.3102/0034654309334431>
- Gravemeijer, K., Bruin-Muurling, G., Kraemer, J.-M., & Van Stiphout, I. (2016). Shortcomings of mathematics education reform in the Netherlands: A paradigm case? *Mathematical Thinking and Learning*, 18(1), 25–44. <https://doi.org/10.1080/10986065.2016.1107821>
- Harding, J. (2013). *Qualitative data analysis from start to finish*. Los Angeles, USA: SAGE.
- Hemmi, K., & Ryve, A. (2015). Effective mathematics teaching in Finnish and Swedish teacher education discourses. *Journal of Mathematics Teacher Education*, 18(6), 501–521
- Herbel-Eisenmann, B. A., Lubienski, S. T., & Id-Deen, L. (2006). Reconsidering the study of mathematics instructional practices: The importance of curricular context in understanding

- local and global teacher change. *Journal of Mathematics Teacher Education*, 9(4), 313–345. <https://doi.org/10.1007/s10857-006-9012-x>
- Ingram, J., Andrews, N., & Pitt, A. (2019). When students offer explanations without the teacher explicitly asking them to. *Educational Studies in Mathematics*, 101(1), 51–66. <https://doi.org/10.1007/s10649-018-9873-9>
- Kaasila, R., Pehkonen, E., & Hellinen, A. (2010). Finnish pre-service teachers' and upper secondary students' understanding of division and reasoning strategies used. *Educational Studies in Mathematics*, 73(3), 247–261. <https://doi.org/10.1007/s10649-009-9213-1>
- Kiemer, K., Gröschner, A., Pehmer, A.-K., & Seidel, T. (2015). Effects of a classroom discourse intervention on teachers' practice and students' motivation to learn mathematics and science. *Learning and Instruction*, 35(C), 94–103. <https://doi.org/10.1016/j.learninstruc.2014.10.003>
- Kieran, C., Forman, E., & Sfard, A. (2001). Guest editorial learning discourse: Sociocultural approaches to research in mathematics education. *Educational Studies in Mathematics*, 46(1), 1–12. <https://doi.org/10.1023/A:1014276102421>
- Klette, K., Blikstad-Balas, M., & Roe, A. (2017). Linking instruction and student achievement: Research design for a new generation of classroom studies. *Acta didactica*, 11(3): 1–19.
- Klette, K., et al. (2018). “Justice through participation: student engagement in Nordic classrooms.” *Education Inquiry* 9(1): 57–77.
- Kovalainen, M., & Kumpulainen, K. (2005). The discursive practice of participation in an elementary classroom community. *Instructional Science: An International Journal of Learning and Cognition*, 33(3), 213–250. <https://doi.org/10.1007/s11251-005-2810-1>
- Kupari, P. (2003). Instructional practices and teachers' beliefs in Finnish mathematics education. *Studies in Educational Evaluation*, 29(3), 243–257. [https://doi.org/10.1016/S0191-491X\(03\)90012-5](https://doi.org/10.1016/S0191-491X(03)90012-5)
- Lampert, M., & Blunk, M. L. (1998). *Talking mathematics in school: Studies of teaching and learning*. Cambridge, UK: Cambridge University Press.
- Lemke, J. L. (1990). *Talking science: Language, learning, and values*. Norwood, USA: Ablex.
- Luoto, J., Klette, K., & Blikstad-Balas, M. (in review) Comparing activity formats and instructional patterns in mathematics classrooms: Cases of instruction in Norwegian and Swedish-Speaking Finnish lower secondary classrooms.
- Mercer, N., & Hodgkinson, S. (2008). *Exploring talk in schools: Inspired by the work of Douglas Barnes*. London, UK: SAGE.
- Mortimer, E., & Scott, P. (2003). *Meaning making in secondary science classrooms*. Maidenhead, UK: McGraw-Hill Education.
- National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, US: NCTM.
- O'Connor, C. (1998). Language socialization in the classroom. In M. Lampert & M. L. Blunk (Eds.), *Talking mathematics in school: Studies of teaching and learning in school* (pp. 17–55). New York, US: University Press.
- O'Connor, C., Michaels, S., Chapin, S., & Harbaugh, A. G. (2017). The silent and the vocal: Participation and learning in whole-class discussion. *Learning and Instruction*, 48, 5–13. <https://doi.org/10.1016/j.learninstruc.2016.11.003>
- Organisation for Economic Co-operation and Development. (2016a). *Equations and inequalities making mathematics accessible to all*. PISA, OECD Publishing. Retrieved from: [https://read.oecd-ilibrary.org/education/equations-and-inequalities\\_9789264258495-en#page5](https://read.oecd-ilibrary.org/education/equations-and-inequalities_9789264258495-en#page5)

- Organisation for Economic Co-operation and Development. (2016b). Selecting and grouping students. In *PISA 2015 results (volume II): Policies and practices for successful schools*. Paris: OECD Publishing.
- Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, 62(3), 307–332. <https://doi.org/10.3102/00346543062003307>
- Patton, M. Q. (2015). *Qualitative Research & Evaluation Methods* (4 ed.). Los Angeles: Sage.
- Pehkonen, E., Ahtee, M., & Lavonen, J. (2007). *How Finns learn mathematics and science*. Rotterdam: Sense.
- Pehkonen, L. (2007). To change or not to change. *Nordic Studies in Mathematics Education*, 12(2), 57–76.
- Pekkarinen, T., & Uusitalo, R. (2012). Peruskoulu-uudistuksen vaikutukset [The reformation of compulsory school]. *Kansantaloudellinen aikakauskirja*, 2(108), 128–139. Retrieved from <https://www.taloustieteellinenyhdistys.fi/images/stories/kak/KAK22012/kak22012pekkari nen.pdf>
- Philipp, R. A. (2007). Mathematics teachers' beliefs and affect. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 257-315). Charlotte, USA: Information Age.
- Portaankorva-Koivisto, P., Eronen, L.-., Kupiainen, S., & Hannula, M. S. (2018). Lukion ensimmäinen yhteinen matematiikan kurssi - mielekästä ja merkityksellistä? [The first common mathematics course in upper secondary – meaningful and worthwhile?]. *FMSERA Journal*, 2(1), 57–65.
- Raymond, A. (1997). Inconsistencies between a beginning elementary teacher's mathematics beliefs and teaching practice. *Journal for Research in Mathematics Education*, 28(5), 550–576.
- Reichenberg, O. (2018). Identifying mechanisms of teaching practices: A study in Swedish comprehensive schooling. *Scandinavian Journal of Educational Research*, 62(2), 213–228.
- Richardson, V. (1996). The role of attitudes and beliefs in learning to teach. In J. Sikula, T. Buttery, & E. Guyton (Eds.), *Handbook of research on teacher education* (Vol. 2, pp. 102-119). New York, US: Macmillan.
- Scott, P. (1998). Teacher talk and meaning making in science classrooms: A Vygotskian analysis and review. *Studies in Science Education*, 32(1), 45–80. <https://doi.org/10.1080/03057269808560127>
- Scott, P., Mortimer, E. F., & Aguiar, O. G. (2006). The tension between authoritative and dialogic discourse: A fundamental characteristic of meaning making interactions in high school science lessons. *Science Education*, 90(4), 605–631. <https://doi.org/10.1002/sce.20131>
- Sfard, A. (2003). Balancing the unbalanceable: The NCTM standards in light of theories of learning mathematics. In J. Kilpatrick, G. W. Martin, & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 353-392). Reston, US: National Council of Teachers of Mathematics
- Silverman, D. (1993). *Interpreting qualitative data: Methods for analysing talk, text and interaction*. London: Sage.
- Simola, H., Kauko, J., Varjo, J., Kalalahti, M., & Sahlström, F. (2017). *Dynamics in education politics: Understanding and explaining the Finnish case*: London, UK: Routledge.
- Skott, J. (2001). The emerging practices of a novice teacher: The roles of his school mathematics images. *Journal of Mathematics Teacher Education*, 4(1), 3–28.
- Skott, J. (2019). Understanding mathematics teaching and learning “in their full complexity.” *Journal of Mathematics Teacher Education*, 22(5), 427–431. <https://doi.org/10.1007/s10857-019-09446-z>

- Solomon, Y., & Black, L. (2008). Talking to learn and learning to talk in the mathematics classroom. In N. Mercer & S. Hodgkinson (Eds.), *Exploring talk in schools: Inspired by the works of Douglas Barnes* (pp. 73-90). London: SAGE Publications Ltd.
- Spillane, J. P. (2002). Challenging instruction for “all students”: Policy, practitioners, and practice. JCPR Working Papers 253, Northwestern University/University of Chicago Joint Center for Poverty.
- Sztajn, P. (2003). Adapting reform ideas in different mathematics classrooms: Beliefs beyond mathematics. *Journal of Mathematics Teacher Education*, 6(1), 53–75.  
<https://doi.org/10.1023/A:1022171531285>
- Taajamo, M., Puhakka, E., & Välijärvi, J. (2014). *Opetuksen ja oppimisen kansainvälinen tutkimus TALIS 2013: Yläkoulun ensituloksia* [The international TALIS 2013 study of learning and teaching]. Retrieved from:  
<https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/75219/okm15.pdf?sequence=1>
- Temple, C., & Doerr, H. (2012). Developing fluency in the mathematical register through conversation in a tenth-grade classroom. *Educational Studies in Mathematics*, 81(3), 287–306. <https://doi.org/10.1007/s10649-012-9398-6>
- Walshaw, M., & Anthony, G. (2008). The teacher’s role in classroom discourse: A review of recent research into mathematics classrooms. *Review of Educational Research*, 78(3), 516–551.  
<https://doi.org/10.3102/0034654308320292>
- Webb, N. M., Nemer, K. M., Chizhik, A. W., & Sugrue, B. (1998). Equity issues in collaborative group assessment: Group composition and performance. *American Educational Research Journal*, 35(4), 607–661.
- Wood, T. (1998). Alternative patterns of communication in mathematics classes: Funneling or focusing? In H. Steinbring, M. G. B. Bussi, & A. Sierpiska (Eds.), *Language and communication in the mathematics classroom* (pp. 167-178). Reston, USA: National Council of Teachers of Mathematics.
- Xu, L., & Clarke, D. (2013). Meta-rules of discursive practice in mathematics classrooms from Seoul, Shanghai and Tokyo. *ZDM*, 45(1), 61–72.
- Xu, L., & Clarke, D. (2019). Speaking or not speaking as a cultural practice: Analysis of mathematics classroom discourse in Shanghai, Seoul, and Melbourne. *Educational Studies in Mathematics*, 102(1), 127–146.

## Appendix. Results: teacher moves\*

Dialogic moves	<i>Anna Ep 1</i>	<i>Anna Ep 2</i>	<i>Anna Ep 3</i>	<i>Bea Ep 1</i>	<i>Bea Ep 2</i>	<i>Bea Ep 3</i>
<i>Asking real/open questions</i>	27	9	0	1	0	0
<i>Spontaneous contributions</i>	26	5	0	8	1	5
<i>Revoicing/reflecting on student responses</i>	5	2	0	0	0	3
<i>Meaning into matter</i>	9	0	0	0	0	0
<i>Promoting disagreement / leaving lack of consensus</i>	0	0	0	0	0	0
<i>Providing neutral responses to students</i>	16	5	0	2	0	0
<i>Teacher prompts students to explain to peers</i>	5	2	0	0	0	0
<i>Teacher encourages students to talk mathematics together</i>	6	2	0	0	0	0
Authoritative moves	<i>Anna Ep 1</i>	<i>Anna Ep 2</i>	<i>Anna Ep 3</i>	<i>Bea Ep 1</i>	<i>Bea Ep 2</i>	<i>Bea Ep 3</i>
<i>Cued elicitation of students' contributions</i>	9	5	0	11	4	5
<i>Sequence of repeated questions</i>	0	2	0	11	1	1
<i>Selecting and/or ignoring students' contributions</i>	0	0	1	0	0	0
<i>Reconstructive paraphrase or recap</i>	0	0	1	2	2	0
<i>Formulaic phrases</i>	0	1	0	0	1	1
<i>Marking significance</i>	1	0	0	2	1	1
<i>Narrative</i>	52	8	2	25	18	14
<i>Promoting/establishing consensus</i>	0	0	0	0	0	0
<i>Providing evaluative responses</i>	9	8	1	20	2	6

\*This includes overlaps, e.g., blended moves when utterances were coded for both authoritative and dialogic moves

