ACTA UNIVERSITATIS GOTHOBURGENSIS
GOTHENBURG STUDIES IN EDUCATIONAL SCIENCES 502

# Does teaching quality matter for student learning outcomes?

## A student perspective of a mathematics classroom

Lena Asp

UNIVERSITY OF GOTHENBURG

# Does teaching quality matter for student learning outcomes?

ACTA UNIVERSITATIS GOTHOBURGENSIS
GOTHENBURG STUDIES IN EDUCATIONAL SCIENCES 502

# Does teaching quality matter for student learning outcomes?

## A student perspective of a mathematics classroom

Lena Asp

UNIVERSITY OF GOTHENBURG

© LENA ASP, 2025
ISBN 978-91-7963-219-9 (printed)
ISBN 978-91-7963-220-5 (pdf)
ISSN 0436-1121

The publication is also available in full text at:
http://hdl.handle.net/2077/85610

Subscriptions to the series and orders for individual copies sent to: Acta Universitatis Gothoburgensis, PO Box 222, SE-405 30 Göteborg, Sweden or to acta@ub.gu.se

Cover artist: Zahra Hasani Yourdshahi

Photographer: Fredrik Asp

Print:
Stema Specialtryck AB, Borås, 2025

# Abstract

Title: Does teaching quality matter for student learning outcomes? A student perspective of a mathematics classroom

Author: Lena Asp

Language: English


ISBN: 978-91-7963-219-9 (printed)

ISBN: 978-91-7963-220-5 (pdf)

ISSN: 0436-1121

Keywords: teaching quality, classroom management, instructional clarity, academic self-concept, academic achievement, equity, TIMSS


High-quality teaching is assumed to provide students with learning opportunities that may mitigate educational inequities and narrow achievement gaps. The underlying question guiding this doctoral thesis is: to what extent does teaching quality matter for student learning outcomes? As a starting point, the thesis assumes that the quality of teaching within the classroom influences both students' affective and academic learning outcomes. Moreover, it is assumed that for teaching and learning to occur, the teacher must establish a positive classroom climate by applying structural-organisational activities. Rationally, it is suggested that success in this respect requires good leadership skills. Once a positive learning environment is in place, high-quality instruction and teaching become possible. By contrast, in classrooms lacking structure and a positive climate, students could not only be hindered in developing a positive attitude and confidence in the subject but could also experience reduced motivation and limited achievement.

Measuring teaching quality, however, is challenging both theoretically, conceptually, and methodologically. Empirical results are inconsistent, while some studies report significant and positive relations between teaching quality and student learning outcomes, others do not. Additionally, there is still limited research on these relationships, particularly from the perspective of younger students. Consequently, the overarching aim of this thesis is to contribute to the research field with empirical studies examining the relationships between two dimensions of teaching quality (classroom management and instructional clarity) and students' mathematics confidence and mathematics achievement while accounting for student background characteristics such as socioeconomic status (SES), language spoken at home, and gender.

The thesis is a secondary analysis of data from the international large-scale assessment Trends in International Mathematics and Science Study (TIMSS) 2019,

Grade 4. Teaching quality is conceptualised using student-reported perceptions of their teacher's instructional activities in the mathematics classroom. The rationale is twofold. First, aggregating students' perceptions of the teacher's actions in the classroom gives a valid and reliable measure of teaching quality. Second, the focus is on the mathematics classroom, as mathematics is primarily taught and learnt in the classroom context, unlike reading, which could be more influenced by factors outside the classroom. In this thesis, teaching quality is pragmatically conceptualised in accordance with the Three Basic Dimensions framework, using the two available student questionnaire scales in TIMSS 2019: classroom management and instructional clarity.

This thesis comprises an integrative essay and three empirical studies addressing teaching quality from different angles. Study I examines the construct validity of the mixed-worded scale of mathematics confidence in TIMSS 2019, Swedish Grade 4. Using confirmatory factor analysis and exploratory structural equation modelling, the nine items in the scale of mathematics confidence are analysed. Possible method effects and response bias are examined along with a semantic validation of the cross-cultural translations. Additionally, using structural equation modelling, the relationships between two teaching quality dimensions and mathematics achievement are examined, with mediation through mathematics confidence. Study II examines the relationships between dimensions of student-perceived teaching quality (classroom management and instructional clarity) and the two outcome variables of mathematics confidence and mathematics achievement. Student background factors are accounted for. Multilevel structural equation modelling is used to separate classroom variation from individual variation when examining the relationships. Study III widens the scope to include four Nordic countries (Denmark, Finland, Norway, and Sweden), to examine teaching quality aspects and the relations to the two outcome variables of mathematics confidence and mathematics achievement. The Nordic context is suitable for analysis as these countries share educational values such as the compensatory task of the educational system, the Nordic Model. Measurement invariance testing is used to assess the possibility of cross-country comparisons. To address the hierarchical structure of the data, multilevel structural equation modelling is applied as it separates individual from classroom variation.

In all, findings from this doctoral thesis give several conclusions. First, there is evidence of method effect from the negatively worded items in the scale of mathematics confidence. Including mixed-worded items for questionnaire scales is a strategy used to keep respondents attentive. However, careful considerations and examinations have to be done regarding their construct validity before including such items in secondary analyses. Therefore, researchers conducting secondary analyses are suggested to examine the items provided by international large-scale assessments (ILSAs) such as TIMSS when conceptualising a latent construct with mixed-worded items. Next, teaching quality factors are significantly related to mathematics confidence and mathematics achievement. The relationship between instructional clarity and mathematics confidence is particularly strong at the student level. Similarly, the relationship between classroom management and mathematics achievement is significant at the student level, but also substantial at the classroom level in Denmark and Sweden. Student background factors related significantly to mathematics achievement, particularly at the classroom level. The findings revealed classroom-level composition effects across the Nordic countries. In Sweden, classroom-level mathematics achievement is influenced by the aggregated SES and the language spoken at home. Classrooms with higher SES and more students speaking Swedish at home performed higher on the mathematics assessment. In Finland, the classroom-level SES significantly influenced both classroom-level mathematics confidence and mathematics achievement.

# Contents

# List of tables and figures

To Fredrik

# Acknowledgements

I would like to express my deepest appreciation to my supervisors. Dr. Alli Klapp has been my main supervisor and a constant presence throughout this doctoral project. Thank you for your expertise, knowledge, support, and friendship. You have strengthened me and continuously challenged me to move forward over these years. Dr. Monica Rosén has been my co-supervisor. I could not have undertaken this journey without her support and reassurance. She encouraged and challenged me until her sudden passing. Her friendship and wisdom are deeply missed. Dr. Victoria Rolfe stepped in as my co-supervisor. I am extremely grateful for her knowledge, expertise, encouragement, and immense support. Thank you, Victoria.

I would like to extend my sincere thanks to the discussants at my planning, mid and final seminars, Dr. Jan-Erik Gustafsson, Dr. Maria Rasmusson, and Dr. Mari-Pauliina Vainikainen. Your thorough reading, detailed commentary, and expertise greatly contributed to this thesis. I would also like to extend my appreciation to the defence opponent Dr. Leonidas Kyriakides.

Many thanks to the entire FUR research group at Gothenburg University for their expertise and for providing a stimulating research environment. A special thanks to Dr Stefan Johansson for his profound knowledge of ILSA data and his support.

This thesis is part of the ASSESS doctoral school, and I am deeply grateful for being entrusted with the opportunity to embark upon this journey. I am grateful for support and friendship within the Gothenburg ASSESS team, especially as we began our doctoral studies during the pandemic and were only able to meet online during the first year. Many thanks to my friends in "Skrivarstugan", your friendship and support have been invaluable. I also wish to thank the entire ASSESS community at GU, SU, and UU, and all my fellow doctoral colleagues, these years would not have been the same without you. A special thank you to those who joined me at conferences over the years. Particularly Mari, Elpis, David, Zahra, and Panos, your friendship, laughter, tears, and mutual support are memories I will always value.

A special thank you to Zahra for the wonderful cover art, you're an amazing artist. I would also like to thank David for reading my manuscript at a late stage. Special thanks to Anna for reading the Swedish version.

On a personal note, I would like to thank all my friends, near and far, for their moral support, encouragement, shared laughter, and all the joyful moments we have experienced throughout this journey.

Last but not least, I would like to thank my children, Erik, Elin, and Elias for filling my life with happiness and laughter. Thank you for your unwavering support, valuable insights, and for being you! Finally, a special thank you to Fredrik. This thesis is dedicated to you. No words can truly express the depth of my appreciation. I am so profoundly grateful for your patience, support, encouragement, and love throughout this doctoral project. Thank you for always being there by my side, for being my best friend and my pillar in life – thank you!

Onsala, May 2025

# Foreword

I have a background as a language teacher, 'lead teacher'[1], head of department, and principal and have enjoyed teaching and interacting with my students in the classroom throughout the years. My doctoral project arose from a wish to examine individual and classroom differences in academic and affective learning outcomes. I wished to understand why some groups of students outperform others across classrooms. Specifically, why there are differences in performance across classrooms, even if the teachers in each classroom have equivalent qualifications? Additionally, I was interested in why the same teacher might experience varying levels of success when teaching different groups of students. An easy and obvious response would be that student learning outcomes are influenced by their background and aptitude. However, as a teacher, I strongly believe that what I, as a teacher, do when teaching, engaging, and interacting with the students in the classroom matters. Therefore, I wished to examine if different aspects of teaching quality related differently to academic achievement and affective outcomes. Maybe this information could provide insight into possible educational inputs on how to mitigate the widening achievement gap between high- and low-achieving students. In Sweden, educational inequity is growing, for example, there is an increasing number of students that leave compulsory school without complete grades. Therefore, to address the overarching aim of this thesis, I used the international large-scale assessment TIMSS to assess what the students reported concerning the teaching quality in their classrooms and what the relationships were to their learning outcomes while accounting for student background factors. However, to be able to examine differences between student and classroom learning outcomes, I embarked on a journey into a previously unfamiliar world, that of statistical methods. These methods serve as essential tools for addressing the research questions of this thesis. Therefore, I wish to emphasise that this is not a thesis on statistical methods, rather, these analytical tools serve as a means to explore what happens inside classrooms.

[1] Referring to the Swedish term 'förste lärare'.

# Chapter 1 Introduction

This thesis is a secondary analysis of international large-scale assessment (ILSA) data, and its overarching aim is twofold. First, it explores the relationships between aspects of teaching quality and student learning outcomes by drawing on student questionnaire data. Second, it examines the construct validity of the student questionnaires scales used in the studies, particularly the use of mixed-worded items.

The first section of Chapter 1 provides background information and the rationale for the thesis. This is followed by a presentation of the aim and research questions guiding the integrated essay. Next, a reader's guide briefly introduces the three empirical studies, accompanied by an overview of the studies. The chapter concludes with a clarification of the structure of the integrated essay.

## Background

Education is recognised as a fundamental human right (UNESCO, 2019). It empowers individuals, fosters more equitable societies, and contributes to improved health and economic prosperity (Gustafsson, 2016; OECD, 2015). Beyond these key benefits, education is a cornerstone for cultivating critical thinking, strengthening the agency of citizenship, and enabling democratic participation, and ultimately, it is a foundation for democracy (Nations, 2021). Extensive research acknowledges the influence of student background factors, such as socioeconomic status (SES), on student achievement (e.g., Coleman et al., 1966; Hattie & Yates, 2013; Sirin, 2005). Previous research also proposes that teacher quality (i.e., characteristics, qualifications, credentials, knowledge, and skills) and teaching quality (i.e., the teacher's observable behaviour, instructions, and practices in the classroom) are essential for student learning outcomes (Blikstad-Balas et al., 2021; Hattie, 2009) and academic achievement (Brophy & Good, 1984; Shulman, 1987). Teaching encompasses the classroom interactions between teachers and students to facilitate students' learning (Cohen et al., 2003; Hiebert & Grouws, 2007). However, teaching is not a consistent phenomenon but fluctuates depending on the student composition and the interactions between the teacher and the students within each classroom (Fauth et al., 2020).

High-quality teaching is suggested to be essential for the compensatory role of the educational system in promoting equity (Guerriero, 2017; OECD, 2017). In the Nordic countries, the compensatory role of education is referred to as the 'Nordic Model'. The Nordic Model emphasises equal opportunities, fairness, inclusion, and equity for each student (Blossing et al., 2014; Frønes et al., 2021; Lundahl, 2016). According to the Nordic Model, each student should have access to high-quality education and educational opportunities regardless of their background or prerequisites (Blossing et al., 2014). However, there are indications of increasing inequity and widening achievement gaps in Nordic countries, particularly Sweden (SOU 2019:40; Yang Hansen & Gustafsson, 2019). These manifested inequities in Sweden are reflected in increasing demographic segregation, where students are self-sorted by free school choice into schools depending on socioeconomic status (SES) or immigration background (Gustafsson et al., 2016; Holmlund et al., 2014; SOU 2017:35; SOU 2019:40).

Teaching quality has been suggested as a key compensatory factor in mitigating the growing achievement gaps, potentially reducing the negative influence of SES on students' academic achievement and affective outcomes (Brophy & Good, 1984; Goe, 2007; Hattie, 2009; Kyriakides et al., 2013; Scherer & Gustafsson, 2015; Wang et al., 1993; Yang Hansen & Gustafsson, 2019). However, measuring teaching quality is a complex and challenging task and previous research has employed various theoretical frameworks, definitions and instruments (Blikstad-Balas et al., 2021; Blömeke & Olsen, 2019; Goe, 2007; Klieme & Nilsen, 2022). In educational effectiveness research, scholars have investigated the relationships between teachers' instructional practices and student learning outcomes. These studies frequently employ secondary analysis with a quantitative approach drawing on international large-scale assessments (ILSAs) and theoretical frameworks such as the dynamic model (Creemers & Kyriakides, 2007) and the three basic dimensions (TBD) framework (Baumert et al., 2010; Kunter et al., 2013; Praetorius et al., 2018). Empirically, the results are inconsistent; some researchers report significant and positive relations between different aspects and conceptualisations of teaching quality and student achievement while others do not (e.g., Bellens et al., 2019; Blömeke et al., 2022; Blömeke & Olsen, 2019; Goe, 2007; Pianta & Hamre, 2009; Praetorius et al., 2018; Senden et al., 2022). Of key importance for these inconsistent patterns is how teaching quality has been conceptualised and what data and methods have been used. Consequently, uncertainty still exists regarding the relationships between teaching quality and student learning outcomes, particularly for younger students in primary school (Klieme & Nilsen, 2022). Therefore, it is crucial to gain further knowledge of whether different aspects of teaching quality relate to student learning outcomes (Guerriero, 2017; Klieme & Nilsen, 2022).

This doctoral thesis therefore sets out to address these research gaps with further empirical information. The presumed relations between teaching quality and student learning outcomes are investigated through a quantitative approach using secondary analysis of data from an international large-scale assessment (ILSA). Over the last three decades, ILSAs such as the Programme for International Student Assessment (PISA) and Trends in International Mathematics and Science Study (TIMSS) have become increasingly influential in shaping educational policy worldwide (Rutkowski et al., 2010). Policymakers frequently use results from these assessments to inform reforms and improve educational systems (Johansson, 2016).

Secondary analysis of ILSA data regularly incorporates contextual questionnaire data, such as students' self-reported background information and attitudes about learning. These data facilitate examinations of plausible factors that may explain variations in student learning outcomes within and between countries. While ILSAs provide robust, high-quality data with high psychometric standards, concerns about reliability and validity must be addressed when building latent constructs using questionnaire-based measures in secondary analyses. Specifically, scales with reverse direction or complex wording may threaten construct validity, introducing measurement errors that may impact the statistical findings (Hooper et al., 2013; Steinmann, Sánchez, et al., 2022; Steinmann, Strietholt, et al., 2022). Previous research has suggested secondary analyses should omit unreliable measures such as negatively worded items, especially for younger students (Hooper et al., 2013; Roszkowski & Soven, 2010). If construct validity issues are left unaddressed, secondary research findings may be affected by method effects, distorting the statistical relationships. However, few studies discuss the validation process of the questionnaire items before conceptualising latent constructs for secondary analysis. Thus, further research is needed to assess the validity of items with reverse direction or complex wording in secondary analysis of ILSA data (Hooper et al., 2013).

The thesis leans on the student-reported perceptions of their teacher's teaching practices in the mathematics classroom reflecting teaching quality. By analysing student-perceived teaching quality in the classroom context, it becomes possible to study the dynamics within the 'black box' of the classroom, capturing the interactions that are difficult to observe directly (Nilsen et al., 2016). The teacher's actions in the classroom are perceived by the students, whose aggregated perceptions provide a reliable and valid measure of teaching quality (Scherer & Gustafsson, 2015).

In Nordic primary schools, a generalist teacher, who is certified to teach the class in at least four subjects, leads a class for three years, from Grades 1 to 3 (age 7-9) before another generalist teacher takes over from Grades 4 to 6 (age 10-12). Throughout primary school, classes encounter additional specialist teachers (e.g. physical education, music). In secondary school (age 13-16), students transition to specialised subject-matter teachers for each subject (Volmari, 2019). Consequently, in Sweden, younger students remain in the same classroom with the same generalist teacher for three years. This structure suggests that primary school classrooms provide a stable social learning environment where teaching and learning occur (Vygotsky, 1978).

The rationale for choosing mathematics as the subject of interest in this thesis is that mathematics is primarily learnt at school, in a classroom environment. Unlike for example reading literacy, which can be acquired via multiple contexts outside of school, learning mathematics depends on formal teaching. Subsequently, it could be assumed that teaching quality would be better captured in a mathematics classroom, where students rely on teacher-led mathematics instruction to acquire mathematical knowledge. Mathematics instruction has distinct characteristics that may differentiate it from other subjects, such as social sciences, where inquiry-based learning, discussions and project-based assignments are more common. In general, mathematics teaching is direct and characterized by procedural fluency, conceptual understanding and problem-solving (Cohen et al., 2003; Lester, 2007). Therefore, mathematics classrooms could include more specific instructions compared to subjects that may emphasise broader instruction.

In Sweden, mathematics teachers in Grade 4 are generalist teachers, meaning they instruct the class in at least four subjects. Hence, students in primary school would encounter the same teacher across multiple subjects, which could create instructional overlaps between subjects. Mathematics instruction in fourth-grade classrooms could be assumed to be more fundamental and more similar to instruction in other subjects. In contrast, eighth-grade mathematics instruction is more specialised, requiring advanced instructional methods to support students' progression. Despite these differences between mathematics instruction and other subjects, certain core instructional principles, such as clear instructions, explanations, and appropriate support, apply across subjects.

## Research aims and outline

The thesis has the assumption that teaching quality is a multidimensional concept situated in a classroom context as a starting point that can explain parts of the variance in students learning outcomes, such as confidence and achievement in mathematics. Teaching is the broad, holistic process aimed at facilitating learning, including pedagogical strategies and interactions. It refers to the observable classroom behaviour of the teacher and encompasses instructional and organisational factors that shape the learning environment. Logically, it could be assumed that the teachers' structural-organisational activities, such as classroom management, are a prerequisite for establishing a positive and supportive climate in the classroom. Without a supportive and orderly environment, meaningful teaching, instruction, or learning cannot occur. Success in this respect requires good leadership skills (Brophy & Good, 1984; Goe, 2007). Once a positive, supportive classroom climate is established, characterised by positive relationships between the teacher and the students, as well as among peers, the teacher can focus on instructional practices, such as giving clear instruction and ensuring students receive time on task to facilitate their learning. Thus, instruction is part of teaching, and it refers to the delivery of specific content and a method to convey information to the students.

Teaching quality could play a role in enhancing student learning outcomes but also help reduce achievement gaps, contributing to increased educational equity (Darling-Hammond, 2021; Guerriero, 2017). Research suggests that teaching quality is particularly important for mitigating the negative influence of low SES on student learning outcomes (Brophy & Good, 1984; Hattie, 2009). In Sweden, the achievement gap between low- and high-SES students has widened. For instance, an increasing number of students currently leave compulsory school without final grades in all subjects which raises concerns about educational equity (Skolverket, 2024a; Yang Hansen & Gustafsson, 2019). These trends highlight the need for a better understanding of how teaching quality may be used to better support disadvantaged students.

Despite its importance, teaching quality is a complex, multidimensional concept that is difficult to conceptualise and measure reliably. Research addressing the validity of teaching quality indicators is still scarce. Similarly, research on the relationships between teaching quality and both academic and affective outcomes remains limited, particularly for younger students in Grade 4 (Klieme & Nilsen, 2022). The primary objective of this thesis is to contribute empirical knowledge on the relationships between student-perceived aspects of teaching quality in mathematics classrooms, students' confidence and achievement in mathematics and student background factors in Sweden and three additional Nordic countries. The Nordic countries, with their similar educational systems and cultural contexts, provide an ideal context for examining various dimensions of teaching quality and its relations to learning outcomes.

By drawing on data from the 2019 Trends in Mathematics and Science Study (TIMSS), Grade 4, three empirical studies examine the relationships between aspects of student-perceived teaching quality in mathematics classrooms and students' confidence and achievement in mathematics. Given that this thesis takes a quantitative approach using secondary analyses of TIMSS questionnaire data, an essential first step is to ensure the construct validity of TIMSS indicators before employing them in statistical analyses. The following overarching research questions guide the thesis:

1. To what extent can the student questionnaire items validly measure the latent constructs of mathematics confidence and two subdimensions of teaching quality, classroom management and instructional clarity?
2. To what extent do aspects of teaching quality matter for student learning outcomes such as mathematics confidence and mathematics achievement and are there differences across groups of students depending on their socioeconomic status, home language, and gender?
3. To what extent are there differences between classrooms in the relationship between aspects of teaching quality and learning outcomes such as mathematics confidence and mathematics achievement?
4. To what extent could aspects of teaching quality mitigate the negative influence of student background factors and decrease the growing achievement gap?

## Guide for readers

The doctoral thesis comprises an integrative essay and three empirical studies addressing teaching quality. Each study builds upon findings from the previous study, gradually expanding the research scope. A condensed overview of the three empirical studies is presented in Table 1.

**Study I** focused on the construct validity of the measures derived from TIMSS 2019 data, which were utilised in the subsequent analyses. Particular focus was on the construct validity of the mixed-worded mathematics confidence scale. The scale comprised nine items, both positively and negatively worded. Confirmatory factor analysis (CFA) and exploratory structural equation modelling (ESEM) were employed to assess the construct validity of the scale. Additional analyses examined possible method effects, response bias, and semantic validation of the cross-cultural translations. A refined conceptualisation and operationalisation of the latent construct of mathematics confidence was presented. Additionally, in a structural equation model (SEM), the relationships between two subdimensions of teaching quality (classroom management and instructional clarity) and mathematics achievement were examined, with mathematics confidence as a mediating factor.

In **Study II**, the focus was broadened as the validated measures were employed to examine the relationships between student-perceived teaching quality (classroom management and instructional clarity) in mathematics classrooms, and the two learning outcomes, mathematics confidence and mathematics achievement in the Swedish educational context. Structural equation modelling (SEM) was employed when exploring these relationships. A multi-level approach was applied to explain differences between classrooms and to keep classroom variation separate from individual variation.

**Study III** widened the scope by examining fourth graders' perceptions of teaching quality in Denmark, Finland, Norway, and Sweden, along with the relations between the two learning outcomes, mathematics confidence and mathematics achievement. Student background factors were included in the analyses. Measurement invariance testing was employed to determine the possibility of conducting cross-country comparisons. Given the hierarchical structure of the data, multilevel structural equation modelling (MSEM) was applied to explain differences between classrooms.

The integrative essay is composed of eight chapters followed by three empirical articles. Chapter 2 presents the theoretical frameworks for the thesis and a conceptual framework for the constructs in focus. Chapter 3 presents the background in a literature review. Chapter 4 gives a brief overview of the Swedish and Nordic educational context. Chapter 5 presents the methodology, and analytical methods, and discusses the reliability, validity, and validation as well as ethical considerations of the thesis. Chapter 6 presents an overview of the main findings of the three empirical studies, and Chapter 7 follows with an integrated discussion of the contribution of the thesis, along with its limitations, future research, and conclusions. The final chapter presents a Swedish summary.

Table 1 Guide for readers

| | Study I | Study II | Study III |
|---|---|---|---|
| Title | Validating the mixed-worded mathematics confidence scale in TIMSS 2019: Examining its relationships to teaching quality and mathematics achievement | The associations between student-perceived teaching quality and mathematics confidence and mathematics achievement: A study of Swedish Grade 4 TIMSS 2019 | Does teaching quality matter for Nordic primary school students' mathematics confidence and mathematics achievement? A multilevel structural equation analysis of Nordic TIMSS 2019 Grade 4 data |
| Author/s | Lena Asp | Lena Asp, Alli Klapp, and Victoria Rolfe | Lena Asp, Alli Klapp, and Monica Rosén |
| Status | Submitted to International *Journal of Educational Research*. Presented at the International Research Conference (IRC) in Dublin, Ireland, June 2023 | Accepted for publication in *Instructional Science*, 2025. Presented at the European Conference on Educational Research (ECER Plus), 2022 | Published in the Journal of *Large-scale Assessments in Education*, 2025, Vol.13, Article 7. Presented at the European Conference on Educational Research (ECER), 2024. |
| Aim | Examine the construct validity of the mixed-format scale of mathematics confidence; the relationships between teaching, possible mediation via mathematics confidence, and mathematics achievement | Examine the relationships between student-perceived teaching quality and two outcomes, mathematics confidence and achievement in Sweden | Examine the relationships between student-perceived teaching quality and two outcomes, mathematics confidence and achievement in four Nordic countries |
| Data & sample | TIMSS 2019 Swedish Grade 4, student questionnaire and mathematics achievement data | TIMSS 2019 Swedish Grade 4, student questionnaire and mathematics achievement data | TIMSS 2019 Nordic Grade 4, student questionnaire and mathematics achievement data |
| Analyses | Confirmatory factor analysis; bi-factor exploratory structural equation modelling; structural equation modelling with mediation | Confirmatory factor analysis; multilevel structural equation modelling, mediation model | Multi-level confirmatory factor analysis; multi-group factor analysis for measurement invariance testing; multi-level structural equation modelling |

# Chapter 2 Theoretical frameworks

The thesis is based on the assumption that teaching and learning take place within the situated classroom interactions between the teacher and peers (Carroll, 1963; Darling-Hammond et al., 2020; Eccles & Wigfield, 2023; Kyriakides et al., 2022). Various theoretical frameworks have been employed to investigate teaching and learning within the educational effectiveness research field. Among these, the *dynamic model of educational effectiveness research*, the *three basic dimensions* (TBD) framework (Senden et al., 2022), and the *conceptual model of determinants of student outcomes* (Nilsen & Gustafsson, 2016) provide key perspectives on how different aspects of teaching contribute to students learning outcomes. This work draws mainly upon two conceptual and theoretical frameworks to investigate two key aspects of teaching quality (classroom management and instructional clarity) and their relationships to students learning outcomes. On the one hand, it leans on the *conceptual model of determinants of student outcomes* (Nilsen & Gustafsson, 2016). On the other hand, it incorporates the *situated expectancy-value theory* (Eccles & Wigfield, 2023; Nilsen & Gustafsson, 2016). Together, these frameworks provide a comprehensive base for examining the relationships between teaching quality and student learning outcomes. To examine the validity of the questionnaire data when operationalising the latent constructs, the thesis draws on Messick's (1989) construct validity theory and latent variable theory, using the frameworks of confirmatory factor analysis and structural equation modelling (Brown, 2015; Kline, 2016) (see Chapter 5, Reliability, validity, and validation).

The first section of this chapter briefly introduces the dynamic model of educational effectiveness research and the three basic dimensions framework. Followed by the conceptual model of determinants of student outcomes. The next section presents the situated expectance-value theory explaining the situated classroom context that influences student motivation and learning. This is followed by a proposed extension of the conceptual model of determinants of student outcomes and the SEVT, which serves as the guiding framework for the thesis. Additionally, the use of the students' self-reports in assessing the classroom context is discussed. Finally, the last sections define the key concepts used throughout the thesis to ensure conceptual clarity.

## Building on the dynamic model and the TBD framework

The thesis is situated within the field of educational effectiveness research (EER), which focuses on understanding factors that contribute to student learning outcomes (Kyriakides & Creemers, 2008). The EER is a line of inquiry rooted in the research from Coleman et al. (1966) on equal educational opportunities for all students (Carroll, 1963; Coleman, 1988; Creemers & Kyriakides, 2007). It aims to identify and analyse the hierarchical nature of the educational system and various factors that contribute to student achievement (Creemers & Kyriakides, 2010; Kyriakides et al., 2022). These include factors such as school leadership, teaching quality, classroom composition, and student background characteristics, that interact across the different levels (e.g., school, classroom and student levels) and directly or indirectly influence student learning outcomes (Creemers & Kyriakides, 2007; Muijs et al., 2014; Reynolds et al., 2014; Scheerens, 2015). To address the different educational levels, the EER encompasses three sub-theories: system effectiveness research, school effectiveness research, and teacher effectiveness research (Scheerens & Blömeke, 2016). System effectiveness research examines the overall effectiveness of the educational system, such as policies, reforms and other structural components (Creemers & Kyriakides, 2007). School effectiveness research addresses both system-level but also different school-specific factors such as the school climate, organisation, and goal attainment (Scheerens, 2015). Finally, teacher effectiveness research addresses the classroom environment and teacher and teaching variables. According to the EER, the classroom level is a key predictor of student learning outcomes (Creemers & Kyriakides, 2010; Muijs et al., 2014).

### Conceptual framework of determinants of student outcomes

Various theories and frameworks have been employed to identify and understand teaching quality (Danielson, 2007; Goe, 2007; Pianta & Hamre, 2009; Praetorius et al., 2018; Senden et al., 2022). Several generic frameworks conceptualise teaching quality with various subdimensions, such as classroom management, cognitive activation, and supportive or social climate. Notable examples are the three basic dimensions (TBD) framework (Klieme & Nilsen, 2022), the classroom assessment scoring system (Pianta & Hamre, 2009), and the dynamic model of educational effectiveness (Creemers & Kyriakides, 2007). The dynamic model addresses the multilevel structure of the educational system and examines various interrelated factors that influence student learning outcomes (Creemers & Kyriakides, 2007). The TBD framework includes three generic subdimensions of classroom teaching: classroom management, supportive climate, and cognitive activation (Praetorius et al., 2020). These dimensions serve as broad measures of teaching quality and can be operationalised flexibly by researchers (Senden et al., 2022).

International large-scale assessment (ILSA) data, such as TIMSS, include measures of both teacher quality and teaching quality, which align to varying degrees with aspects of teaching quality outlined in the dynamic model (Kyriakides & Creemers, 2008; Nilsen & Gustafsson, 2016) and the TBD framework (Klieme & Nilsen, 2022; Praetorius et al., 2018). Building on these two frameworks, Nilsen and Gustafsson (2016) proposed a conceptual model that illustrates the hierarchical structure of the educational system and the determinants of student outcomes (see Figure 1). From the perspective of this thesis, this conceptual framework serves as the central theoretical lens for understanding the relationship between teaching quality in the classroom and student learning outcomes (Nilsen & Gustafsson, 2016, p. 4).

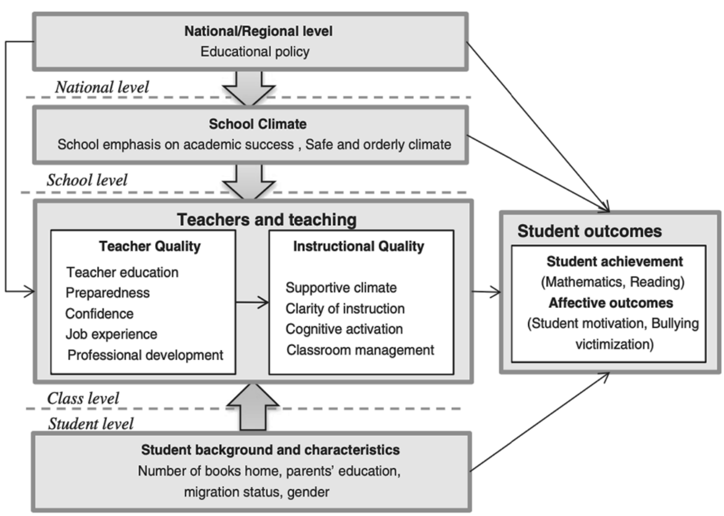


Figure 1 Conceptual model of determinants of student outcomes
(Nilsen, Gustafsson & Blömeke, 2016, p. 4)

The conceptual model proposed by Nilsen and Gustafsson (2016) outlines the hierarchical structure of the educational system, identifying four levels that influence student learning outcomes. The first level is the student level representing the student background characteristics that can explain part of student learning outcomes (Sirin, 2005). The background characteristics include factors such as aptitude, perseverance, gender, SES, and migration status of their parents or guardians. The second level is the classroom level including two key determinants: teacher quality and instructional quality. Teacher quality includes factors such as teacher education, job experience, and confidence. Instructional quality includes four subdimensions: classroom management, supportive climate, clarity of instruction, and cognitive activation (Creemers & Kyriakides, 2007). The third level is the school level with factors such as school climate, school emphasis on academic success, and perceived safety. The fourth level is the regional and national level, where governmental and political decisions shape the Education Act, curricula, and various educational policy documents. This level is the overarching level in which schools operate. These four levels interact, both directly and indirectly, as presented by the arrows in Figure 1 and explain variations in student outcomes, including both academic achievement (e.g., mathematics) and affective outcomes (e.g., student motivation). While all four levels contribute to student learning, (Kyriakides et al., 2020), the main focus of the thesis is on the student and classroom levels. Given that each educational system functions within a specific cultural context with distinct educational values and policies, Nilsen et al. (2016) proposed that the student, classroom, and school levels of the model may vary considerably within and between countries.

While the conceptual model presented by Nilsen and Gustafsson (2016) effectively describes the complexity and dynamics of the hierarchical educational system and determinants of student learning outcomes, it does not address the situated learning context. Research has established that students' motivation and self-concept are shaped not only by individual student and home-related factors but also by the situated learning context and the interactions between peers and the teacher in the classroom (Eccles & Wigfield, 2002; Wang et al., 2019). Furthermore, the social learning context and the teacher-student interactions are acknowledged to influence learning outcomes (Eccles & Wigfield, 2023; Marsh & Martin, 2011; Shavelson et al., 1976). Thus, teaching and learning take place within the situated classroom context, where the teacher and the students interact and jointly co-construct the teaching quality. Therefore, to better understand the classroom processes, this study incorporates the situated expectancy-value theory (SEVT) as a theoretical framework (Eccles & Wigfield, 2023).

## Situated expectancy-value theory

The situated expectancy-value theory (SEVT), rooted in the expectancy-value theory (Eccles & Wigfield, 2002), provides a broad framework for understanding the psychological processes, that guide and motivate an individual's achievement-related choices (Eccles & Wigfield, 2023). It explains how students' beliefs in their ability and expectancy to succeed are shaped by various situational and individual factors. According to the SEVT (see Fig. 2), students' expectancy for success is influenced by situational factors, such as the social context and the classroom environment. These factors include students' relationships with classmates, a sense of belonging, and classroom interactions. Within the classrooms, peers show varying levels of motivation and aptitude, while collaborating, competing, and comparing themselves with one another. The classroom climate and the teacher's instructional practices contribute to the student's perceptions of their ability and expectancy to succeed. The teacher influences the students' beliefs through instruction, feedback, support, and lesson design. Thus, teaching quality influences how the students engage with their learning, either reinforcing or diminishing their self-confidence. However, the SEVT recognises that not only the social context influences students' beliefs in their ability and expectancy to succeed but also individual factors.

These individual factors include the students' perceptions of a task that are shaped by how enjoyable, important, and valuable they find it, as well as the effort they perceive is required of them to successfully complete the task. In turn, these individual factors shape students' motivation and engagement (Eccles & Wigfield, 2023). Extensive research has shown that affective factors, such as motivation or self-concept, are among the strongest non-cognitive predictors of mathematics achievement, though their relative influence may vary (Hattie, 2009; Marsh et al., 2005; Stankov & Lee, 2017; Stankov et al., 2014). Therefore, including the SEVT framework in the thesis facilitates the understanding of how the student's social context could either support or hinder their belief in their ability to succeed, and how the student's social context could influence their learning outcomes. The SEVT framework displays how the dynamics between the situational factors in various classroom interactions, the teaching quality, and individual factors shape student learning experiences and outcomes.

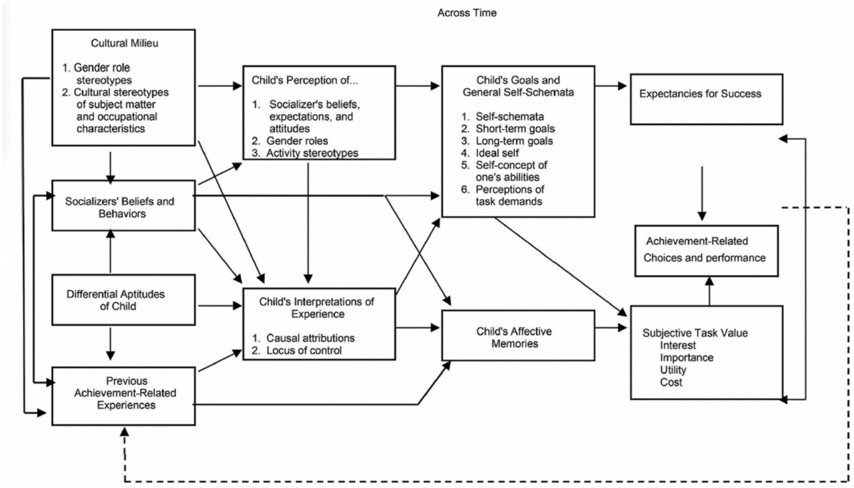


Figure 2 The situated expectancy-value model of achievement choices
(Eccles, Wigfield, 2023, p.2)

In their SEVT model, Eccles and Wigfield (2023, p. 2) illustrate the intricate relationships that directly and indirectly influence students' achievement choices and performances. This section briefly explains these relationships from a classroom perspective (see Fig. 2). Students' aptitude and their belief in success, develop from birth and are shaped by their various life experiences. In the SEVT model, the situational and individual factors shape how the students interpret their surroundings and the classroom context which influence their achievement-related choices and performances (Eccles & Wigfield, 2023). On the left side of the SEVT model, the broader cultural setting is illustrated, emphasising the role of the situational factors, in this case, the classroom context. Students are situated within a '*Cultural Milieu*', such as the classroom, with different cultural stereotypes and gender roles. The '*Socializers' Beliefs and Behaviors*' refer to the interactions between peers and teachers in the classroom context. The cultural milieu of the classroom influences the peers (the socialisers), while peers also influence an individual student's aptitude via social interactions and comparisons. The middle section of the model depicts students' developmental processes, focusing on how students interpret their experiences and develop their self-concept. These interpretations shape affective memories, beliefs about personal ability to succeed, and subjective task value. The right side of the model visualises how the relationships between students' subjective task value and expectancy for success influence students' academic choices and subsequently their performance.

To conclude, while the conceptual model of determinants of student outcomes by Nilsen and Gustafsson (2016) captures the complexity and dynamics of the hierarchical educational system and factors related to student learning outcomes, it does not address the classroom's situated context as the SEVT does (Eccles & Wigfield, 2023). The following section will address this gap with a proposed extension of the conceptual model of determinants of student outcomes.

## Conceptualising the situated classroom context

It is well established that both student background characteristics and the situated classroom context influence student learning outcomes (Eccles & Wigfield, 2023; Marsh & Martin, 2011; Shavelson et al., 1976). However, a comprehensive conceptual model that integrates factors from the situated classroom context remains underdeveloped. To bridge this gap, this thesis builds on the conceptual model by Nilsen and Gustafsson (2016) and the SEVT framework (Eccles & Wigfield, 2023), to propose an extended conceptual model of teaching quality and the situated classroom context (SCC). This model expands the classroom level to include not only teacher and teaching characteristics but also the situated classroom context that shapes student learning outcomes. Each student brings unique background characteristics, aptitudes, academic emotions, perceptions of a learning activity and its importance to the classroom. The composition of students within a classroom shapes the situated classroom context (SCC) by influencing interactions, relationships, climate, and learning outcomes, as described in the SEVT framework (Eccles & Wigfield, 2023). The SCC is jointly constructed by both teachers and the students, and is shaped by their characteristics, and developed continuously through their interactions within each classroom (Fauth et al., 2020).

This *conceptual model of teaching quality* and *the situated classroom context* serves as a guiding framework for examining and understanding teaching quality as a classroom-specific construct in this thesis. It provides a structured approach for analysing how the classroom environment interacts with teaching quality and relates to student learning outcomes (see Fig. 3). The next section will further elaborate on this conceptual model.

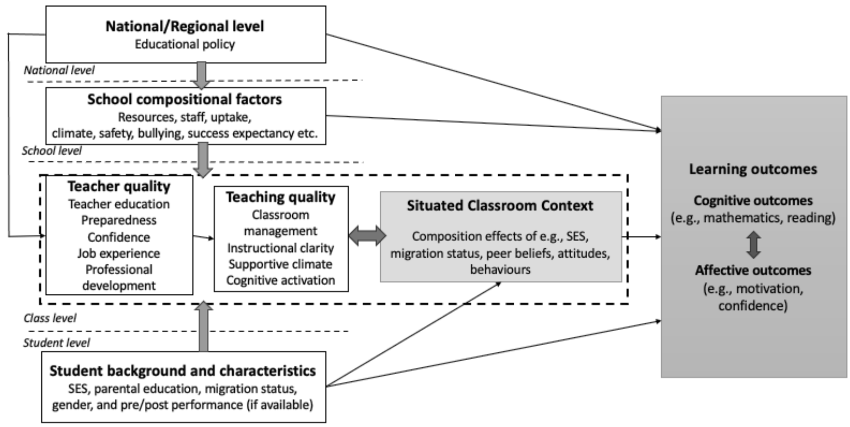


Figure 3 Conceptual model of teaching quality and the situated classroom context
A proposed extension of the conceptual model of determinants of student outcomes by Nilsen et al. (2016)

It is widely recognized that students' learning takes place within the social context of a classroom, where teachers and peers support, encourage or distract one another (Bruner, 1996; Coleman, 1988; Vygotsky, 1978). Each student brings unique individual background characteristics, such as aptitude, academic emotions, and perceptions of an activity and its importance, which in turn shape their classroom experiences. Any shift in the classroom dynamics affects both the teacher's and the students' behaviours, self-beliefs, and learning outcomes (Eccles & Wigfield, 2023). This thesis proposes an extension of the conceptual model by Nilsen and Gustafsson (2016) incorporating teaching quality and the situated classroom context. The *conceptual model of teaching quality and the situated classroom context* (see Figure 3), includes three key components at the classroom level: teacher quality, teaching quality, and the situated classroom context (SCC). There is a dynamic relationship between these three key components that is co-created by the teachers and the students in each unique classroom.

The *conceptual model of teaching quality and the situated classroom context* highlights how the clustering of students in classrooms creates varying composition effects at the classroom level. Composition effects refer to the aggregated characteristics of students within a classroom and their relationship to both individual and overall class learning outcomes (Burke & Sass, 2013). The learning environment can be shaped by different kinds of composition effects, such as socioeconomic status (SES), linguistic characteristics, affective attitudes, and learning abilities. One key composition effect is SES which influences students' relationship to learning outcomes and also to teaching quality. Previous research suggests that high-SES classrooms may benefit more from teaching quality, as students in these classrooms are often more familiar with the teacher's communication styles and interactions (Atlay et al., 2019). In contrast, students in low-SES classrooms may have limited home educational resources and parental support making it more difficult for them to comprehend these communication styles. As a result, they may feel more distanced from the teacher's interactions (Sortkaer, 2019). Linguistic diversity is another composition effect. In more diverse and multilingual classrooms, students may require additional instruction and language support, while monolingual classrooms, with a linguistically homogenous student group, may require less language support (Hansson, 2012). Students' affective attitudes toward learning contribute to composition effects. In classrooms with less motivated students, disruptive behaviour may be more frequent, reducing the overall learning opportunities for the class. In contrast, classrooms with highly motivated students may give a calmer and more positive learning environment (Vansteenkiste & Ryan, 2013). Another significant composition effect refers to students' aptitude and learning abilities. In a classroom with a majority of low-achieving students, the teacher must prioritise basic knowledge requirements and provide additional instruction and support. In contrast, in a classroom with high-achieving students, more focus may be on challenging tasks (Brophy, 2000).

Additionally, the Big-Fish-Little-Pond Effect (Marsh et al., 2019), illustrates how a student's self-concept is influenced by their relative ranking within the classroom. A high-achieving student in a classroom where most peers are high-achievers may develop a lower self-concept compared to equally capable students in lower-achieving classrooms. Thus, accounting for different kinds of composition effects when examining factors that could related to student learning outcomes is of importance.

The proposed *conceptual model of teaching quality and the situated classroom context* includes two student learning outcomes, a cognitive outcome, such as mathematics achievement, and an affective outcome, such as mathematics self-concept. These outcomes are interrelated, reflecting a reciprocal relationship, where high achievement promotes high self-concept, which in turn leads to improved achievement (Marsh & Martin, 2011; Marsh et al., 2005; Ramazan et al., 2023). See Figure 3.

When analysing cross-sectional data such as TIMSS, it is important to recognise that the relationships identified through statistical analyses reflect directional associations between the variables. While these models specify the direction of effects based on theory, they do not establish causality. Disentangling the relationships between different composition effects, such as affective attitudes and learning abilities, is a complex task, as there could be a question of reverse causality (Marsh & Craven, 2006). A key challenge is determining what causes what. For example, high teaching quality may improve students' achievement, but it is also possible that high-achieving classrooms perceive the teaching as of better quality. The possibility of reverse causality complicates the interpretation of the relationships between teaching quality and student achievement. For instance, low teaching quality may lead to lower student achievement, but it could also be that low-achieving classrooms contain a large proportion of students who perceive the teacher's instructions as difficult to understand and therefore perceive the teaching quality as lower (Nilsen et al., 2018; Ramazan et al., 2023). However, not all composition effects are subject to reverse causality. Factors such as gender, SES, and immigration background are fixed and cannot be influenced by the classroom processes. Thus, while these factors can shape student learning outcomes, they are not affected by them.

To conclude, students are part of a situated classroom context where teaching and learning take place. At the classroom level, each student brings unique background characteristics, aptitude, academic emotions, and perceptions of an activity and its importance to the classroom. In turn, the composition of the classroom influences the teaching and learning processes. The teaching and learning processes and the students' experiences in the classroom are suggested to be mutually reinforcing. Consequently, teachers' instructional practices are shaped by the classroom dynamics, which in turn, will influence both the students' and the teacher's behaviours (Eccles & Wigfield, 2020). The proposed *conceptual model of teaching quality and the situated classroom context* will serve as a guiding framework for this thesis, providing a structured approach to examining classroom factors related to student learning outcomes.

## The student perspective

Self-reported questionnaire data from students and teachers are commonly used when researching the classroom context. There is a general agreement on the face validity of students' and teachers' self-reports, as students and teachers mostly consider the questionnaire items reasonable and relevant (Aditomo & Köhler, 2020; Lüdtke et al., 2009). However, self-reported questionnaire data could be related to issues of bias, reliability, and validity. For example, teachers might not assess their ability objectively, while student ratings might be affected by difficulty in understanding the items, or if students dislike a teacher, they might respond more harshly (Kunter & Baumert, 2006). Previous research suggests that, compared to teacher ratings, student ratings have high predictive validity and are strongly associated with achievement across several subjects (Gaertner & Brunner, 2018; Scherer & Gustafsson, 2015; Wagner et al., 2016). Teacher and student ratings reflect different aspects of the teaching quality in the classroom context, both are valid measures, and the decision to use teacher or student self-reported data should depend on the specific research aim (Kunter & Baumert, 2006).

Students can be considered experts in assessing the teaching quality in their classroom, as they encounter numerous teachers during their schooling and may be better at evaluating classroom dynamics than external observers (Walberg & Haertel, 1980). The teaching and instruction will evoke individual experiences and emotions, which the student interprets, shaping the student's motivation, learning and achievement-related behaviours (Eccles & Wigfield, 2023). According to Eccles and Wigfield (2023), students develop internal and external causal attributions to explain their success or failure. Internal attribution occurs when students attribute success to internal factors, such as effort or ability. In contrast, external attribution occurs when failure is attributed to external factors, such as teaching quality or perceived unfairness from the teacher.

Teaching quality is presumed to be a classroom construct that is reliably captured by student-self-reported ratings (Scherer & Gustafsson, 2015). Student ratings provide researchers with insight into the processes of the classroom 'black box'. In each classroom, students are clustered together and their observations can be aggregated at the classroom level (Scherer & Gustafsson, 2015). By aggregating students' perceptions of teachers' classroom behaviour, individual biased ratings may be reduced (Aditomo & Köhler, 2020). The aggregated student ratings at the classroom level constitute a "reflective construct that reflects a specific teacher characteristic" (Lüdtke et al., 2008, p. 220). This also applies to self-reported data from primary school students (Fauth et al., 2014).

Consequently, the relationships presented in the proposed *conceptual model of teaching quality and the situated classroom context* (Fig. 3) are examined using student-reported perceptions of their teacher's practices and behaviour as indicators of teaching quality.

# Definition of main concepts

## Defining teacher quality

Although teacher quality is not the primary unit of analysis in this thesis, it is pertinent to provide a brief overview of the concept of teacher quality.

Teacher quality is a multidimensional construct encompassing several interrelated subdimensions, including teacher knowledge, qualifications, competence, specialisation, work experience, perceived confidence, and preparedness (Darling-Hammond, 2021; Goe, 2007). Among these, teacher knowledge is central and can be categorised into three knowledge domains: *content knowledge* (CK), *pedagogical content knowledge* (PCK), and *general pedagogical knowledge* (GPK) (Shulman, 1987). CK refers to teachers' knowledge and mastery of the subject matter, while PCK denotes subject-specific instructional strategies. GPK refers to generic teaching-related strategies (Baumert et al., 2010; Leijen et al., 2022) that are applicable across subjects (Shulman, 1987). GPK has been defined as "…subject transcendent knowledge about learning and teaching processes, classroom management and educational context to support students' development and motivation" (Leijen et al., 2022, p. 218). This suggests that effective teaching depends not only on subject-specific knowledge but also on the teacher's ability to manage the classroom environment and adapt teaching practices to meet each student's diverse needs. This includes adapting instructions, interactions, and strategies to meet each student's aptitude, and motivation, as well as the social context of the classroom (Goe, 2007). The relative importance of GPK and PCK varies across educational stages and student age. In primary education, where general teaching strategies and classroom management play a more central role, GPK is more important. As students move on to higher levels of education, PCK becomes increasingly essential as the complexity of subject-specific content increases (Goe, 2007). Both GPK and PCK are important for improving teacher effectiveness (Baumert et al., 2010; Guerriero, 2017).

Previous research on teacher quality has revealed inconsistencies in both its conceptualisation and in the findings regarding its relationship with student learning outcomes (Blömeke et al., 2013). Certain subdimensions of teacher quality have been found to be directly related to students' academic achievement, including teacher certification (Darling-Hammond, 2000), years of teaching experience (Toropova et al., 2019), and subject-matter specialisation (Goe, 2007). This relationship was found to be stronger in Grade 8 than in Grade 4 (Nilsen et al., 2018). Another key factor influencing teacher quality is the variation in the design of teacher education across countries (Blömeke et al., 2016). A recent study using Swedish TIMSS 2019 Grade 4 data reported a positive relationship between teachers' formal competence, measured as a latent factor, and students' mathematics achievement (Lindström et al., 2024). However, not all studies have found direct relationships between teacher quality and academic achievement (Blömeke et al., 2022; Blömeke & Olsen, 2019). Blömeke et al. (2022) found that teacher quality did not directly relate to academic achievement but teaching quality mediated the relationship between teacher quality and student achievement. Their findings suggest that teachers' pedagogical and instructional knowledge are predictors of classroom management and instructional clarity, which in turn influence student learning outcomes.

## Defining teaching quality

In educational research, the definitions *teaching quality* and *instructional quality* have often been used interchangeably. To address this inconsistency, Charalambous and Praetorius (2022) emphasised the need for a standardised vocabulary and recommended adopting teaching quality as the preferred term. At the beginning of my PhD journey, I used the definition of *instructional quality* (Study II), but as my studies and research progressed, I aligned with the recommendation for standardised terminology (Charalambous & Praetorius, 2022) and adopted *teaching quality* in subsequent studies (Study I and III) and in this integrative essay.

Teaching quality is a multidimensional concept referring to the teacher's observable behaviour in the classroom. It comprises several interrelated subdimensions, such as instructional strategies and practices designed to engage students in teacher-student interactions and support their learning (Pianta & Hamre, 2009). In addition to the teacher's observable behaviour and practices, high-quality teaching also requires general pedagogical knowledge (GPK) and pedagogical content knowledge (PCK) (Guerriero, 2017). These knowledge domains help the teacher to manage the classroom effectively and adapt instructions to the needs of the students. Moreover, the social context of the classroom plays a crucial role in shaping teaching and learning processes. Classrooms are situated and co-constructed contexts, where the interactions between students and the teacher simultaneously influence and shape the learning environment (Eccles & Wigfield, 2020).

Previous research on teaching quality has employed both qualitative and quantitative approaches, including various methods, conceptualisations, and perspectives. Teaching quality has been studied from perspectives such as the students', teachers', and external observers' perspectives. Different frameworks define the main aspects of teaching quality (Senden et al., 2022). The conceptual model proposed by Nilsen and Gustafsson (2016) defines teaching quality as comprising four subdimensions: supportive climate, clarity of instruction, cognitive activation, and classroom management (see Fig. 1). Previous research has conceptualised teaching quality using one or several of these sub-dimensions (Kyriakides et al., 2013; Panayiotou et al., 2021). Some studies include three subdimensions of teaching quality (Senden et al., 2022) whereas others focus on a single subdimension (Blömeke & Olsen, 2019; Toropova et al., 2019). In this thesis, teaching quality is pragmatically operationalised comprising the two subdimensions available in the TIMSS 2019 student questionnaire: classroom management and instructional clarity. Previous research has identified classroom management and instruction clarity as core aspects of teaching across different educational systems and cultural contexts. Both aspects are considered fundamental aspects of teachers' GPK, facilitating teachers to adapt their teaching to the situated context of each classroom (König & Blömeke, 2012).

*Classroom management* refers to the structural-organisational activities the teacher uses to create a positive, calm, and productive learning environment. These activities are considered generic across all subject matters. The teacher establishes social norms and rules to facilitate instruction and engage students in a positive working climate, thus providing them with learning opportunities (Nilsen & Gustafsson, 2016). The situated context of each classroom varies across time, cultures, and contexts. This means that classroom management is dynamic and evolves accordingly. The teacher has to adjust their actions to suit individual students and the specific moment in the classroom (Wubbels, 2011).

*Instructional clarity* refers to how effectively teachers use pedagogical techniques to support student learning. In the classroom, teachers must manage a multitude of instructional processes and complex social interactions, adapting instructions, explanations, scaffolding strategies to meet each student's proficiency and giving appropriate feedback for increased learning (Baumert et al., 2010; Goe, 2007; Hafsteinsdóttir et al., 2021; Vygotsky, 1978). Instructional clarity encompasses not only providing students with clear instructions and explanations but also offering each student with appropriate support and challenges, as well as linking new concepts to previously learnt knowledge (Nilsen & Gustafsson, 2016). Thus, presenting students with more demanding tasks is insufficient. Depending on each student's aptitude and cognitive ability, appropriate instructional support and scaffolding are also needed to promote understanding and learning (Vygotsky, 1978). This requires both general pedagogical knowledge (GPK) and pedagogical content knowledge (PCK), including subject-specific instructional strategies and didactics. Such instructional strategies vary between subjects. For example, mathematics instruction tends to be more direct and is characterised by procedural fluency, conceptual understanding, and problem-solving, compared to other subjects (Cohen et al., 2003; Lester, 2007). However, the degree of subject-specific instruction also varies across educational stages. In primary education, GPK plays a more prominent role. In Grade 4 mathematics classrooms, mathematics instructions are more generic and focused on fundamental concepts, classroom routines, and general teaching strategies. Whereas in an eighth-grade mathematics classroom, the increased complexity of the subject-specific content requires more specialised content knowledge (CK) and a subject-specific approach regarding the instructional methods. Consequently, instructional strategies at primary school could be considered generic and part of GPK, as they are applicable across multiple subjects.

## Defining academic self-concept

When examining factors influencing student learning outcomes, it is important to consider students' self-perceptions and self-beliefs, as these are significant predictors of academic achievement (Marsh & Craven, 2006; Shavelson et al., 1976; Stankov & Lee, 2017). Self-perceptions and self-beliefs are important psychological constructs that shape an individual's thoughts, behaviour, and emotions (Eccles & Wigfield, 2023). In social sciences research different constructs are used when discussing self-perceptions and self-beliefs. For example, in educational research constructs such as self-esteem, self-efficacy, self-confidence, and self-concept are often examined (Eccles & Wigfield, 2002; Marsh & Craven, 2006). The definition of self-esteem refers to an individual's overall evaluation of their worth and value (Rosenberg et al., 1995). The next construct, self-efficacy, denotes an individual's belief in their ability to succeed in a specific task that directly influences their motivation, effort, and persistence while actively engaging with the task (Bandura, 1997). The third definition, self-confidence, refers to a more general belief in one's ability to succeed in a certain domain (e.g., mathematics) Unlike self-efficacy, self-confidence is less task-specific and is influenced by recent success or failures, consequently making it fluctuate over time (Stankov et al., 2014).

The fourth construct, the self-concept, denotes an individual's perceptions and beliefs in their ability to succeed which is shaped by their previous experiences and interactions with peers (Eccles & Wigfield, 2023; Marsh & Craven, 2006; Shavelson et al., 1976). Self-concept has a hierarchical structure, with an overarching general self-concept divided into two domains: non-academic self-concept and academic self-concept. Non-academic self-concept includes social, emotional, and physical aspects, while academic self-concept is divided into subject-specific subdomains, such as mathematics, science, and history (Marsh & Craven, 2006; Marsh & Martin, 2011; Shavelson et al., 1976).

Academic self-concept (ASC) is a multidimensional concept, which has been conceptualised and operationalised in various ways and is often measured through subjective self-reports (Shavelson et al., 1976). In educational research, ASC is one of the strongest non-cognitive predictors of mathematics achievement (Hattie, 2009; Marsh et al., 2005; Stankov & Lee, 2017; Stankov et al., 2014). Additionally, ASC has been suggested as a tool for addressing social inequity in education, as promoting a positive ASC may contribute to narrowing achievement gaps (Marsh & Craven, 2006). Teaching quality has been found to be interrelated with students' cognitive appraisals, influencing both academic achievement and emotions related to achievement activities (Lazarides & Raufelder, 2021).

The relationship between academic self-concept (ASC) and academic achievement (ACH) is complex, and has been explored through three different models, each reflecting different perspectives, directions, and relations. The self-concept enhancement model proposes that the academic self-concept is a determinant of academic achievement (ASC → ACH), as students with high ASC are more likely to have higher academic achievement (Marsh, 1994). The skill development model suggests the reverse relationship, where academic achievement is a prerequisite for academic self-concept (ACH → ASC), meaning that high academic achievement improves students' ASC (Calsyn & Kenny, 1977). The reciprocal effects model suggests that academic achievement and ASC reinforce each other mutually in a reciprocal relationship (ACH ↔ ASC), creating a feedback loop in which students' academic self-concept shapes, and is shaped by, their academic achievement (Marsh et al., 2005). In the reciprocal effects model, environmental factors are acknowledged to influence this reciprocal relationship (Ramazan et al., 2023). Based on theory and empirical research, the reciprocal effects model was developed after methodological advances, such as structural equation modelling, that enabled deeper analyses of these interactions (Marsh et al., 2005). Students' self-concept in mathematics, is linked to the social context and the classroom interactions, which in turn influence students' achievement (Eccles & Wigfield, 2023; Marsh & Martin, 2011; Shavelson et al., 1976). This highlights the importance of both the individual student's perceptions as well as the collective classroom environment in shaping students' experiences and learning outcomes.

The TIMSS 2019 Questionnaire Framework (Mullis & Martin, 2017), reports that a subject-specific self-concept is measured using the 'student confidence in mathematics' scale in the student questionnaire. Previous TIMSS assessment cycles have "shown a strong relationship between students' academic self-concept and their achievement", as reflected in the 'student confidence in mathematics' scale (Hooper et al., 2017, p. 72). However, TIMSS does not provide a theoretical rationale for defining the scale as 'student confidence in mathematics' instead of 'mathematics self-concept', nor does TIMSS elaborate on conceptual foundations or the items within the scale (Eklof, 2006). Additionally, TIMSS offers no theoretical underpinning for how the scale is conceptualised and operationalised, or how it aligns with previous research on academic self-concept. In the TIMSS questionnaire, students assess their mathematics confidence reflected in their perceived ability to succeed which is based on prior experiences and comparisons with their peers (Hooper et al., 2017; Marsh & Craven, 2006). In this thesis, the definitions of academic self-concept and mathematics confidence will be used interchangeably to refer to the students' self-reported beliefs in their ability to succeed academically in mathematics.

## A conceptual model of teaching quality and the situated classroom context

Taken together, the theories and frameworks discussed in this chapter suggest that student learning outcomes are associated with both teaching quality and student background characteristics. However, existing frameworks have, to a lesser extent, addressed how the situated classroom context could relate to teaching quality when examining student learning and its variations. While the conceptual model of determinants of student outcomes by Nilsen and Gustafsson (2016), emphasises the various aspects of teacher and teaching quality at the classroom level, less attention is paid to the role of the situated classroom context and the nesting of students within classrooms. Teaching quality is not solely determined by the teacher but is co-constructed through the situated interactions between students and the teacher in the classroom. Students bring different background characteristics, aptitudes, and motivation to the classroom composition while engaging in the classroom processes, which include teacher-student interactions and students' affective interpretations of these classroom processes. This situated classroom context is highlighted in the SEVT model (Eccles & Wigfield, 2023). The SEVT model examines how individual student's aptitude and beliefs shape their motivation and academic achievement (Eccles & Wigfield, 2023). It includes the classroom interactions between the teacher and students, and between peers. Subsequently, by building on both the SEVT model and conceptual model of determinants of student outcomes (Nilsen and Gustafsson (2016), this thesis proposes an extended model for examining teaching quality and the situated classroom context. Thus, the *conceptual model of teaching quality and the situated classroom context* is introduced as a guiding framework for examining and understanding teaching quality as a classroom-specific construct and its relationship to student learning outcomes (see Fig. 3).

# Chapter 3 Literature review

To contextualise the thesis, this chapter presents previous research on the relationships between teaching quality and student learning outcomes. First, the literature review discusses research on the relationships between teaching quality, with the subdimensions of classroom management and instructional clarity, and mathematics achievement. The next section presents previous research on the relationships between teaching quality, students' mathematics confidence, and academic achievement. The last section addresses validity concerns associated with using questionnaire data in secondary analysis.

## Teaching quality and student learning outcomes

Teaching quality is a multidimensional construct lacking consensus within the research field on its operationalisations, conceptualisations, or theoretical framework (Nilsen & Gustafsson, 2016; Senden et al., 2022). As a result, research on the relations between teaching quality and student learning outcomes has reported inconsistent findings, with variations depending on the subject matter, country, student age, self-reports from students or teachers, and differences in how teaching quality is conceptualised and operationalised. Despite these inconsistencies, teaching quality is recognised as a key determinant of student learning outcomes (Darling-Hammond, 2021; Guerriero, 2017; Hattie, 2009). In a second-order meta-analysis, Hattie (2009) reported a positive relationship between student learning outcomes and teachers' behaviour and interactions in the classroom.

Several studies have examined the relationships between teaching quality and student learning outcomes drawing on PISA data. One such study, a longitudinal study of 194 German mathematics classrooms in PISA 2003, examined the relationships between students' mathematics achievement, mathematics motivation, teacher competence, and teaching quality (Kunter et al., 2013). Their study investigated various teacher-related factors, such as mathematical pedagogical content knowledge (PCK), enthusiasm, self-regulatory style, cognitive activation, and classroom management, and their associations with student-perceived support, enjoyment and achievement in mathematics. Kunter et al. (2013) employed a 1-year repeated-measures design to assess the effects of teaching quality on both academic achievement and motivation. Their findings showed that teachers with greater PCK and more enthusiasm provided better learning support and demonstrated more effective classroom management. They reported that these factors had a positive influence on both students' academic achievement and motivation.

Using the same German PISA 2003 data, along with the national follow-up in 2004 (*N* = 3,738), Atlay et al. (2019) examined the relationships between teaching quality and mathematics achievement. Teaching quality was measured through student-reported classroom management, cognitive activation, and supportive climate. The analysis controlled for prior mathematics achievement to account for baseline differences in student performance. In this study, Atlay et al. (2019) found a significant positive relationship between classroom management and mathematics achievement. Whereas no significant relationship was found for supportive climate or cognitive activation (Atlay et al., 2019). Contrary to their hypothesis, Atlay et al. (2019) found that the achievement gap increased, as students with high socioeconomic status (SES) benefitted more from cognitive activation and a supportive climate than the students with lower SES. This finding aligns with a Swedish study showing that high-SES students benefitted more from teachers' mathematics instructions, thereby contributing to widening the Swedish achievement gap (Hansson, 2012).

Teacher quality and teaching quality have been the focus of several studies based on TIMSS data. Blömeke et al. (2016) used TIMSS 2011 data from 205,515 students in 10,059 fourth-grade classrooms in 47 countries to investigate the relationships between teacher quality, teaching quality, and mathematics achievement. Teacher quality was measured through formal teacher certification, professional development activities, and a sense of preparedness. Teaching quality was measured using six items in the teacher questionnaire, capturing clear instruction, cognitive activation, and supportive climate. Blömeke et al. (2016) showed that teacher quality, specifically teacher education and specialisation, was significantly related to both teaching quality and student achievement. However, there was substantial variation across countries. Blömeke et al. (2016) reported that students with higher mathematics achievement were generally taught by more experienced teachers, and these teachers had higher instructional quality across countries. However, the directions of these relationships varied, with some countries showing significant effects in the opposite direction (Blömeke et al., 2016, p. 38).

Similarly, using TIMSS 2011 Grade 8 data, Nilsen and Bergem (2020) investigated how teacher competence and teacher education related to educational equity in Finland, Norway, and Sweden. Their study focused on teacher competence and instructional quality, measured by six indicators from the teacher questionnaire reflecting the subdimensions of instructional clarity, cognitive activation, and supportive climate. Their findings were consistent with previous research (e.g., Blömeke & Olsen, 2019; Nilsen et al., 2018), showing a significant relationship between teaching quality and mathematics achievement in some Nordic countries. However, teacher competence showed varying relationships with equity. In Finland, the impact of SES was of greater significance in classrooms where teachers specialised in mathematics and reported having high instructional quality. Similarly, in Sweden, high teacher competence increased the impact of SES on achievement. In contrast, in Norway, teacher competence reduced the impact of SES on mathematics achievement. Nilsen and Bergem (2020) concluded that, in their study of TIMSS 2011, teacher competence contributed to increasing the mathematics achievement gap between low- and high-SES students. They explained that competent teachers were unequally distributed between low- and high-SES classrooms. However, this pattern was not found in Norway, as teacher competence seemed to mitigate educational inequities related to SES. Additionally, Nilsen and Bergem (2020) discussed construct under-representation in instructional quality, as their study used six items from the teacher questionnaire, two items for each subdimension (instructional clarity, cognitive activation, and supportive climate).

## Classroom management and student learning outcomes

Classroom management refers to the structural-organisational activities undertaken by the teacher to instruct and engage students while establishing rules and maintaining order within the classroom (Emmer & Stough, 2001). Effective classroom management requires teachers to anticipate and navigate the complex social interactions and the multitude of processes occurring in the classroom (Marzano et al., 2003). More precisely, to support student learning, teachers must adapt their structural-organisational strategies, instructional methods, and scaffolding techniques to align with each student's proficiency level at a given time (Baumert et al., 2010; Goe, 2007; Hafsteinsdóttir et al., 2021; Vygotsky, 1978). The more time a student actively engages in learning activities without disruptions, the greater their opportunity to learn and achieve (Guerriero, 2017).

Meta-analyses have shown that classroom management has a substantial effect on student achievement. A meta-analysis of 100 studies found that strong teacher-student relationships served as the foundation for effective classroom management (Marzano et al., 2003). Similarly, a meta-analysis of 47 studies, exploring the effects of classroom management and students' academic and social-emotional learning concluded that classroom management is a prerequisite for students' learning opportunities (Korpershoek et al., 2016). Their meta-analysis reported that while classroom management was positively and significantly related to students' academic learning outcomes, it did not significantly relate to students' motivation. In a second-order meta-analysis by Hattie (2009), classroom management had an effect size of 0.35 on achievement, slightly below the 0.40 threshold considered to indicate a particularly strong impact (Hattie, 2009). Meanwhile, a meta-analysis by Wang et al. (1993) found that the student learning outcomes were not only influenced by students' cognitive ability, motivation, and home environment but also by teacher-student interactions and classroom management. Their findings, aligned with the Carroll (1963) model, emphasising that the more time students can focus on learning tasks without disruptions, the greater their academic achievement (Wang et al., 1993). Beyond academic achievement, it has been suggested that effective classroom management may mitigate negative classroom dynamics. For instance, effective classroom management has been shown to act as a buffer against bullying (Gomes et al., 2020; Rutkowski, L., & Rutkowski, 2016).

Previous research based on TIMSS data has found positive relationships between classroom management and mathematics achievement. A study drawing on TIMSS 2015 from Norway, Germany, and the Flanders region of Belgium, investigated the relationships between instructional quality and mathematics achievement (Bellens et al., 2019). Instructional quality was measured through classroom management (5 items), cognitive activation (5 items), and supportive climate (5 items in Germany and Flanders, but 10 items in Norway). Their study also accounted for student background factors, including the number of books at home as a proxy for SES and language spoken at home as an indicator of ethnicity. Bellens et al. (2019) found significant and positive relationships between classroom management and mathematics achievement across all three countries. Similarly, a longitudinal study using TIMSS data from German secondary mathematics classrooms followed students through two follow-ups in Grades 7 and 8 (Kunter et al., 2007). This study examined the relationships between mathematics interest, and classroom management, measured by rule clarity and monitoring (Kunter et al., 2007). Contrary to their expectations, Kunter et al. (2007) found significant positive relationships between classroom management and mathematics achievement at the individual level but not at the classroom level.

## Instructional clarity and student learning outcomes

Teachers' instructional clarity is a multidimensional construct referring to how efficiently teachers use pedagogical techniques to provide students with clear instructions, appropriate support, and challenges while linking new concepts to prior knowledge (Nilsen & Gustafsson, 2016). Previous research has conceptualised and operationalised instructional clarity in various ways (Titsworth et al., 2015). In TIMSS, instructional clarity was first introduced as a distinct scale in the 2019 assessment. However, in earlier TIMSS assessments, the supportive climate scale included indicators related to instructional clarity (Nilsen & Gustafsson, 2016).

Meta-analyses have shown that instructional clarity has a substantial influence on student achievement. In a second-order meta-analysis, Hattie (2009) found that dimensions of instructional clarity had a large effect on students' academic learning, well above the suggested threshold of 0.40. For example, teacher clarity had an effect size of 0.75 on academic achievement. When teachers provided students with appropriate challenges or support through scaffolding, the effect size increased to 0.82. Furthermore, when teachers integrated new instruction with students' prior knowledge, the effect size reached 0.93 (Hattie, 2009). Similarly, two meta-analyses exploring the relationship between teacher clarity and student learning found that teacher clarity had a stronger relationship to students' affective learning than their cognitive learning (Titsworth et al., 2015). However, Titsworth et al. (2015) cautioned that these findings may be subject to positive bias, as clarity is often assumed to be positively associated with achievement. Additionally, their findings suggest that some students may feel annoyed or frustrated by excessive instructions, and in these cases, instructional clarity could negatively influence students' affective outcomes (Titsworth et al., 2015).

Previous research using TIMSS data has found positive relationships between instructional clarity and mathematics achievement. A study of Swedish TIMSS 2003 Grade 8 data found that mathematics achievement was higher when teachers provided appropriate scaffolding and support (Hansson, 2012). This study emphasised the importance of scaffolding, particularly for second language learners, highlighting the need for teachers to offer targeted support.

A more recent study, using data from all participating European countries in TIMSS 2015, found that student-perceived instructional clarity was positively associated with students' intrinsic motivation (Konstantinidou & Kyriakides, 2022). Similarly, Blömeke and Olsen (2019) analysed TIMSS 2011 data from students in Grades 4 and 8 across five countries (England, Norway, South Korea, Thailand, and Tunisia) to examine the relationships between instructional clarity, academic achievement, and non-cognitive outcomes. Their study found that instructional clarity had a direct positive effect on students' enjoyment of learning mathematics and science.

## Mathematics confidence, teaching quality, and student learning outcomes

Student's academic self-concept is one of the strongest non-cognitive predictors of academic achievement (Hattie, 2009; Marsh et al., 2005; Stankov & Lee, 2017; Stankov et al., 2014). Previous research has demonstrated a reciprocal relationship between academic self-concept and academic achievement (e.g., Creemers & Kyriakides, 2007; Marsh & Craven, 2006; Marsh et al., 2005). A prominent example of this reciprocal relationship was found in a longitudinal study by Marsh et al. (2005), involving 5,659 German 7th graders. Their findings showed that mathematics interest and achievement were not only strongly influenced by academic self-concept, but also reciprocally reinforced each other. Marsh et al. (2005) concluded that "the most effective strategy is to improve academic self-concept, interest, and academic achievement simultaneously" (Marsh et al., 2005, p. 413). As a result, academic self-concept functions both as an outcome variable and as a mediating variable, shaping and explaining academic achievement (Chen & Lu, 2022; Marsh & Craven, 2006).

Students' general self-confidence and academic achievement improve when they have positive social relationships with their teachers, and when their learning needs are effectively met (Fan & Williams, 2018; Hattie & Yates, 2013; Ryan & Deci, 2000). Conversely, students' low academic self-concept has been linked to lower academic achievement (e.g., Lazarides & Ittel, 2012; Liu & Meng, 2010). Students with low academic self-confidence and who perceive themselves as having limited ability to succeed are more likely to feel demotivated, which may hinder their academic performance (Legault et al., 2006). Previous research has found that self-confidence and self-esteem are generally high in younger students but tend to decline with age (Orth et al., 2012). Additionally, gender differences have been found, with boys tending to rate their abilities higher than girls (Orth et al., 2012).

Research investigating the relationship between student-perceived academic achievement and affective outcomes, such as motivation and self-concept, has identified a cross-level paradox (Shen & Pedulla, 2000; Shen & Tam, 2008). At the individual level, a positive relationship is observed between academic self-concept and academic achievement. However, at the next level, such as classroom or school, a negative relationship emerges between academic self-concept and academic achievement (Chen & Hastedt, 2022; Marsh et al., 2019). This paradoxical negative relationship at the higher level may occur when students attend high-achieving schools and perceive themselves as having a lower self-concept compared to students of equal ability attending low-achieving schools. This paradox is referred to as the Big-Fish-Little-Pond Effect (Marsh et al., 2019). A TIMSS 2015 study, involving students in Grades 4 and 8 across 57 countries, investigated this paradox by examining the non-cognitive scales of academic confidence, enjoyment of learning, and achievement outcomes in mathematics and science (Chen & Hastedt, 2022). To ensure the validity of their findings, the researchers tested for measurement invariance across the educational systems. They confirmed the non-cognitive scales to be measurement invariant and consistent across countries. However, Chen and Hastedt (2022) highlighted inconsistencies in the operationalisations of the construct of academic self-concept in previous research into the Big-Fish-Little-Pond Effect. They argued that establishing a consensus on which items to include in such studies is vital for facilitating meaningful comparisons between studies. A recent study of Swedish 15-year-old students ($N$ = 24,771) showed that students in schools with higher average achievement reported lower self-concept in mathematics and language compared to those in schools with lower average achievement, even when individual achievement levels were comparable (Klapp et al., 2025). The study also found that the cognitive ability of the individual student moderated the negative Big-Fish-Little-Pond Effect in the mathematics domain, students with higher cognitive abilities were less affected than their peers with lower cognitive abilities.

Teachers' instructional support has been identified as a significant contributor to students' mathematics self-concept (Ramazan et al., 2023). Analysing data from the USA PISA 2012 sample, Ramazan et al. (2023) examined how the relationship between instructional support and mathematics self-concept varied depending on students' primary language spoken at home. Students were grouped into two categories, those who spoke English at home (86%) and those who spoke Spanish or other languages at home (14%). Ramazan et al. (2023) found that for students in the 'other-language-at-home' group, the relationship between teachers' instructional support and mathematics self-concept was nonsignificant. Additionally, students in this 'other-language-at-home' group who perceived higher levels of instructional support from their teachers had lower mathematics achievement. Ramazan et al. (2023) concluded that students' mathematics self-concept was reinforced when students perceived that their teachers liked them, treated them fairly, and created safe and orderly classrooms. This finding aligns with previous research by Lazarides and Raufelder (2021), which reported classroom management and teacher-student relations as significant predictors of mathematics self-concept.

A recent study drawing on TIMSS 2019 data from Grade 8 in England and Hong Kong explored the relationships between classroom management, instructional clarity and three outcomes: achievement, boredom, and enjoyment in mathematics (Chen & Lu, 2022). The study also examined the mediating role of mathematics self-concept and students' value of mathematics. The sample included only eighth-grade mathematics students whose parents were born in Hong Kong ($N$ = 1,146) and in England ($N$ = 1,981). Chen and Lu (2022) found that instructional clarity was significantly related to mathematics self-concept for students in both Hong Kong and England. Additionally, their findings revealed significant indirect relationships between instructional clarity and mathematics achievement, mediated through academic emotions, in both countries, although stronger in Hong Kong (Chen & Lu, 2022). However, their study also reported a significant negative direct effect between instructional clarity and achievement for students in England, while this relationship was nonsignificant for students in Hong Kong (Chen & Lu, 2022). This finding aligns with previous research (Ramazan et al., 2023; Titsworth et al., 2015). Chen and Lu (2022, pp. 8-9) explained that "excessively extensive or comprehensive instruction may restrict students' opportunities for individual discovery, diminish their motivation to learn, and ultimately compromise their academic performance". Their study also found that classroom management was positively associated with mathematics achievement for students in England. By contrast, in Hong Kong, a negative perception of classroom management was associated with lower self-concept, a relationship not observed in England (Chen & Lu, 2022). Although their study accounted for students' background variables (gender, books as a proxy for SES, and age), the authors chose not to present these estimates, citing clarity as a reason.

Together, the studies discussed in this chapter highlight the challenges in conceptualising and operationalising latent constructs. The variation in indicators used to define academic self-concept across different studies complicates direct comparisons and limits the generalisability of the findings (Liu & Meng, 2010).

## Questionnaire data and validity

Self-reported survey data is widely used in research, both in questionnaires constructed by the researcher and in secondary analysis of existing data. To ensure valid, reliable, and unbiased research, questionnaire items must demonstrate high validity and reliability (Messick, 1989; Spector, 1992). However, the way the questionnaire items are constructed can introduce method effects and response biases, that could influence the accuracy of the findings (DiStefano & Motl, 2006).

### Likert scales

International Large-Scale Assessment (ILSA) questionnaires, such as TIMSS, use rating scales where respondents indicate the extent to which they agree or disagree with the statements. A commonly used format in questionnaires is the Likert scale. This scale typically features multiple response points, which are presented in either ascending or descending order. Previous research has indicated a left side selection bias, leading to a response-order effect, which may inflate responses on the left-side (Chyung et al., 2018). To mitigate this issue, using ascending scales has been suggested as a more reliable approach (Chyung et al., 2018). TIMSS employs various types of Likert scales. Some consist of four response options (e.g., 1 = I agree a lot, 2 = I agree a little, 3 = I disagree a little, and 4 = I disagree a lot). Others include a neutral midpoint (3 = I neither agree nor disagree), while some use a reversed continuum, where items range from both disagree to agree depending on the wording of each item (Hooper et al., 2013).

### Mixed-worded scales

Questionnaires can include mixed-worded scales to ensure the construct validity of latent variables. Such scales are suggested to maintain the respondents' attention and engagement while reducing acquiescence response bias, where respondents tend to consistently agree with statements without careful consideration (Marsh, 1994; Roszkowski & Soven, 2010). There are four ways in which a mixed-worded scale can be presented: a direct positive mode ('Mathematics is fun'), a negated positive mode by adding a negative word ('Mathematics is not fun'), a direct negative mode ('Mathematics is boring'), and a negated negative mode ('Mathematics is not boring') (Colston, 1999).

In a study examining the reverse directional items in the TIMSS 2011 scale of 'students confident in mathematics', Hooper et al. (2013) found that items phrased in opposite directions of the scale showed different psychometric properties compared to those aligned with the scale. This finding was stronger for Grade 4 students than for Grade 8 students. Similarly, Michaelides (2019) analysed the same dataset but focused on fourth-grade students. Findings showed signs of bias as students with lower linguistic ability failed to interact consistently with the negatively worded items that had a reverse direction (e.g. mathematics is not boring). Additionally, these items were found to be non-metric invariant across countries, indicating that students did not respond to the meaning of the items consistently (Michaelides, 2019). In secondary analyses involving scales with items of opposite directions the items are recoded to ensure consistency in direction. In this process, items expressing unfavourable agreement are reversed, so that positive agreement receives a high value on the Likert scale (e.g., 4) and negative agreement a low value on the Likert scale (e.g., 1) (van Sonderen et al., 2013).

Questionnaires using mixed-worded scales sometimes show inconsistencies in respondents' answers due to various factors. Respondents could be influenced by social desirability, making them agree with positively phrased statements, regardless of their true opinion (Spector, 1992). For example, when engaging with a statement such as 'I am good at mathematics, respondents may agree because it is a socially desirable response. However, when presented with negatively worded items, respondents may give untruthful answers due to various factors. For instance, respondents may misunderstand or not read the statement thoroughly, experience fatigue, or fail to interpret the item correctly (Merritt, 2012; Steinmann, Sánchez, et al., 2022; van Sonderen et al., 2013). Additionally, negative wording may increase the linguistic complexity of the item, making the item more difficult to interpret and more cognitively demanding (Michaelides, 2019; Wang et al., 2018). This issue is particularly relevant for younger students or those with limited reading or language proficiency, as they could experience negatively worded items as more cognitively demanding (Marsh, 1994; Steinmann, Strietholt, et al., 2022; Van Dam et al., 2012). Given this, it has been suggested that negatively worded items "should not be used in secondary data analyses" (Marsh et al., 2013, p. 124).

Supporting this concern, a study on measurement invariance of attitudinal scales in TIMSS 2019, comprising 58 countries in Grade 4 and 39 countries in Grade 8 (Reynolds et al., 2022), found that the non-invariance observed in some items belonging to attitudinal scales was related to negatively worded and reverse coded items. Additionally, Reynolds et al. (2022) suggested that cultural differences in how teachers use praise to shape student behaviours may contribute to the observed non-invariance in some mixed-worded attitudinal scales.

Previous research comparing items in mixed-worded scales has shown that negatively worded items can reduce the scale's internal consistency. These items often produce factor loadings below 0.4, a threshold below which items are considered unreliable for measuring latent constructs (Chiu, 2008). As a result, combining positively and negatively worded items may complicate the factorial structure of an instrument and introduce method effects (Michaelides, 2019; Wang et al., 2018). Some studies suggest that removing negatively worded items can improve the scale's reliability and internal consistency (Roszkowski & Soven, 2010; Ye & Wallace, 2014). For example, when removing negatively worded items, the internal consistency was improved for the scale 'psychological sense of school membership' (PSSM) scale (Ye and Wallace, 2014). In addition, removing negatively worded items may help avoid construct-irrelevant multidimensionality and reduce method effects (Scherer & Nilsen, 2016). Other strategies include separating positively and negatively worded items into two separate sets (Roszkowski & Soven, 2010; van Sonderen et al., 2013), or using a transformed index consisting of three levels (e.g., Berger et al., 2020). However, some researchers choose to include some or all scale items in the analyses, regardless of their wording direction (e.g., Eklof, 2007). If scale items are not properly recoded and aligned, mixed-worded scales may produce biased results, leading to method effects which may compromise the validity of the analysis (Spector, 1992).

## Method effects and response bias

Method effects and response bias are well-known threats to construct validity in questionnaire-based research (Messick, 1989). Method effects arise when respondents react to the format of the instrument, rather than the underlying construct it is intended to measure. Response bias refers to respondents answering items systematically based on factors such as cognitive complexity, social desirability, or method issues of the questionnaire (Messick, 1989). When positively and negatively worded items within a scale are found to measure different constructs, they can introduce method effects and response bias, and thus, threaten construct validity, (Marsh, 1994; Scherer et al., 2016; van Sonderen et al., 2013) which could lead to construct-irrelevant variance (Messick, 1989). Such variance is unrelated to the intended latent construct and reduces the factorial and statistical validity of a scale (DiStefano & Motl, 2006), and may challenge the internal consistency reliability of a scale (Roszkowski & Soven, 2010; Ye & Wallace, 2014).

Hooper et al. (2013) concluded that reverse directional items in the TIMSS 2011 scale 'students confident in mathematics' introduced construct-irrelevant variance. Their analysis of this mixed-worded scale showed that simply reverse-coding negatively worded items was insufficient to render them equivalent to the positively worded items. Consequently, the inclusion of such items in secondary analysis may threaten construct validity as it could introduce construct-irrelevant variance (Messick, 1989).

Taken together, previous research highlights the presence of method effect and response bias with mixed-worded scales. If secondary analyses drawing on ILSA questionnaire data do not account for these issues, the construct validity of the scale could be compromised. Despite this risk, few secondary studies using ILSA data address the construct validity of the indicators used in their analyses. Little attention has been given to how item phrasing, reverse item direction, or translation differences may introduce method effects and influence the relationships found in the analysis (Rosén & Nilsen, 2024).

# Chapter 4 The cultural context

The following chapter presents the cultural context of the thesis. Sweden is the primary focus of Studies I and II. In Study III, the scope expands to include four Nordic countries: Denmark, Finland, Norway, and Sweden.

## The Nordic context

The Nordic countries comprise Denmark, Finland, Iceland, Norway, and Sweden, and additional autonomous regions. Åland is an autonomous region of Finland and Greenland and Faroe Islands are autonomous territories of Denmark (Nordic Co-operation, nd). The Nordic Council serves as the official body for inter-parliamentary co-operation across the Nordic nations. The populations of these countries are relatively small. Sweden is the most populated country with approximately 10.5 million inhabitants, followed by Denmark and Norway each with around 6 million. Finland has a population of roughly 5.5 million, and Iceland is the smallest country, with around 400,000 inhabitants (Nordic Co-operation, nd).

Ages for starting school differ slightly across the Nordic countries. While all Nordic children begin either preschool or comprehensive compulsory education (school) at the age of six, an important distinction exists. In Norway and Iceland, children start school the year they turn six. In contrast, children in Denmark, Finland, and Sweden begin a mandatory preschool class at the age of six and start school the following year. As a result, when participating in TIMSS, Norway samples Grade 5 students, while the three other Nordic countries sample students in Grade 4 (Nordic Co-operation, nd).

Teacher education and certification in the Nordic countries vary in structure and specialisation. Currently, Finland, Norway and Sweden use a generalist approach, meaning that a generalist class teacher is responsible for multiple subjects during the first three years in primary school (Grades 1-3) (Volmari, 2019). In Sweden, class teachers for Grades 1-3 are certified to teach Swedish, mathematics, English, social science, and science/technology. For Grades 4-6, Swedish class teachers are certified in four subject matters (Swedish, mathematics, English, and one of the eligible subjects of PE, music, social science, or science/technology). In contrast, Danish primary students are taught by semispecialised subject teachers (Volmari, 2019). A notable distinction is that Danish primary teacher education is run by vocational colleges, unlike the university-based programmes in the other Nordic countries (SOU 2024:81). In secondary school, all Nordic countries use specialised subject teachers for the three final years of compulsory education (Volmari, 2019).

The Nordic countries have many similarities, such as a common history, shared political, cultural, and societal similarities. Additionally, these countries have common values around welfare, democracy, equity, openness, and gender equality (Lundahl, 2016; Nordic Co-operation, nd; Reimer et al., 2018). The common core values are reflected in the respective educational systems through what is often referred to as the 'Nordic Model' (Reimer et al., 2018). The Nordic Model, also known as 'A School for All', represents equal opportunities, fairness, inclusion, and equity in education (Blossing et al., 2014; Frønes et al., 2021; Lundahl, 2016). Nordic educational values are also considered child-centred, with a strong emphasis on meeting the needs of every child or student (Andersen, 2010). The Nordic, non-streamed, comprehensive school system is based on egalitarian principles, ensuring equal educational opportunities for all students (Blossing et al., 2014). This means that in all Nordic countries, each student, regardless of their background, should have access to high-quality education (Blossing et al., 2014). In line with these values, education is publicly funded and free of charge across all Nordic countries.

At the same time, differences exist across the Nordic countries in areas such as language, climate, and educational policies. Although the Nordic countries share many values, their educational systems show national variations. Such variations include differences in policy documents, curricula, teacher education, and educational reforms. During the 1990s, all Nordic countries implemented educational reforms relating to decentralisation, funding, and parental school choice, but the extent and impact of these reforms varied between the countries (Lundahl, 2016). While education is publicly funded, there are independent organisations and companies running schools to varying extents (Frønes et al., 2021). In Sweden, the marketisation of the educational system has gone further than in other Nordic countries. Profit-making companies are allowed to operate publicly funded schools to a larger extent in Sweden (Lundahl, 2016). This shift has contributed to growing differences in student performance at the school level in Sweden (Yang Hansen & Gustafsson, 2019).

Differences in student performance also appear within public school systems. In Finland, there is evidence of classroom variation in performance (Yang Hansen et al., 2014). Although Finland is known for high-quality education and educational equity, an early indirect streaming system exists that is not apparent at first glance (Lundahl, 2016). Mainly in urban areas, parents can choose specialised selective classes when their child begins Grades 1, 3 or 7 (Koivuhovi et al., 2022). These specialised classes focus either on academic subjects (e.g., mathematics; language) or non-academic subjects (e.g., music; sports) and students are selected via aptitude tests (Kosunen et al., 2020). This ability-based differentiation with students gathered in 'classes with emphasis', leads to a segregation with a 'cream-skimming effect', resulting in increased between-class differences (Bernelius & Vaattovaara, 2016).

## The Swedish educational context

In Sweden, schooling is compulsory from the age of six until the end of Grade 9, typically when students are 15 to 16 years old (SFS 2010:800). After completing compulsory school, students can apply for a non-compulsory three-year upper secondary education programme, which can be either academic or vocational in orientation.

### The Swedish educational context and reforms in the 1990s

Until the late 1980s, Swedish schools were governed centrally by the state but during the 1990s, the Swedish educational system underwent a transformative decade with several reforms that completely changed its governance, funding, and structure (Sundberg, 2005). In 1991, reforms introduced decentralisation and deregularisation, transferring responsibility for schooling from the national government to Sweden's 290 municipalities. As a result, the state was no longer in charge, instead, municipalities became responsible for providing students with equal education and equal access to it. In 1992, several key reforms were implemented. The 'independent school reform' allowed private actors to establish and run schools alongside the municipal schools. The independent schools could be operated by either private, profit-making companies or non-profit organisations and foundations (Skolverket, 2014). Like municipal schools, independent schools were tax-funded and free of charge for students. The same year, in 1992, the so-called 'school choice reform' was introduced, granting parents the right to choose which school their children would attend, regardless of whether it was administered by municipal or independent school authorities. A school authority is defined as a local school entity responsible for providing the attending students with an equitable, high-quality education, as stipulated in the Swedish Education Act (SFS 2010:800). Another reform of 1992 was the 'school voucher system reform', in which public educational funding followed the student to the school of their choice (Blossing et al., 2014). In 1994, a new curriculum and grading system were implemented (Sundberg, 2021).

The reforms introduced in the 1990s led to the marketisation process of the Swedish school system, resulting in increased competition between schools, both municipal and independent, to attract students through parental choice (Fjellman, 2019; Holmlund et al., 2014). As a result of these changes, the proportion of students attending independent schools has grown significantly. In 2022, 17.5% of students in compulsory education and 35.8% of students in upper secondary education attended independent schools (Skolverket, 2023a). During the academic year of 2023/2024, Swedish compulsory education was provided by 3,830 municipal school authorities across 290 municipalities and 816 school authorities operated by 536 independent school organisations (Skolverket, 2023a). It is important to note that, in the Swedish context, there is a difference between independent schools and private schools. Both independent schools and municipal schools are publicly funded and free of charge for students. In contrast, Swedish private schools are financed through student tuition fees. However, such private schools are rare, with fewer than a handful operating across Sweden.

## Manifested inequalities in Sweden

Swedish governmental reports and studies have highlighted growing inequities within the educational system, including increasing demographical segregation, sorting of students by family background, and widening achievement gaps (Gustafsson et al., 2016; Holmlund et al., 2014; SOU 2017:35; SOU 2019:40). These inequities are manifested in different ways. For instance, the academic achievement gap is widening as there is an increasing number of students who leave compulsory school without meeting the eligibility requirements for upper-secondary education. In 2023, 16.3% of the students did not meet the minimum grade requirements (Skolverket, 2024a).[2] There is also a gender gap in academic

---

[2] According to the minimum requirements for admission to upper-secondary education in Sweden, students must have at least an E grade in Swedish, English, mathematics. In addition, vocational programmes require passing grades in five other subjects, while university-preparatory programmes require passing grades in nine additional subjects.

achievement, manifested in the Grade Point Averages (GPA[3]). In 2023, the average GPA at the end of compulsory schooling was 236.4 points for girls and 219 points for boys (Skolverket, 2024a).

Further inequities in the Swedish school system are manifested in both geographical and pedagogical segregation. Students in rural areas often have limited school choices, as most independent schools are concentrated in the three largest urban areas in Sweden (Fjellman, 2019). This geographical distribution correlates with a socioeconomic divide. Students with higher levels of SES are typically concentrated in larger urban areas, while students with lower levels of SES are more commonly concentrated in rural areas (SOU 2019:40). Pedagogical segregation is reflected in the unequal distribution of qualified teachers (Hansson & Gustafsson, 2016). Schools differ regarding the proportion of certified teachers with appropriate subject qualifications (Skolverket, 2019). In 2022 in Swedish compulsory education, municipal schools employed an average of 72% qualified teachers, compared to 62% in independent schools (Skolverket, 2022b). Particularly concerning is the subject 'Swedish as a second language', which has the lowest proportion of qualified teachers at both compulsory and upper-secondary levels (Skolverket, 2022b). This is notable given that students who study Swedish as a second language belong to the group of students whose performance in ILSAs has declined (OECD, 2023).

Finally, one important factor threatening educational equity to take into consideration is problematic school attendance, absenteeism, and school refusal (Kearney et al., 2019; SOU 2016:94). Previous research has shown that increased absenteeism is associated with students' lack of motivation and perceptions of a hostile school environment (Gren Landell, 2021; Kearney et al., 2019). Similarly, students with low academic motivation are at greater risk of school absenteeism and dropout (Balkis, 2018). A report from the Swedish School Inspectorate concluded that the rise in problematic school attendance was linked to students' negative perceptions of schools, including poor teacher relationships, a lack of appropriate scaffolding and challenges, and feelings of unsafety (Skolinspektionen, 2016; Skolverket, 2021). The inequalities and widening achievement gaps in the

3 To clarify, in Sweden, students in Grade 9 receive final grades in 16 subject matters (17 if the student has studied a modern language which is an elective option). The grading system is currently letter-based: A is the highest grade (equivalent to 20 points), followed by B (17.5 points), C (15 points), D (12.5 points), E (10 points), and F, which is a failing grade. Eligibility for upper-secondary school is determined by a merit score, calculated from the total number of points from the student's 16 best subjects. This means a student who took 16 subjects can receive a maximum of 320 points.

Swedish educational system discussed throughout this section have been suggested to be related to, at least in part, the educational reforms implemented in the 1990s (SOU 2019:40).

## The Nordic counties and international large-scale assessments

All Nordic countries have consistently participated in several international large-scale assessments (ILSAs) such as TIMSS (Trends in Mathematics and Science Study), PIRLS (Progress in International Reading Literacy Study), and PISA (Programme for Internal Student Assessment). TIMSS, launched in 1995, assesses students in Grades 4 and 8 in mathematics and science every four years. PIRLS, introduced in 2001, assesses fourth-grade students' reading ability every five years. PISA, launched in 2000, evaluates 15-year-old students' performance in mathematics, science, and reading every three years. All Nordic countries have participated in PISA since the start in 2000. In TIMSS, only Sweden and Norway have participated with Grade 8 since 1995. Since 2011, all Nordic countries, except Iceland, have participated in TIMSS Grade 4.

Nordic countries generally score above the international mean in ILSAs in mathematics. These Nordic countries are characterised by high educational equity as students' socioeconomic status (SES) has a relatively small impact on performance (Mittal et al., 2021). High educational equity implies equal learning opportunities for all students, regardless of their SES, ethnicity, home language, gender, or cultural capital (Bourdieu, 1986). Although the Nordic countries have good quality education, they have, as other countries participating in PISA, experienced a decline in results (OECD, 2023) and widening achievement gaps with increasing educational inequities related to students' SES (Kavli, 2018). Norway, for instance, is among the countries in the world where a student's SES has a low impact on student achievement, but at the same time, Norway has fewer high-achieving students (Mittal et al., 2021). In Denmark, there is a strong influence of students' SES on achievement (Buchholtz et al., 2020). In Sweden, there are larger differences between schools and classrooms than in the other Nordic countries (Reimer et al., 2018; Skolverket, 2018). Finland is among the top-performing countries in ILSA and most equitable countries in the world with small differences between schools. For years, international educational policymakers have looked at Finland for inspiration and guidance on how to improve educational performance and equity (Elliott et al., 2019). Although Finland continues to be among the top-achieving countries worldwide, it has experienced a decline in ILSA scores since 2006 (Skolverket, 2024b).

According to the Swedish National Agency for Education, SES-related differences in educational outcomes are more pronounced in Sweden than in any of the Nordic countries participating in ILSAs (Skolverket, 2023b). The PIRLS 2021 study showed a decline in reading literacy among students who do not speak Swedish at home, whereas students who always speak Swedish at home remained at the same level (Skolverket, 2023b). Similarly, in PISA 2022, students who reported not speaking the test language at home scored lower than peers who did, 29 points lower in Finland and 27 points lower in Sweden, relative to the national average achievement levels (OECD, 2023).

## The Nordic countries and TIMSS

The Nordic countries have performed above the international average in TIMSS, although there are variations between countries and across cycles. In mathematics, fourth-grade students in Denmark and Finland have outperformed their peers in Sweden and Norway (Kavli, 2018). Swedish fourth-grade students in mathematics have scored higher than Norwegian students. However, this comparison must be interpreted with caution, as Norwegian students were one year younger than their Nordic peers in TIMSS cycles from 1995 to 2015. To address the discrepancy of age and years of schooling, Norway began sampling Grade 5 students from TIMSS 2015 onwards, (Kavli, 2018). Comparing the results in mathematics performance between 2011 and 2015, there has been an improvement in Denmark and Sweden, but a decline in Finland and Norway (Reimer et al., 2018). Sweden and Norway are the only Nordic countries to have participated with Grade 8 students since TIMSS began in 1995. Both countries experienced a decline in mathematics performance from 1995, followed by an ongoing improvement from 2002 in Norway and from 2011 in Sweden (Reimer et al., 2018). In science at Grade 4, Finnish students are at the top among the Nordic countries, followed by Sweden, Denmark, and Norway (Kavli, 2018). However, despite Finland's strong overall performance, TIMSS 2015 revealed that Finnish fourth-grade students had the largest achievement gap related to SES among the Nordic countries (Mittal et al., 2021).

Recent TIMSS cycles confirm that Nordic countries continue to perform above the international average, though there are some variations. In TIMSS 2019, Finland achieved the highest average score among the four participating Nordic countries (Skolverket, 2020). All four countries scored above the TIMSS scale centrepoint of 500. In TIMSS 2023, mathematics achievement among the Nordic students in Grade 4 remained relatively stable. All countries, except Denmark, continued to perform above the OECD average (Skolverket, 2024b). Beyond overall performance, studies have examined the relationship between teaching quality, SES, and academic achievement. Drawing on TIMSS 2011 Grade 8 data from Finland, Norway, and Sweden, Nilsen and Bergem (2020) found that teachers with higher qualifications were more likely to work in high-SES schools. In Finland, the relationship between SES and achievement was stronger at the classroom level than at the school level. Similarly, a study using TIMSS 2015 Grade 8 data from Norway and Sweden, found that teaching quality, measured through teachers' specialisation and professional development, mediated the influence of SES on science achievement in Sweden, but not Norway (Nilsen et al., 2021).

## Sweden and TIMSS

Despite improved mathematics performance in TIMSS, Sweden continues to face growing achievement gaps. According to the Swedish National Agency for Education, results from TIMSS 2023 show that Swedish fourth-grade students achieved the highest average mathematics performance since Sweden began participating in TIMSS in 2007 (Skolverket, 2024b). However, despite this improvement, the achievement gap continues to widen. In the mathematics assessment in TIMSS 2023, fourth-grade students from high SES backgrounds outperformed their low SES peers by an average of 82 points (Skolverket, 2024b). Similarly, students with at least one custodian of Swedish origin had an average of 540 points and thus outperformed their peers with two custodians born outside Sweden with about 40 to 45 points in TIMSS Grade 4 (Skolverket, 2024b). These results align with previous research indicating widening achievement gaps and growing educational inequities in Sweden (Gustafsson et al., 2013; Skolverket, 2018; Yang Hansen & Gustafsson, 2019).

Pedagogical segregation has also been shown as a contributing factor to growing educational inequities in Sweden. In addition, Swedish classrooms are becoming more homogenous in terms of student background characteristics, particularly concerning SES and language background (Hansson, 2012; Yang Hansen & Gustafsson, 2019). A study of TIMSS 2003 Grade 8 mathematics classrooms found that Swedish mathematics teachers, probably unintentionally, contributed to widening the achievement gap in low-SES classrooms and classrooms with students born outside Sweden (Hansson, 2012). In these classrooms, teachers were found to be less supportive and less organised compared to teachers in classrooms with higher SES and higher language competence. Particularly, second-language learners were found to need more support and scaffolding to succeed in mathematics. Hansson (2012) concluded that many students had not received appropriate teaching to support their learning needs, resulting in inequivalent teaching across classrooms. This was interpreted as a sign of pedagogical segregation.

In the TIMSS survey, students' mathematics confidence was significantly related to mathematics achievement. In the TIMSS 2023 mathematics assessment, there is in average a 90 points difference between Swedish fourth-grade students who reported high mathematics confidence compared to their peers with low mathematics confidence (Skolverket, 2024b). However, the proportion of students reporting very high mathematics confidence has declined over time. In TIMSS 2011, 40% of Swedish students reported very high mathematics confidence. This figure dropped to 37% in TIMSS 2019, and further to 28% in TIMSS 2023 (Skolverket, 2024b). Subsequently, the proportion of fourth-grade students reporting low mathematics confidence has increased, from 13% in TIMSS 2011 to 15% in TIMSS 2019, and 25% in TIMSS 2023 (Skolverket, 2024b). These patterns may be associated with issues considering teacher qualification and professional development. In TIMSS 2023, only 15% of Swedish Grade 4 mathematics teachers reported having the highest teacher qualification, compared to 90% in Finland, 43% in Norway, and the EU and OECD average of 40% (Skolverket, 2024b). The percentage of Swedish fourth-grade mathematics teachers who have participated in competence-enhancing training has declined over time. In TIMSS 2011, 60% of Swedish teachers reported participating in such training, dropping to 23% in TIMSS 2019, and a modest increase to 33% in TIMSS 2023 (Skolverket, 2024b).

In sum, although Sweden remains a high-achieving country in ILSAs, the country faces issues of educational equity, particularly compared to other Nordic countries.

# Chapter 5 Methodology

## Data

This thesis conducts secondary analysis using data from the international large-scale assessment Trends in Mathematics and Science Study (TIMSS), with a focus on teaching quality. In addition to achievement tests, TIMSS includes questionnaires designed to measure a range of theoretical constructs, including frameworks for assessing teaching quality (Klieme & Nilsen, 2022; Nilsen & Gustafsson, 2016). The TIMSS sampling procedure is well suited for examining class-level effects and classroom dynamics which is an advantage for research on teaching quality. Unlike PISA, which samples individual students from different classrooms, TIMSS samples intact classrooms, enabling analysis of teaching practices within the shared classroom context (Bellens et al., 2019; Klieme & Nilsen, 2022).

Data from the fourth-grade cohort of TIMSS 2019 is used. The samples in Studies I and II included 3,965 Swedish fourth-grade students. In Study III, the sample was expanded to include fourth-grade students in Denmark ($N$ = 3,227), Finland ($N$ = 4,696), and Norway ($N$ = 3,951).

## TIMSS

TIMSS, administered by the International Association for the Evaluation of Educational Achievement (IEA), is an international large-scale assessment that has monitored trends in mathematics and science achievement every four years since 1995 (Kelly et al., 2020). Test items are developed through extensive international collaboration to evaluate students' proficiency in mathematics and science at Grades 4 and 8 (Mullis & Martin, 2017). All test items are reviewed by participating countries to ensure their relevance and appropriateness across national contexts. TIMSS adheres to the ethical principles outlined in the Declaration of Helsinki. All data are publicly available via the international TIMSS data repository[4].

TIMSS employs a rigorous two-stage stratified cluster sampling design to ensure national representativeness across participating countries (Martin et al.,

[4] https://timss2019.org/international-database/

2017). In the first stage, a representative sample of schools is randomly selected. In the second stage, one or more intact classrooms are randomly chosen within each selected school (Mullis et al., 2020). To meet international requirements, the sample in each participating country must include at least 150 schools and one or more classes per grade, resulting in approximately 4,000 students (Mullis & Martin, 2017). TIMSS uses a matrix sampling design for the achievement tests in mathematics and science. Each student completes only a subset of the full test-item pool and also responds to background questionnaires covering several contextual factors, such as family background, peer interactions, school environment, and learning conditions (Mullis & Martin, 2017). To generate achievement scores that are comparable across all students, TIMSS uses an imputation method based on item response theory (Martin et al., 2020). Since schools are randomly selected, sampling weights are applied in the analysis to maintain representativeness and account for sampling error or missing data. Depending on the type of analysis, different sets of weights are required (Rutkowski et al., 2010). The cross-sectional design of TIMSS provides statistical relationships at one point in time rather than establishing causal relationships. Further technical and methodological information is available in The Technical Report for TIMSS 2019 (Martin et al., 2020).

## TIMSS questionnaires

TIMSS collects contextual data through questionnaires, administered to students in connection to the test, and to their teachers, school principals, and, at Grade 4, to parents. These questionnaire data can be examined alongside mathematics and science achievement scores to provide insight into factors relating to student achievement. In the student questionnaire, students respond to items covering basic demographic information about their home and school lives, including their attitudes towards learning mathematics and science, and perceptions of school climate (Fishbein et al., 2021). These questionnaire items can be combined to measure latent theoretical constructs, such as mathematics confidence, that are relevant for understanding students' performance on the assessment.

For secondary analysis of the questionnaire items, TIMSS provides both raw data in the form of scale scores and the transformed categorical indices derived from the questionnaire items. The transformed scales are constructed using a partial credit model within the framework of item response theory (IRT) and are presented as an index with three levels (high – medium - low), referred to as trichotomised scores (Yin & Fishbein, 2020). Trichotomisation of the continuous scale scores is intended to support researchers by simplifying secondary analyses and interpretations (Foy et al., 2020). However, according to Marsh et al. (2013), trichotomised scores tend to be less reliable than continuous scale scores and may reduce both the reliability and predictive validity of models including independent latent variables. Consequently, while TIMSS provides both raw scale scores and transformed trichotomised indices, this thesis relies on the raw scale score to preserve measurement precision when operationalising the independent variables.

# Variables

## Teaching quality

The TIMSS 2019 Context Questionnaire Framework includes student-reported measures of teaching quality, specifically classroom management and instructional clarity (Mullis & Martin, 2017). This thesis conceptualises teaching quality using two subdimensions drawn from the TBD framework and conceptual model of determinants of student outcomes (Nilsen & Gustafsson, 2016; Praetorius et al., 2018). While this selection introduces a risk of construct under-representation (Messick, 1984), previous research leaning on the TBD framework has also conceptualised teaching quality using one or more of the subdimensions (Senden et al., 2022). Accordingly, when conceptualising teaching quality, a pragmatic approach is adopted drawing on the constructs available in the TIMSS 2019 student questionnaire.

*Classroom management* is measured using six items reflecting the frequency of student-perceived disruptive behaviour during mathematics lessons, such as "There is disruptive noise" and "My teacher has to keep telling us to follow the classroom rules". All items address the classroom context and ask students how often disruptive behaviour occurs during mathematics lessons. Responses are given on a 4-point Likert scale, indicating the levels of agreement with the statements. Low values, on the left-hand side of the scale, indicate agreement with negative statements and high values indicate disagreement. Consequently, lower values on the scale reflect a more negative and disorderly classroom climate, whereas higher values indicate a more positive and calmer classroom climate.

*Instructional clarity* is measured using five out of six student-reported items reflecting the perceived clarity of instruction during the mathematics lesson, such as "My teacher is easy to understand", "My teacher has clear answers to my questions", "My teacher is good at explaining", "My teacher does a variety of things to help us learn", and "My teacher explains a topic again when we don't understand". These five items refer to the teacher as the subject of the phrase, such as "My teacher…". One item of the scale has the respondent as the subject of the phrase ("I know what my teacher expects me to do"). This item was excluded from the analyses of this thesis due to its poor factorial structure and limited statistical validity. Students indicate their level of agreement on a 4-point Likert scale, with lower values, on the left-hand side of the scale, reflecting more positive perceptions. Therefore, the items were reverse coded for all analyses so that higher values reflect greater perceived instructional clarity.

Although these items do not address particular mathematics teaching strategies in detail, it is important to recognise that mathematics instruction is characterised by distinct pedagogical practices, such as procedural fluency, conceptual understanding, and problem-solving, which differ from other subjects (Cohen et al., 2003; Lester, 2007). Nevertheless, given that the respondents are 10-year-old students in Grade 4, the mathematics instructional practices they experience are generally less complex compared to those at higher educational levels, such as in Grade 8.

### Mathematics confidence

The TIMSS 2019 student questionnaire includes the scale 'Students Confident in Mathematics', a subject-specific proxy for academic self-concept which is strongly related to mathematics achievement (Lee & Stankov, 2018; Mullis & Martin, 2017). The mathematics confidence scale comprises nine items rated on a 4-point Likert scale, where students indicate their agreement with statements such as: "I usually do well in mathematics"; "Mathematics is harder for me than for many of my classmates"; "I am good at working out difficult mathematics problems"; "Mathematics makes me confused". Four of the nine items are positively worded, while five are negatively worded. The positively worded items were reverse-coded so that higher values reflect greater mathematics confidence. As the mathematics confidence scale includes both positively and negatively worded items, which could increase cognitive load and linguistic complexity, construct validity was examined in Study I. Accordingly, the factorial structure of the positively and negatively worded items was examined to assess potential method effects associated with item wording (Hooper et al., 2013; Michaelides, 2019; Steinmann, Sánchez, et al., 2022).

## Student background variables

Educational research employing analyses such as structural equation modelling (SEM) often include student background variables to account for any confounding influences, to examine their relationship with student learning outcomes, and to improve model validity. These background characteristics shape students' learning and also enable researcher to address questions of educational equity (Bourdieu, 1986). In Studies II and III, student background factors were included from the TIMSS 2019 student questionnaire.

To measure students' socioeconomic status (SES), the number of books at home, as reported by students, was used as a proxy for SES. Although TIMSS provides a composite SES index (Home Resources for Learning), this index is considered a less valid measure of SES due to cultural diversity across participating countries in TIMSS (Rutkowski et al., 2013). Previous research has shown that the number of books at home is a strong predictor of academic achievement and it has been used as a valid SES proxy in the Nordic countries for a long period (Blömeke et al., 2016; Nilsen & Gustafsson, 2016; Rolfe, 2021). Using the number of books as a proxy for SES aligns with the theory of cultural capital by Bourdieu (1986). However, the increasing use of digital and audiobooks may reduce the long-term validity of this SES proxy. In the TIMSS questionnaire, students report the number of books at home on a 5-point Likert scale (BOOKS: 1=0-10; 2=11-25; 3=26-100; 4=101-200; and 5=over 200). The student questionnaire for Grade 4 does not include any measures of parental or guardian occupation. However, this measure is included in the questionnaire for Grade 8 students.

The second and third student background variables included in Studies II and III were students' home language and gender. While the TIMSS questionnaires do not include a direct measure of students' immigrant status, previous research has validated the use of students' self-reported frequency of speaking the test language at home as a proxy (Rutkowski, L., & Rutkowski, 2016). In the questionnaire, students indicated how often they spoke the test language at home (LANG) on a 4-point Likert scale where 1= never; 2= sometimes; 3= almost always; and 4= always. The student-reported gender was also included in the analysis (SEX: dummy coded 0= boy; 1= girl) (Mullis & Martin, 2017). However, given that schools and classrooms in the Nordic countries are not segregated by gender, this variable was excluded from the final analyses.

## Mathematics achievement

Mathematics achievement served as the academic achievement outcome variable in the thesis and was modelled using the five plausible values provided in the TIMSS 2019 data. In TIMSS, students' mathematics knowledge is assessed across the domains of numbers, measurement and geometry, and data, for example interpreting tables. A complex matrix-sampling booklet design is used, in which each student only completes a subset of items from a larger item pool, establishing a reliable estimation of their proficiency while limiting their test burden (Mullis & Martin, 2017). Students' achievement levels are estimated using item response theory (IRT), producing five plausible values to reflect the students' latent mathematics proficiency (Foy et al., 2020). To account for an unbiased and reliable estimation, all five plausible values were included in the SEM analysis, using the 'imputation' option in Mplus (von Davier, 2020; von Davier et al., 2009).

There is a pedagogical motive for having mathematics achievement as an outcome variable in this thesis. Unlike subjects such as reading literacy or English, which students often learn while playing online games or watching YouTube (Sundqvist & Sylvén, 2016), students primarily learn mathematics in formal classroom contexts guided by a teacher. Thus, mathematics achievement would be more likely to reflect the teaching quality the students receive in the classroom. Using students' perceptions of the mathematics classrooms would provide a more valid and reliable measure of teaching quality than relying on assessments of reading literacy or teacher self-reports (Reynolds et al. 2014). Mathematics instruction is typically characterized by procedural fluency, conceptual understanding and problem-solving (Cohen et al., 2003; Lester, 2007), which may differ from instructional practices in other subjects. However, while mathematics instruction involves subject-specific components, it is generally more generic in primary school, reflecting broader aspects of general pedagogical knowledge (GPK). Thus, mathematics instructions for 10-year-olds would generally be more generic and less complex compared to mathematics instructions at higher educational levels. Consequently, certain basic teaching instructions could be assumed to be generic practices, which are applicable across subject matters.

## Missing data

Missing data can arise in several ways, each requiring careful consideration to ensure valid statistical analysis. Although missing data are coded in the same way in TIMSS, they can arise in four different ways (Martin et al., 2020). First, the data may be omitted or invalid. This occurs when a respondent fails to answer an item, leaves the item blank, or provides an uninterpretable response. Second, an item may not have been administered to the respondent. For instance, if the item was excluded from the country-specific questionnaire version, such as items from the home resources scale, or due to technical failure or incorrect translation. Third, data may be logically not applicable when a prior reply filters out a follow-up question. Fourth, missingness may occur from students not reaching certain items in the achievement tests due to time limitations (Martin et al., 2020). In structural equation models (SEM) conducted in Mplus, missing data are handled using full information maximum likelihood (FIML), assuming the data to be missing at random. FIML requires a large sample but provides more precise and less biased estimates (Brown, 2015; Kline, 2016).

## Analytical methods

The following sections provide an overview of the main statistical methods employed in this thesis, a secondary analysis of TIMSS 2019 data using a quantitative approach. The analytical approach is based on structural equation modelling (SEM) by conceptualising teaching quality as a classroom-level construct linked to student learning outcomes. Drawing on the proposed *conceptual model of teaching quality and the situated classroom context* (elaborated on in Chapter 2), this doctoral thesis examines the relationships between two student-perceived teaching quality factors, classroom management and instructional clarity, students' mathematics confidence, and their mathematics achievement. The data were retrieved from the TIMSS data repository and prepared using the IDB Analyzer provided by the IEA (nd). The IDB Analyzer ensures that the complex hierarchical structure of the data is correctly merged and properly processed. The prepared data were then imported into IBM SPSS (version 27) for evaluation of descriptive statistics, including estimates such as means, standard deviations, missingness, and reliability estimates. Subsequent analyses were conducted in Mplus 8.6 (Muthén & Muthén, 2017). For further detailed descriptive statistics and information regarding the analytical methods, and software utilised, please refer to the three empirical studies.

TIMSS 2019 includes observable variables from the context questionnaire items which are combined into measurement models to capture theoretical latent constructs such as students' affective attitudes. These observable items, also referred to as indicators, are combined to create a latent variable. A latent variable captures a theoretical latent construct which cannot be observed or measured directly, such as mathematics confidence. An advantage of using latent constructs in statistical analysis is that they provide a more accurate representation of the underlying construct, and they are without the measurement error found in a single individual variable (Brown, 2015). Thus, aggregating observable indicators into a latent variable improves both validity and reliability (Field, 2018).

## Confirmatory factor analysis

To assess the construct validity of a theoretical latent variable such as mathematics confidence, confirmatory factor analysis (CFA) is used (Brown, 2015). CFA is used to evaluate the measurement model by examining the relationship between the observed indicators and the underlying latent construct. These relationships are expressed by factor loadings, which, when standardised, range from 0 to 1. The values of factor loadings represent the strength of the relationship between each observed indicator and the latent variable. A factor loading of 1 indicates a strong relationship, while a factor loading of 0 indicates no relationship. Factor loadings above 0.4 are considered acceptable for inclusion in the measurement model although suggested cut-offs vary across research disciplines (Field, 2018; Kline, 2016). Construct validity is assessed through both the factor loadings and the model fit. Indicators with factor loadings below 0.4 may require further examination. See Appendix A, Table A1 for information on the initial CFA measurement models and the standardised factor loadings for the Swedish sample.

The analysis of the mathematics confidence scale showed signs of potential method effects relating to the inclusion of both positively and negatively worded items. To investigate these validity concerns, the validation process in Study I extended the CFA to include exploratory structural equation modelling (ESEM). ESEM combines the advantages of both exploratory factor analysis (EFA) and CFA (Asparouhov & Muthén, 2009; Kline, 2016) by accounting for cross-loadings and possible multidimensionality of the data (Wang et al., 2018). For a more detailed description of the ESEM process and results, see Study I.

## Measurement invariance

To assess whether the latent constructs (classroom management, instructional clarity, and mathematics confidence) were measured equivalently and were comparable across the four Nordic countries, a series of multigroup confirmatory factor analyses (MGCFA) was conducted. This included testing increasingly restrictive levels of measurement invariance (MI) for each construct separately. The first three levels of invariance were examined. First, *configural invariance* assesses the overall model and determines if the instrument has the same factor structure across groups without constraints. Second, *metric invariance* refers to whether the instrument has equivalent factor loadings across groups. If metric invariance is met, comparisons of the relationships between constructs can carried out (He et al., 2019). Third, *scalar invariance* determines if the instrument has equal intercepts across groups (Kline, 2016; Teig et al., 2024). No difference between models means full invariance indicating that the same latent construct is measured across all groups. In contrast, significant differences may indicate cultural differences in how the respondents understand the items of the instrument, limiting the validity of cross-country comparisons (He et al., 2019). Reaching strict scalar invariance is suggested as unrealistic using ILSA data (Reynolds et al., 2022).

To evaluate if the model fits the data, several fit indices were used as the chi-square tests can be sensitive to large samples (Rutkowski & Svetina, 2014). The following fit indices were considered in this evaluation process: comparative fit index (CFI), Tucker-Lewis fit index (TLI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR). This thesis followed the guidelines suggested by Rutkowski and Svetina (2014), with more liberal cutoffs for MI analysis using ILSA data. For samples larger than 600, acceptable model fit was defined with CFI and TLI values above .95, and RMSEA and SRMR values below .08. For assessing relative changes in model fit indices, Rutkowski and Svetina (2014) recommended more liberal cut-offs for metric invariance ($\Delta$CFI = .020; $\Delta$RMSEA = .030), and for scalar invariance ($\Delta$CFI = .010, $\Delta$RMSEA = .0100). These cut-offs were followed in Study III, aligning with previous research (He et al., 2019). For a more detailed description of the MI testing and results, see Study III.

Previous research drawing on TIMSS data has identified challenges related to measurement invariance, particularly for non-cognitive constructs. In an analysis of non-cognitive measures from 29 countries participating in PISA 2015 and TIMSS 2015, He et al. (2019) found that self-reported Likert-scales were more sensitive to cross-cultural differences in item interpretation which limited score comparability. Similarly, a recent study examining the relationship between teaching quality and achievement in PISA 2012 and PISA 2022 concluded that measurement invariance was not supported for the construct of classroom management in countries such as France, Finland, and Norway (Liu et al., 2024). Liu et al. (2024) argued that classroom management was a context-dependent construct that could differ between classrooms and in how it was perceived by the students. He et al. (2019) pointed out that TIMSS implicitly assumed measurement invariance of item parameters, by relying on the partial credit model, whereas PISA uses the generalised partial credit model which accounts for cultural variation (Yin & Fishbein, 2020). Reynolds et al. (2022) examined the comparability of the non-cognitive scales in the student questionnaire in TIMSS 2019 across all 58 participating countries in Grades 4 and 8. Their findings supported cross-country comparability for these constructs, which is of importance for secondary analysis of ILSA data. However, the authors highlighted that "strict scalar invariance in the traditional sense may be unrealistic in an ILSA context" (Reynolds et al., 2022, p. 10) and suggested that the observed non-invariance may be related to negatively worded items in the attitudinal scales (Reynolds et al., 2022).

## Structural equation modelling

To test the theoretical hypotheses regarding the relationships between the latent variables, the validated measurement models are combined using structural equation modelling (SEM) (Kline, 2016). In a structural model, a simple regression includes at least two latent variables: a predictor (independent variable, $x$), referred to as exogenous in SEM, and a dependent outcome variable ($y$), referred to as endogenous in SEM. The dependent variable is regressed on the independent variable, resulting in a regression coefficient estimate that expresses the strength and direction of the relationship between these two variables ($x \rightarrow y$), also referred to as the total effect ($c$) (Kline, 2016). A positive regression coefficient indicates a direct relationship between the independent and dependent variables, while a negative coefficient indicates an inverse relationship between the independent and dependent variables. In SEM, the relationships between variables are often expressed as standardised regression coefficients (β), indicating how much a one standard deviation change in the independent variable is expected to change the dependent variable. The regression coefficient estimates can range from -1 to 1. An coefficient close to -1 or 1 indicates a strong relationship, while a coefficient close to 0 indicates a weak or no relationship (Kline, 2016).

TIMSS data is hierarchically structured, with students nested within intact classrooms, which in turn are nested within schools, and then within countries. Given this nested structure of the data, all models in this thesis account for clusters by running complex models with classrooms as cluster units. A key methodological consideration when analysing hierarchical and nested data is whether to conduct single-level SEM or multilevel SEM (MSEM). Not addressing the data at the appropriate level may result in both statistical and conceptual issues. Statistically, aggregating individual-level data to higher levels may result in an insufficient number of units at the higher levels, thus reducing statistical power (Hox, 2002). Conceptually, there could be issues if the aggregated data are interpreted at the wrong level. For instance, making individual-level conclusions based on group-level data is known as an ecological fallacy, while drawing group-level conclusions based on individual-level data, is referred to as an atomistic fallacy (Hox, 2002).

## Intra-class correlations coefficients

To gain information about the nested structure of the data, the intra-class correlation coefficients (ICC) determine whether to run a single-level structural equation model (SEM) or multilevel SEM (MSEM) (Khine et al., 2013). ICC estimates indicate the proportion of variance in a specific variable that can be attributed to the group-level factors, such as classrooms, rather than individual-level differences among students (Hox et al., 2011). An ICC estimate close to 0 indicates that most of the variance is at the individual level, supporting the use of a single-level SEM. In contrast, ICC estimates above .05 indicate group-level differences related to the cluster, such as classrooms, warranting the use of MSEM (Hox et al., 2011). However, these cut-offs can vary across research disciplines, and even small ICCs (e.g., .03) may warrant MSEM to avoid Type I errors.

MSEM accounts for the hierarchical structure of the data by dividing the variance into two levels, the within-group (the student level) and the between-group (the classroom level). This method recognises the clustering of the data, allowing researchers to examine variability across clusters such as classrooms. SEM and MSEM do not underestimate standard errors in complex sampling designs with nested data and reduce random error and noise (Kline, 2016). Furthermore, MSEM addresses validity issues concerning student ratings by separating individual- and group-level variation (Scherer & Gustafsson, 2015).

In Study III, ICC estimates across the Nordic countries indicated a substantial proportion of variance at the classroom level for several variables used which supported the use of MSEM. As presented in Table 2, ICC estimates for mathematics achievement are high for all Nordic countries included. For instance, the Swedish ICC for mathematics achievement was .19, indicating that 19% of the variance in student mathematics achievement was attributed to differences between classrooms. In contrast, the ICC estimates for mathematics confidence were below .03 across all countries, which would suggest single-level SEM was appropriate (Hox et al., 2011).

Table 2 Estimated classroom-level intra-class correlation coefficients (ICCs)

| TIMSS 2019 Grade 4 | Denmark | Finland | Norway | Sweden |
|---|---|---|---|---|
| Mathematics achievement | .10 | .18 | .12 | .19 |
| SES (BOOKS) | .08 | .08 | .11 | .16 |
| Home language (LANG) | .03 | .13 | .07 | .17 |
| Gender (SEX) | .003 | .004 | .004 | .01 |
| Classroom management | .08 - .17 | .12 - .19 | .11 - .18 | .14 - .25 |
| Instructional clarity | .06 - .12 | .08 - .14 | .05 - .10 | .07 - .12 |
| Mathematics confidence | .02 - .02 | .02 - .02 | .02 - .03 | .01 - .03 |

This thesis employed both single-level SEM and multilevel SEM. The modelling choices were guided by the different research aims of the studies. In Study I, single-level SEM was used as the focus was on examining the construct validity of the mathematics confidence scale and not on differences between classrooms. However, the hierarchical structure of the data, the high ICC estimates, and the aim to examine classroom-level differences relating to student-perceived teaching quality justified the use of MSEM in both Studies II and III (Lazarides & Buchholz, 2019). Secondary analysis based on hierarchical data such as TIMSS has to include appropriate sampling weights corresponding to the structure of the clustered sample and the correct grouping variable to ascertain valid and reliable statistical analysis (Fishbein et al., 2021; Klieme & Nilsen, 2022; Martin et al., 2020), see the separate studies for more information.

## Mediation

Mediation analysis in SEM is used to examine whether the relationship between an independent variable and a dependent variable could be passed through a third, mediating variable. It is possible to assume, based on theory, that this relationship could be influenced by an additional independent variable, suggesting the suitability of a mediation SEM model. A simple regression includes one independent variable and one dependent variable with a total effect (*c*) between them. In mediation, as displayed in Figure 4, a third variable is taken into account, and the relationship is decomposed into two paths, one direct path estimating the direct effect (*c'*) and the other indirect path via the mediating variable (*ab*) estimating the indirect effect ($c - c' = ab$) (Kline, 2016). To establish mediation, significant relationships between the independent and the mediating variables, and between the mediating variable and the outcome variable are required (MacKinnon et al., 2000). Mediation analysis is often used on longitudinal data, with the mediator measured at a different time point from the predictor and outcome variables (Kline, 2016). Mediation models drawing on data with a cross-sectional design such as TIMSS, do not allow for any causal inferences, but only present relationships (Kline, 2016).

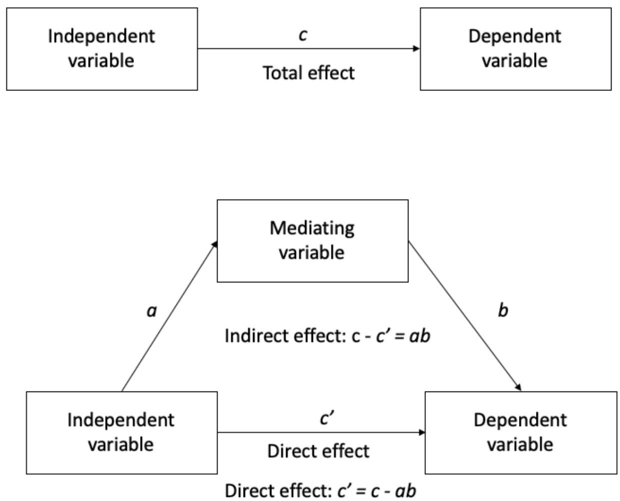


Figure 4 A single regression and a mediating regression SEM

An example from Study I of a single regression SEM (above) and a mediating SEM with total effect, direct effect and indirect effect (MacKinnon et al., 2004).

In Study I, mediation models were included to test possible indirect effects of the two teaching quality factors (classroom management and instructional clarity) on mathematics achievement, with mathematics confidence as the mediating variable (see Fig. 4).

## Reliability, validity and validation

The following sections discuss the quality and usefulness of the data, and the findings presented in this thesis by addressing aspects of reliability, validity, and validation.

## Reliability

*Reliability* refers to the consistency of a measurement instrument to produce comparable results across different situations (Field, 2018). This thesis is a secondary data analysis, relying on reliable sampling methods and assessments of the IEA and TIMSS (Mullis & Martin, 2017). TIMSS employs a rigorous two-stage clustered sampling design to ensure a reliable and representative sample (Martin et al., 2017). In the first stage, schools are randomly selected, in the second stage, one or two intact classrooms within each school are randomly sampled, and all students in these classrooms participate (Mullis & Martin, 2017). This sampling procedure is a reliable way to achieve a representative sample of students. Hence, it does not give a representative sample of teachers (Fishbein et al., 2021).

Following a rigorous test development process and a robust scaling methodology, the reliability of the outcome variable mathematics achievement is considered to be high (Martin et al., 2017; von Davier, 2020). The inclusion of all five plausible values of mathematics in the analysis strengthens the reliability.

## Validity

*Validity* is a central concept in both research and assessment, and it is important to clarify what is meant by validity and how it is applied in this thesis. In contemporary research and educational measurement, both validity and reliability are central aspects of the development, interpretation, and use of tests (Messick, 1989; Newton & Shaw, 2014). According to the Standards for Educational and Psychological Testing, validity is defined as the degree to which an instrument successfully measures the intended construct (AERA, 2014). Consequently, validity is not a dichotomous characteristic, meaning that a measurement instrument can never be valid or invalid, but it is valid to a certain extent (Linn, 2010).

There are several aspects of validity relevant to the interpretation of the findings in this doctoral thesis (Messick, 1989; Shadish et al., 2002). *Construct validity* is a way to describe to what extent an instrument accurately captures the theoretical latent construct it was intended to measure (Messick, 1989). A latent theoretical construct, such as motivation, cannot be directly observed, only indirectly measured via indicators that are presumed to reflect the underlying construct. However, the construct validity of an abstract construct such as motivation cannot be proven but only supported to a certain degree (Spector, 1992). The statistical analysis in the thesis leans on the structural equation modelling (SEM) framework, which includes the measurement modelling part and the structural modelling part. In the measurement modelling, confirmatory factor analysis (CFA) was used to evaluate construct validity and the model fit of this theory-driven model (Wagemaker, 2020).

Messick (1989) presented construct validity as a unitary concept, outlining two main threats to construct validity, *construct-irrelevant variance* and *construct under-representation*. First, *construct-irrelevant variance* refers to when items included to reflect a latent construct, tap into other constructs rather than the intended. This could be the case if items included in a measurement model are perceived as complex and linguistically demanding, or if respondents are affected by external factors such as low motivation or anxiety. As an example, in the CFA analyses conducted in this thesis, item 1 in the latent construct of instructional clarity had a factor loading of .378, below the recommended cut-off of .40. When this item was excluded from the following CFA model, the overall model fit improved significantly. In Study I, the construct validity of the mixed-format scale of mathematics confidence was examined with CFA, ESEM, and cross-cultural semantic validation. The statistical analysis revealed signs of multidimensionality and indications of construct-irrelevant variance. The factorial structure showed evidence of low validity, suggesting method effects related to five negatively worded items. Item 7, in the mathematics confidence scale, had a factor loading of .394, indicating a weak relationship to the intended construct. It was suggested that item 7 may capture the latent construct of a supportive teacher rather than mathematics confidence. Based on these findings, it was suggested that only the three positively worded items with strong psychometric properties be retained for subsequent analysis.

The second major threat to construct validity is *construct under-representation*, that is when a measurement instrument fails to capture important aspects of the intended construct. Construct under-representation occurs when an insufficient number of items are included to reflect a latent construct. As a consequence the measurement instrument is too narrow to accurately capture the aspects of the construct (Messick, 1989). The use of student-perceived teaching quality could be argued to represent a mono-method bias in this thesis. However, existing research has demonstrated that aggregated student ratings are a valid and reliable approach to operationalise the classroom level constructs such as classroom management, and instructional clarity (Marsh et al., 2012; Scherer & Gustafsson, 2015).

When evaluating the quality of the findings, further forms of validity were considered, including convergent, discriminant, statistical conclusion, internal, and external validity. *Convergent validity*, meaning that the items reflect the intended construct, was assessed by examining the factor loadings of the items onto the latent construct. The factor loadings indicate how well the different items capture the latent construct with values ranging from 0 to 1. A value of 1 indicates that the item fully measures the latent construct and a value of 0 no relationship. Appendix A, Table A1 presents the CFA results of the Swedish sample, including the model fit and the standardised factor loadings.

*Discriminant validity* ensures that latent constructs are distinct from other latent constructs and do not have a high correlation. Discriminant validity was supported when the results showed low correlations between the different latent constructs, confirming that these constructs are separate. In the present studies, the latent constructs (classroom management, instructional clarity, and mathematics confidence) were found to reflect separate dimensions. To further ensure validity and reliability, the model fit and the factorial structure of the models were assessed through CFA (Brown, 2015).

*Statistical conclusion validity* refers to the extent to which appropriate statistical methods have been implemented in the analysis, the accuracy of the relationships between the independent and the dependent variables, and the accuracy and reliability of the conclusions drawn from the findings. TIMSS data have high psychometric standards with a reliable, robust sampling procedure (Martin et al., 2017). This minimizes the threat of low statistical power, unreliability of measure, selection bias, and sampling issues. However, secondary analysis requires correct methodological considerations and appropriate handling and interpretation of data. For this reason, the IDB Analyzer, an analytical software tool provided by the association IEA, was used to correctly merge and prepare complex data files (Meinck et al., 2020). Additionally, this thesis addressed the hierarchical and nested structure of the TIMSS data, by using multilevel structural equation models (MSEM) and accounting for the classroom clustering with complex models and appropriate sampling weights using the software Mplus 8.6 (Rutkowski et al., 2010).

*Internal validity* describes the extent to which a measurement instrument and its design can provide evidence supporting causality and discusses the directions of the relationships between the variables included. As this doctoral thesis is based on TIMSS which has a cross-sectional design, all relationships found in the analysis present the correlated relationships at the specific measurement and no causal claims can be drawn (Rutkowski et al., 2010). Given the cross-sectional design of TIMSS data, other internal validity considerations such as attrition, maturation, or instrument decay are not applicable.

The final consideration relates to *external validity*, referring to the generalisability of the results and assessing whether reliable and valid inferences can be drawn from the findings (Shadish et al., 2002). External validity threats may arise if there is selection bias in the sample and thus not representative of the intended population. However, as this thesis is a secondary study based on IEA data from TIMSS 2019, which selects a representative sample of the target population, such threats can be considered to be minor. Nevertheless, the interpretation, presentation, and dissemination of findings must be conducted carefully, correctly, and transparently. The findings presented in the thesis may have implications for educational policy, teacher education, and teachers' professional practice. For instance, results indicating the importance of teachers' general pedagogical skills, such as classroom management and structural organisational activities, aimed at improving student learning outcomes, may influence teacher education and working teachers' professional development.

## Validation

*Validation* refers to the evaluation process of the appropriateness, meaning, and social consequences of the interpretations of research findings or test scores (Kane, 2006). Kane (2006) describes validation as an ongoing, iterative process that the researcher uses to build evidence to support the credibility, trustworthiness, and implications of the research findings. This process requires considerations of whether the evidence and theoretical rationale support the intended use and interpretation of research findings or test scores.

As part of the validation process of this thesis, one step focused on evaluating the construct validity of the questionnaire items, with particular attention to possible method effects and response bias, using statistical analysis such as CFA. Response bias can arise if scale items tap into different constructs, resulting in construct-irrelevant variance, or if too few items are included in the scale, leading to construct under-representation (Messick, 1989; van Sonderen et al., 2013). To assess these threats to construct validity, the factorial and statistical validity of the items included in each scale were assessed (DiStefano & Motl, 2006). Indications of low factorial structure would suggest weak internal consistency and would negatively influence both construct validity and reliability (Roszkowski & Soven, 2010; Ye & Wallace, 2014).

Special attention was given to whether aspects of the scale design could introduce possible method effects or response bias which might threaten construct validity (Scherer et al., 2016). Questionnaire scales often combine positively and negatively worded items, such as those in the mathematics confidence scale in TIMSS 2019, to maintain respondents' attention and reduce acquiescence response bias (Marsh, 1994; Roszkowski & Soven, 2010). However, negatively worded items may be more cognitively demanding, particularly for younger respondents, and would require careful consideration in the validation process, as they could be a threat to construct validity, (Marsh, 1994; Van Dam et al., 2012).

## Sematic validation

In a semantic validation, the linguistic and conceptual equivalence of the translated questionnaire items included in the constructs of classroom management, instructional clarity, and mathematics confidence were assessed. The student questionnaire items employed in the thesis were translated from the source language, English, into the target language, Swedish. Although it is often assumed that items remain linguistically equivalent throughout the translation process, differences in phrasing and cultural context may occur, potentially affecting both item clarity and meaning at face value. To address these concerns, the validation process of this thesis included a pragmatic, semantic cross-cultural validation to determine whether the translated items maintained their original meaning, clarity, comprehensibility, and contextual appropriateness (Ercikan et al., 2015; Osborn, 2004). This process aimed to assess both conceptual and linguistic equivalence between the source and the target language. Study I included the semantic validation of the mathematics confidence scale.

The semantic validation process adopted Kane's argument-based approach, including both an interpretive argument and a validity argument to assess the item equivalence (Kane, 2006). The interpretive argument examined the construct equivalence, that is, whether the intended meaning of each original item was kept in the translated version (Ercikan et al., 2015). During this phase, the linguistic, content, and cultural equivalence of the translated wordings and expressions of each item were reviewed followed by interpretations on which inferences were based (Ercikan et al., 2015). The translations were evaluated for ambiguity and for the presence of words or phrases that may have different connotations for different groups. If the cultural equivalence is not maintained and an item is interpreted differently across cultural contexts due to the translation, it may cause item bias or cultural bias (Ercikan et al., 2015; He et al., 2022). If items are evaluated as culturally biased, He et al. (2022) recommended excluding such items from secondary analysis. The next step in the semantic validation process included the validity argument. In this step, an evaluation of the interpretive argument was made to ensure that reasonable inferences are drawn from the interpretive argument. No new meaning should have been added to the translations, which should be accurately phrased to ensure that the translated items convey the intended meaning (Ercikan et al., 2015; Hooper, 2022).

The semantic validation process also included analysis of the syntax of the questionnaire items to ensure consistency in the grammatical structure. For instance, items may differ in terms of the subject performing the action expressed in the item. Some items may have the student as the agent (e.g., I know what my teacher expects me to do), the whole class as the agents (Students don't listen to what the teacher says), or the teacher as the agent (My teacher is good at explaining mathematics), or some of the items could assess the respondent's feeling (Mathematics makes me nervous). Thus, if the syntax differs across translations, for instance, if an item has one agent or recipient of the action in the source text but a different agent or recipient in the target text, this could cause response bias, weakening the factorial structure and construct validity (Wagner et al., 2013). The three latent constructs (classroom management, instructional clarity, and mathematics confidence) analysed in the thesis were included in the semantic validation process.

*Classroom management*

The semantic validation of the classroom management scale is presented in Table 3, comparing the source and target language versions of each item. The first column displays the English source text, and the second column presents the Swedish target text. For each item, the top line presents the phrasing as it appears in the questionnaire, while the line below shows the literal translation of the corresponding item within brackets [ ]. To exemplify, the English version of item number 5 (MS4e) is "Students interrupt the teacher" with the corresponding Swedish questionnaire item "Elever avbryter läraren", as they appear in the student questionnaire. The literal back-translation is presented below within brackets. In the classroom management scale, two items with a near identical translation were marked with one asterisk in Table 3 and are discussed in more detail in the following section.

Two items in the classroom management scale showed minor semantic deviations from the English source text. In item 4 (MS4d), the source text uses '*for*', which was translated into Swedish as '*innan*'. A literal back-translation of '*innan*' would be '*before*'. Although they seem similar, there is a subtle difference in meaning. The word '*before*' emphasises the duration of time that passes before the students become quiet, whereas '*for*' highlights the act of waiting as a condition until the students settle down.

In item 6 (MS4f), the source text '*keep telling*' was translated into Swedish as '*påminna*'. While both phrases involve the teacher telling or prompting the students, they differ in nuance and emphasis. The English '*keep telling*' implies repeated instruction as the students are not compliant. By contrast, the Swedish '*påminna*' suggests that the teacher gives a gentle reminder to the students to comply, but it is less frequent and urgent.

*Instructional clarity*

The instructional clarity scale consisted of six items (see Table 4), of which five items displayed minor differences in phrasing or emphasis between the English source text and the Swedish target text, which call for further discussion. These translations are marked with an asterisk.

In item 1 (MS3a), the English phrase '*expect me to do*' emphasises the action or the task the teacher wants and expects the student to carry out, and the student is aware of this expectation. The Swedish translation '*förväntar sig av mig*' expresses a more general expectation from the teacher for a certain student behaviour or performance, rather than a specific action or task as the English phrase '*to do*' is excluded in the Swedish version. The lack of the verb '*to do*' in the Swedish version broadens the interpretation to include expectations related to behaviour, attitude, or performance. This broader and maybe more ambiguous phrasing could explain why this item showed low factor loadings in the confirmatory factor analysis (CFA) and was therefore excluded from the subsequent studies of this thesis.

In item 2 (MS3b), the English version begins with the subject of the sentence, '*my teacher*', emphasising the teacher's instructional qualities. In contrast, the Swedish version includes a dummy subject '*det*' and moves '*my teacher*' to the object position of the phrase. This syntactic change gives more focus on the act of understanding rather than on the teacher's ability to communicate with the students, shifting the focus of the item slightly.

Table 3 Semantic validation of the classroom management scale

| Scale name | # | Indicator<br>English | Indicator<br>Swedish |
|---|---|---|---|
| *Classroom management* | | | |
| | | How often do these things happen in your mathematics lessons?<br>[How often do the following happen in your mathematics lessons]<br><br>4-point Likert scale: Every or almost every lesson – about half the lessons – some lessons – never | Hur ofta händer följande på dina matematiklektioner?<br>[Hur ofta händer dessa saker på dina matematiklektioner]<br><br>4-point Likert scale: Varje eller nästan varje lektion – ungefär hälften av lektionerna – vissa lektioner – aldrig |
| MS4a | 1 | Students don't listen to what the teacher says<br>[students are not listening to what the teacher is saying] | Elever lyssnar inte på vad läraren säger<br>[Elever lyssnar inte på vad läraren säger] |
| MS4b | 2 | There is disruptive noise<br>[It is a disturbing noise] | Det är störande oljud<br>[Det finns störande ljud] |
| MS4c | 3 | It is too disorderly for students to work well<br>[It is too disorderly for students to be able to work well] | Det är för stökigt för att elever ska kunna arbeta bra<br>[Det är för stökigt för elever att arbeta bra] |
| MS4d * | 4 | My teacher has to wait a long time **for** students to quiet down<br>[My teacher has to wait a long time before the students quiet down] | Min lärare måste vänta länge **innan** elever tystnar<br>[Min lärare måste vänta länge för att eleverna ska tystna] |
| MS4e | 5 | Students interrupt the teacher<br>[Students interrupt the teacher] | Elever avbryter läraren<br>[Students interrupt the teacher] |
| MS4f * | 6 | My teacher has to **keep telling** us to follow the classroom rules<br>[My teacher has to remind us to follow the rules in the classroom] | Min lärare måste **påminna oss** om att följa reglerna i klassrummet<br>[Min lärare måste hela tiden påminna oss om att följa klassrumsreglerna] |

Note. * = near identical translation of the words marked in bold, these are discussed; [literal translation of the corresponding item]

In item 3 (MS3c), the English version uses the phrase '*has clear answers*', implying that the teacher has the knowledge and a clear answer available. The Swedish version uses the verb '*gives'*, shifting the focus from the teacher possessing the knowledge to the active behaviour. While this could be considered a minor change, this may still influence how students interpret the teacher's clarity.

In item 5 (MS3e), the English version uses the phrase '*a variety of things*', which is translated into Swedish as '*different things*'. The Swedish version suggests less pedagogical diversity. Moreover, the English verb '*learn'* is translated as '*understand*' in Swedish, shifting the focus from the learning process to the outcome of understanding.

In item 6 (MS3f), the English version includes the words '*a topic*' and '*again*' whereas the Swedish version has omitted '*topic'* and translates '*again*' with '*once again*'. The omission of '*a topic*' in the Swedish version removes the reference to specific content, thus shifting the focus from what is explained to the act of explaining. Additionally, the use of '*once again*' in the Swedish version suggests that the teacher will only explain one more time and not indefinitely, a minor change of the time aspect.

*Mathematics confidence*

In the mathematics confidence scale, four of the nine items showed minor differences in phrasing and emphasis in the Swedish version, which calls for further discussion. These items are marked with one asterisk in Table 5.

In item 1 (MS5a), the English version uses '*I*' as the subject, focusing on the student's ability and performance. The Swedish translation uses the dummy subject '*it*', shifting the focus to a more general experience of the situation rather than the student's personal experience.

In item 7 (MS5g), the Swedish version uses the preposition '*i*' rather than '*på*'. While both prepositions can be used interchangeably in this context, there is however a minor difference in nuance. The phrase '*bra i matematik*' refers to being good or having the ability within a specific subject such as mathematics, while '*bra på matematik*' refers to a broader ability or aptitude in mathematics and not just having knowledge of mathematics. In the CFA and the ESEM analyses, the psychometric analysis of this item showed low factor loadings. It was suggested in Study I, that this item reflected students' perceptions of teacher support rather than the construct of mathematics confidence. This item was omitted in the subsequent studies of this thesis.

Table 4 Semantic validation of the instructional clarity scale

| Scale name | # | Indicator<br>English | Indicator<br>Swedish |
|---|---|---|---|
| *Instructional clarity* | | | |
| 4-point Likert scale | | | |
| | | How much do you agree with these statements about your mathematics lessons?<br>[How well do you think the following statements are true about your mathematics lessons]<br><br>4-point Likert scale<br>Agree a lot – agree a little – disagree a little – disagree a lot | Hur bra tycker du att följande stämmer om dina matematiklektioner?<br>[Hur mycket instämmer du i dessa uttalanden om dina matematik-lektioner]<br><br>4-point Likert scale<br>Stämmer precis – stämmer ganska bra – stämmer inte så bra – stämmer inte alls |
| MS3a * | 1 | I know what my teacher **expects me to do**<br>[I know what my teacher expects of me] | Jag vet vad min lärare **förväntar sig av mig**<br>[Jag vet vad min lärare förväntar sig att jag ska göra] |
| MS3b * | 2 | **My teacher** is easy to understand<br>[It is easy to understand my teacher] | **Det** är lätt att förstå min lärare<br>[Min lärare är lätt att förstå] |
| MS3c * | 3 | My teacher **has** clear answers to my questions<br>[My teacher gives clear answers to my questions] | Min lärare **ger** tydliga svar på mina frågor<br>[Min lärare har tydliga svar på mina frågor] |
| MS3d | 4 | My teacher is good at explaining mathematics<br>[My teacher is good at explaining mathematics] | Min lärare är bra på att förklara matematik<br>[Min lärare är bra på att förklara matematik] |
| MS3e * | 5 | My teacher does **a variety of things** to help us **learn**<br>[My teacher does different things to help us understand] | Min lärare gör **olika** saker för att hjälpa oss att **förstå**<br>[Min lärare gör en mängd olika saker för att hjälp oss att lära] |
| MS3f * | 6 | My teacher explains a **topic again** when we don't understand<br>[My teacher explains once again when we don't understand] | Min lärare förklarar **en gång till** när vi inte förstår<br>[Min lärare förklarar ett ämne igen när vi inte förstår] |

Note. * = near identical translation of the words marked in bold, these are discussed; [literal translation of the corresponding item]

In item 8 (MS5h), the English version compares the difficulty of mathematics to all other subjects at school. The Swedish version claims that mathematics is the hardest subject, omitting any comparisons to other subject matters. While both versions focus on the student's experience of difficulty, the English version includes the comparison which is absent in the Swedish translation.

In item 9 (MS5i), the English version uses '*mathematics*' as the subject, suggesting that mathematics is the source of confusion. In contrast, the Swedish version uses '*I*' as the subject, focusing on the student's personal experience and emotional reaction to working with mathematics rather than attributing the confusion to mathematics.

To conclude, there are indications of response bias in the translation of the items discussed above. The measurement invariance testing of the latent constructs and the analysis of their factorial structure in the CFA, and ESEM, showed similar results. One way to handle such discrepancies could be to include a qualitative approach with interviews or a think-aloud protocol and examine students' perceptions of the items. Alternatively, a quantitative approach with methods such as comparative pilot studies of both original items and culturally equivalent translations to evaluate the variance explained (Fenn et al., 2020). Further studies are needed to examine cross-cultural bias in translated questionnaire items.

Table 5 Semantic validation of the mathematics confidence scale

| Scale name | # | Indicator English | Indicator Swedish |
|---|---|---|---|
| *Mathematics confidence* | | | |
| | | How much do you agree with these statements about mathematics?<br>[How well do you think the following statements are true about mathematics]<br>4-point Likert scale<br>Agree a lot – agree a little – disagree a little – disagree a lot | Hur bra tycker du att följande stämmer om matematik?<br>[Hur mycket instämmer du i dessa uttalanden om matematik]<br>4-point Likert scale<br>Stämmer precis – stämmer ganska bra – stämmer inte så bra – stämmer inte alls |
| MS5a * | 1 | **I** usually do well in mathematics<br>[It usually goes well for me in mathematics] | **Det** brukar gå bra för mig i matematik<br>[Jag brukar klara mig bra i matematik] |
| MS5b | 2 | Mathematics is harder for me than for many of my classmates<br>[identical] | Matematik är svårare för mig än för många av mina klasskamrater<br>[Matematik är svårare för mig än för många av mina klasskamrater] |
| MS5c | 3 | I am just not good at mathematics<br>[I am simply not good at mathematics] | Jag är helt enkelt inte bra i matematik<br>[Jag är bara inte bra på matematik] |
| MS5d | 4 | I learn things quickly in mathematics<br>[identical] | Jag lär mig snabbt i matematik<br>[Jag lär mig snabbt i matematik] |
| MS5e | 5 | Mathematics makes me nervous<br>[identical] | Matematik gör mig nervös<br>[Matematik gör mig nervös] |
| MS5f | 6 | I am good at working out difficult mathematics problems<br>[I am good at solving difficult mathematics tasks] | Jag är bra på att lösa svåra matematikuppgifter<br>[Jag är bra på att lösa svåra matematikproblem] |
| MS5g * | 7 | My teacher tells me I am good at mathematics<br>[identical] | Min lärare säger att jag är bra **i** matematik<br>[Min lärare säger att jag är bra **på** matematik] |
| MS5h * | 8 | Mathematics is harder for me **than any other subject**<br>[Mathematics is the hardest subject for me] | Matematik är **det svåraste ämnet** för mig<br>[Matematik är svårare för mig än något annat ämne] |
| MS5i * | 9 | **Mathematics** makes me confused<br>[I get confused by mathematics] | **Jag** blir förvirrad av matematik<br>[Matematik gör mig förvirrad] |

Note. * = near identical translation of the words marked in bold, these are discussed; [literal translation of the corresponding item]

## Ethical considerations

This thesis is a secondary analysis of data retrieved from a publicly open webpage administered by the International Association for the Evaluation of Educational Achievement (IEA). The dataset is fully anonymised, with no possibility to identify students, teachers, or schools. As such, ethical authorisation was not a requirement. Research conducting secondary analysis of international large-scale assessments (ILSAs) has gained influence on educational policymakers and society (Grek, 2009; Johansson, 2016) and, consequently, findings from secondary analyses could result in policy decisions that ultimately have an impact on both individuals and populations (Suri, 2020). Consequently, ethical considerations regarding the dissemination of the thesis have to be clear and detailed and include limitations (Gustafsson, 2018). This thesis complies with the ethical procedures of IEA, follows the principles of good research practice and ethical guidelines provided by the Swedish Research Council (Vetenskapsrådet, 2024), and abides by the guidelines in the Declaration of Helsinki stipulating that considerations for individual rights and dignity should precede those of science and society.

# Chapter 6 Results

This chapter provides a summary of the three empirical studies of this thesis. The studies were guided by four overarching research questions, as presented in Chapter 1, 1) *To what extent can the student questionnaire items validly measure the latent constructs of mathematics confidence and two subdimensions of teaching quality, classroom management and instructional clarity?* 2) *To what extent do aspects of teaching quality matter for student learning outcomes such as mathematics confidence and mathematics achievement and are there differences across groups of students depending on their socioeconomic background, home language, and gender?* 3) *To what extent are there differences between classrooms in the relationship between aspects of teaching quality and learning outcomes such as mathematics confidence and mathematics achievement?* 4) *To what extent could aspects of teaching quality mitigate the negative influence of student background factors and decrease the growing achievement gap?* In sum, these questions address the validity of the measurement instruments, the role of teaching quality concerning student learning outcomes, and the relationships between classroom composition and variation in student learning outcomes across classrooms.

## Study I: Validating the mixed-worded mathematics confidence scale in TIMSS

Using Swedish TIMSS 2019 Grade 4 data, Study I examined the construct validity of the mixed-worded mathematics confidence scale, employing confirmatory factor analysis (CFA), exploratory structural equation modelling (ESEM), and a semantic validation of the translated items.

Results showed possible indications of method effects and response bias in the mixed-worded mathematics confidence scale. Firstly, the inter-item correlation matrix suggested that some respondents may not have adjusted their responses appropriately, as they failed to reverse their ratings on the negatively worded items. This lack of variation between positively and negatively worded items may indicate method effects (Steinmann, Strietholt, et al., 2022). One reason could be attributed to the wording of the items as some items could be perceived as more linguistically complex and cognitively demanding, particularly for younger students. Secondly, the semantic validation process found possible cross-cultural bias related to the phrasing or translation of questionnaire items. The item phrasing and translation have to ensure conceptual and linguistic equivalence between the source text and the target text (Ercikan, 1998). However, findings indicated possible cross-cultural bias in four of the nine items, suggesting that these items were not conceptually and linguistically equivalent. The bi-factor ESEM model showed signs of multidimensionality within the mixed-worded scale (Asparouhov & Muthén, 2009; van Zyl & Klooster, 2022). Several items had factor loadings below .40, suggesting limited reliability for these items (Chiu, 2008; Kline, 2016). Notably, item 7, although positively worded, had low factor loadings. This item could be argued to reflect a different latent construct, such as the 'supportive teacher' construct, rather than mathematics confidence. Results from Study I revealed possible method effects and response bias related to the negatively worded items, as well as construct-irrelevant variance in item 7. Therefore, it is reasonable to consider retaining only three positively worded items (items 1, 4, 6) in subsequent analyses.

Additionally, by leaning on the TBD framework and the dynamic model (Creemers & Kyriakides, 2010; Praetorius et al., 2020), Study I examined the direct and indirect relationships between the two factors of student-perceived teaching quality (classroom management and instructional clarity) and mathematics achievement, with mathematics confidence as a mediator using SEM. Results revealed substantial, positive relationships between mathematics confidence and mathematics achievement, aligning with previous research (e.g., Lee & Stankov, 2018; Marsh & Craven, 2006; Stankov et al., 2014). Both direct and indirect positive relationships were found between classroom management and mathematics achievement. However, an unexpected finding was the suppression effect found in the relationship between instructional clarity and mathematics achievement, mediated by mathematics confidence, warranting further studies (Kline, 2016; MacKinnon et al., 2000).

## Study II: The associations between student-perceived teaching quality and mathematics confidence and mathematics achievement

Using the Swedish TIMSS 2019 Grade 4 sample, Study II examined the relationships between aspects of student-perceived teaching quality and two learning outcomes, mathematics confidence and mathematics achievement, at both student and classroom levels. Teaching quality was conceptualised through the subdimensions of classroom management and instructional clarity. Previous research has found that teaching quality influences not only the individual student's learning but also the learning of the entire classroom (Darling-Hammond, 2021; Hattie, 2009), and is considered a classroom-level construct (Scherer & Gustafsson, 2015). Consequently, multilevel structural equation modelling (MSEM) was used to distinguish between student-level and classroom-level variance and examine the relationships between student-perceived teaching quality and, on the one hand, students' mathematics confidence, and on the other hand, mathematics achievement. The analyses also accounted for student background factors, such as gender, students' socioeconomic status (SES), which was proxied by the single measure of number of books at home, and immigration status, proxied by the single measure language spoken at home. The number of books has been validated as a reliable measure of SES ( Rutkowski, D., & Rutkowski, 2013) and language spoken at home as a reliable measure of immigration status (Rutkowski, D., & Rutkowski, 2013; Rutkowski, L., & Rutkowski, 2016).

The SEM models showed significant and positive relationships between both teaching quality aspects and the two outcomes, mathematics confidence and achievement. The results revealed a significant, substantial, and positive relationship between instructional clarity and mathematics confidence at both student and classroom levels. Classroom management was found to be positively and significantly related to mathematics achievement at the student level when accounting for student background factors. Moreover, student background factors were substantially related to mathematics achievement at both levels, while only minor relationships were observed with mathematics confidence, and only at the student level.

Students' socioeconomic status (SES) was substantially and significantly related to mathematics achievement at both the student and classroom levels, highlighting the importance of classroom composition for student learning outcomes. Notably, the strength of the relationship between classroom-level SES and mathematics achievement was initially high, with a standardised regression coefficient of .81. This indicates that students in classrooms with a higher average SES tended to perform substantially better than classrooms with lower SES. However, when accounting for both dimensions of teaching quality and student background factors in the final model, the strength of this relationship was reduced to .66 at the classroom level. At the student level, the strength of the relationship between SES and mathematics achievement remained stable, with a standardised regression coefficient of .30. These findings would suggest that part of the variance in classroom-level mathematics achievement, which was previously attributed to SES, can, in this model, be accounted for by differences in teaching quality and student background factors. This reduction in the strength of the relationship may indicate that teaching quality can partially mitigate the relationship between classroom-level SES and classroom-level mathematics achievement. Alternatively, the decrease may reflect a classroom dynamic, given that in high-achieving classrooms, which often require less time spent on maintaining order, teachers would be able to deliver higher teaching quality. These findings highlight the importance of considering classroom-level processes, when examining factors relating to student learning outcomes. Results revealed that teaching quality was significantly related to student learning outcomes, particularly at the classroom level.

## Study III: Does teaching quality matter for Nordic primary school students' mathematics confidence and mathematics achievement?

Study III expanded the scope of analysis and examined fourth-grade students' perceptions of teaching quality in the Nordic countries and its relationship to two outcomes: mathematics confidence and mathematics achievement. The Nordic context is suitable for analysis as these countries share the idea of a compensatory mission of the educational system, often referred to as the Nordic Model (Blossing et al., 2014; Lundahl, 2016).

Four Nordic countries participated in TIMSS 2019 at Grade 4 level: Denmark, Finland, Norway, and Sweden. The total sample comprised 15,839 students, nested in 966 classrooms, with an average class size of 16 students. By drawing on this data, Study III addressed three research questions related to teaching quality. First, it assessed whether the three latent constructs of classroom management, instructional clarity, and mathematics confidence, were measured equivalently across countries and their suitability for country comparisons. To assess cross-country comparability, a series of increasingly restrictive measurement invariance (MI) tests were conducted using multigroup confirmatory factor analysis (MGCFA), testing each of the three latent constructs separately. Second, given the hierarchical structure of TIMSS and the large intra-class correlation coefficients (ICC), the relationships between the teaching quality dimensions of classroom management and instructional clarity and the two outcomes: mathematics confidence and mathematics achievement were examined using multilevel structural equation modelling (MSEM). Third, Study III examined whether the two teaching quality aspects could mitigate the negative influence of lower socioeconomic background on student learning outcomes in the Nordic countries. The student background factors included measures of socioeconomic status, measured by the proxy number of books at home, and immigration status which was measured by the language spoken at home. These factors have been found to have high reliability and validity in previous research (Rolfe, 2021; Rutkowski, D., & Rutkowski, 2013).

Findings showed that measurement invariance supported both configural and metric invariance of all three scales of classroom management, instructional clarity, and mathematics confidence (using three positive items). However, scalar invariance was not supported. Nevertheless, as metric invariance was reached, correlational comparisons across countries can validly be drawn from the findings, thus allowing for cross-country comparisons (He et al., 2019).

The results revealed substantial and positive relationships between mathematics confidence and instructional clarity across all four Nordic countries, at both levels. These findings suggest that the variation in mathematics confidence between classrooms was, to a large extent, related to how the students perceived their teacher's instructional clarity. The higher the students assessed their teacher's instructions, the higher the mathematics confidence in the entire classroom. Student background factors were also significantly related to mathematics confidence. In all Nordic countries, positive relationships were found at the student level between SES and mathematics confidence. Interestingly, in Finland this relationship was significant and substantially positive at the classroom level, suggesting that differences in classroom mathematics confidence were related to the SES composition of the classroom. This finding raises concerns regarding selection bias in Finland, if students with similar SES are grouped together in the same classroom. One explanation could be the possibility of choosing specialised classes in mathematics, language, music, or sports which have been found to cause increasing differences between classroom achievement related to students' SES (Kosunen et al., 2020).

The results regarding the relationships between the two teaching quality factors and mathematics achievement were inconsistent across countries. Instructional clarity was positively related to mathematics achievement at the student level in Denmark, Finland, and Norway. The relationships between mathematics achievement and classroom management were significant and positive at the student level in Finland, Norway, and Sweden. At the classroom level, there were positive relationships between classroom management and mathematics achievement in Denmark and Sweden. This result would suggest that differences in classroom-level mathematics achievement in these two countries may partly be explained by students' belonging to classrooms perceived to be well-managed and calm.

Student background factors were found to relate significantly to mathematics achievement, aligning with previous research. The SES of the student showed a substantial relationship to mathematics achievement across all four countries at both student and classroom levels. The classroom-level relationships were substantial, indicating that Nordic classrooms with higher average SES tended to perform better on the mathematics assessment. This finding concurs with research from Atlay et al. (2019), which found that students with higher SES would benefit more from teaching quality than students with middle and lower SES. Thus, increasing the achievement gap. The language spoken at home was also significantly related to mathematics achievement at the student level across all countries. At the classroom level, this relationship was also positive in Norway and Sweden, indicating that classrooms with a higher proportion of students who spoke the test language at home performed better on mathematics assessments. These findings suggest mathematics performance is higher in classrooms where students are more language-aligned with many students speaking the test language at home compared to multilingual classrooms.

Taken together, Study III presented empirical findings of significant relationships between teaching quality aspects, such as classroom management and instructional clarity, and both mathematics confidence and mathematics achievement across all Nordic countries. However, the variation in results across countries suggests that no universal solution exists, rather, each country may need to address context-specific factors that could explain variations in student learning. Additionally, there were indications of educational inequity as the results indicated that differences in student learning outcomes were related to classroom composition effects, such as SES and language spoken at home. However, the findings also suggest that well-managed classrooms appeared to decrease the influence of SES, potentially narrowing the achievement gap. Therefore, by improving teachers' general pedagogical knowledge, particularly classroom management and instructional practices, student learning outcomes could be enhanced and thus also improve educational equity across the Nordic countries.

# Chapter 7 Discussion and conclusion

In this chapter, the main findings from the empirical studies are reviewed. The thesis had two primary aims. First, it addressed the construct validity of student questionnaires of TIMSS 2019 Grade 4, particularly the mixed-worded scale of mathematics confidence. Second, it examined the relationships between aspects of student-perceived teaching quality and student learning outcomes. By analysing classroom processes from the student perspective, the study provided insights into students' situated learning context. The discussion is framed by the proposed *conceptual model of teaching quality and the situated classroom context* (see Fig. 5), which integrates teaching quality aspects from the dynamic model, the TBD framework, and the conceptual framework of determinants of student leaning outcomes (Creemers & Kyriakides, 2010; Nilsen et al., 2016; Praetorius et al., 2018) with the situated expectancy-value theory (SEVT) (Eccles & Wigfield, 2023) as discussed in Chapter 2. Three themes have emerged across the empirical studies: 1) the situated classroom context and student learning outcomes; 2) teaching quality and student learning outcomes; and 3) mixed-worded items in secondary analyses.

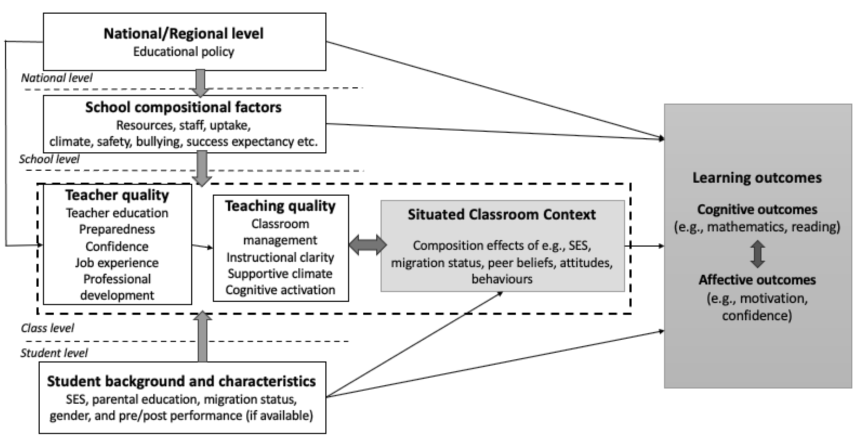


Figure 5 Conceptual model of teaching quality and the situated classroom context

## The situated classroom context and student learning outcomes

Findings from Study II and III revealed that classroom-level differences in mathematics achievement related to the classroom composition, supporting the *conceptual model of teaching quality and the situated classroom context*. There was evidence of composition effects, defined as the influence, within a classroom, of aggregated student background characteristics on learning outcomes (Burke & Sass, 2013). In particular, classroom-level socioeconomic status (SES) related strongly to mathematics achievement across all Nordic countries (Denmark, Finland, Norway, and Sweden). Teaching quality, operationalised through the dimensions of classroom management and instructional clarity, appeared to increase the influence of SES on mathematics achievement, thus increasing educational inequity in several Nordic countries. This finding aligns with previous research suggesting that high-SES classrooms benefit more from high-quality teaching due to these students having greater familiarity with the teaching, interaction, and communication styles of the teacher (Atlay et al., 2019; Hansson, 2012; Nilsen & Bergem, 2020). In contrast, students with limited access to home educational resources and lower levels of parental support could experience greater difficulties in interactions and comprehension (Sortkaer, 2019). For instance, Atlay et al. (2019) found that the achievement gap increased as high-SES students benefitted more from cognitively demanding instruction and a supportive learning environment than their low-SES peers. Similarly, Hansson (2012) found. mathematics instructions to be more beneficial for high-SES students in Sweden, contributing to a widening achievement gap.

Results from this thesis revealed substantial and positive relationships between classroom-level SES and mathematics achievement across all four Nordic countries. The higher the classroom-level SES, the higher the classroom-level mathematics achievement. In Finland, there were substantial significant relationships between classroom-level SES and mathematic achievement as well as mathematics confidence. One probable explanation could be the hidden stratification in Finland as parents can choose specialisation classes for their children, either in language, music, or sport (Kosunen et al., 2020). However, in Sweden, the relationship between classroom-level SES and mathematics achievement was substantially lower when accounting for the teaching quality factors. Consequently, it could be assumed that high-quality teaching significantly influenced Swedish classroom-level mathematics achievement and increased mathematics achievement in low-SES classrooms, subsequently narrowing the achievement gap.

Moreover, Study III provided evidence of linguistic composition effects on classroom-level mathematics achievement in Norway and Sweden. Specifically, mathematics achievement was significantly higher in classrooms where a majority of students spoke the test language at home. This finding suggests that linguistically homogenous classrooms may provide students with greater learning opportunities. By contrast, multilingual classrooms may require additional language support, an issue made more challenging given the shortage of certified 'Swedish as second language' teachers in Sweden (Skolverket, 2022a). Given there is less need for linguistic support in linguistically homogenous classrooms as fewer language-related instructional challenges may arise, while multilingual classrooms may require additional scaffolding and language support (Hansson, 2012).

Across the Nordic countries, results showed notable classroom differences for the relationships between classroom-level instructional clarity and mathematics confidence. The most confident mathematics classrooms were those with teachers who were perceived to provide high-quality instruction. This suggests that instructional clarity may contribute to widening the achievement gap, thus reflecting indications of pedagogical segregation across the Nordic countries (Hansson, 2012). Moreover, the findings indicate a composition effect of affective attitudes as classroom-level mathematics confidence may influence the learning environment. For instance, in classrooms with students less confident in mathematics, there may be more disruptive behaviour and subsequently fewer learning opportunities, while more confident classrooms would provide a calmer and more supportive learning environment (Vansteenkiste & Ryan, 2013). Likewise, poor teacher-student relationships and low student academic confidence and motivation may increase school absenteeism and problematic school attendance (Gren Landell, 2021; Skolverket, 2021). Subsequently, it could be proposed that higher classroom-level instructional clarity and mathematics confidence positively influence classroom-level mathematics achievement (Stankov & Lee, 2017) and school attendance (Skolinspektionen, 2016; Skolverket, 2021). Additionally, there may be composition effects relating to students' aptitude and learning abilities. In classrooms with lower-achieving students and less prior knowledge, more of the teaching and instruction would likely focus on meeting basic knowledge requirements, while classrooms with higher-achieving students may focus on more cognitively challenging tasks (Brophy, 2000).

The findings of the constituent studies in the thesis underscore the importance of accounting for the situated classroom context and diverse composition effects, such as SES, linguistic background, affective attitudes, and academic aptitude, when examining factors related to student learning outcomes. However, disentangling the relationships between various composition effects and learning outcomes is complex, given the possibility of reverse causality (Marsh & Craven, 2006). For instance, if low classroom-level achievement correlates with low teaching quality, it is unclear if ineffective teaching leads to lower achievement, or if lower-achieving classrooms perceive instructions as less clear and more difficult to understand and thus rate teaching quality lower (Nilsen et al., 2018; Ramazan et al., 2023). The inconsistent relationships found in Study III between teaching quality and student learning outcomes across the Nordic countries, despite the similar cultural and educational systems, suggest that country comparisons may be complex. While within-country analyses can be compared to other within-country analyses and provide valuable insights, the extent to which educational outcomes are comparable across countries participating in international large-scale assessments (ILSAs) remains open to debate.

## Teaching quality and student learning outcomes

Existing research suggests that teaching quality plays a significant role in influencing student learning outcomes (Darling-Hammond, 2021; Guerriero, 2017; Hattie, 2009). However, given the multidimensional nature of the teaching quality construct with a lack of consensus regarding its operationalisation and conceptualisation (Nilsen & Gustafsson, 2016; Senden et al., 2022), this thesis aimed to contribute empirical evidence to the field by drawing on TIMSS 2019 Grade 4 data and the *conceptual model of teaching quality and the situated classroom context*. The results showed the utility of the TIMSS student questionnaire items in capturing two valid subdimensions of teaching quality (classroom management and instructional clarity). Although this pragmatic conceptualisation may introduce the risk of construct under-representation (Praetorius et al., 2020), the strong significant relationships observed in the statistical analyses provided support that these measures validly reflect dimensions of the broader teaching quality construct. Findings revealed substantial differences across Nordic classrooms in student-perceived classroom management and instructional clarity, even greater than those relating to classroom-level SES and language spoken at home, and almost comparable in strength to classroom-level differences in mathematics achievement.

The results from Studies II and III revealed that instructional clarity related substantially to mathematics confidence, both at the individual level and especially at the classroom level. Across the Nordic countries, there were highly significant relationships between instructional clarity and mathematics achievement at the student level, aligning with previous research (Blömeke & Olsen, 2019; Hattie, 2009; Konstantinidou & Kyriakides, 2022; Titsworth et al., 2015). However, Titsworth et al. (2015) reported a significant negative relationship between teacher clarity and student learning outcomes, suggesting that students perceive instruction as excessive or too difficult. This negative relationship was found in Study I, when mathematics confidence was included as a mediating variable. Though, when student background factors were included in Studies II and III, this relationship was positive, confirming the findings by Blömeke and Olsen (2019).

In Study III, classroom-level mathematics achievement was positively and significantly related to classroom management in all Nordic countries except Finland. This suggests that classrooms with less reported disorderly behaviour showed higher levels of mathematics achievement. A plausible explanation is that in well-managed classrooms, teachers would have more opportunities to provide students with appropriate support and scaffolding. Consequently, a calmer and more orderly environment would give students more time on task and increase their learning opportunities (Doyle, 2013; Nilsen & Gustafsson, 2016). This result is contrary to findings by Kunter et al. (2007) who reported significant positive relationships between classroom management and mathematics interest at the individual level, but not at the classroom level.

Taken together, these findings suggest that teaching quality plays a meaningful role concerning student learning outcomes. The studies revealed a strong positive relationship between classroom management and mathematics confidence and an even stronger relationship between instructional clarity and mathematics confidence. This result is particularly relevant given that mathematics confidence is considered a strong predictor of student academic achievement (Lee & Stankov, 2018; Marsh et al., 2013; Marsh & Craven, 2006; Shavelson et al., 1976; Stankov & Lee, 2017).

## Method effects of mixed worded items

Secondary analysis relies on the quality, validity, and reliability of the data used. However, discussions regarding the operationalisations of latent constructs and the validity of items included in the measurement models are often lacking in the literature. For instance, latent constructs such as mathematics confidence have been conceptualised and operationalised in various ways across ILSA studies, often without sufficient verification or explanations for their operationalisation. Some secondary studies include all items from the questionnaire scale (Ivanova & Michaelides, 2022; Lee & Chen, 2019), while others separate positively and negatively worded items into distinct factors (Roszkowski & Soven, 2010; van Sonderen et al., 2013). Some studies use the transformed scale provided by TIMSS (e.g., Berger et al., 2020), whereas others include only a subset of positively worded items (Chen, 2022; Chen & Lu, 2022; Eklof, 2007; Min et al., 2016), and some rely on a single-item measure from a larger scale (Kim & Sidney, 2024; Stankov et al., 2014). These varying approaches to operationalisations highlight the need for construct validity studies and greater transparency in measurement modelling decisions.

In Study I, the construct validity of the mixed-worded scale measuring the latent construct mathematics confidence was examined. Results showed evidence of the method effect associated with the negatively worded items in the scale. The negatively phrased items could be more linguistically and cognitively demanding for primary school students, and the variations in phrasing, particularly the shift in the agent between items, may contribute to confusion (Michaelides, 2019; Reynolds et al., 2022). When interpreted through the lens of the *conceptual model of teaching quality and the situated classroom context*, these findings suggest that the classroom's composition influences the extent of method effects. For instance, in classrooms where students have a lower reading ability or language proficiency, the risk of misunderstanding negatively worded items is greater than in classrooms with students with higher levels of reading and language skills.

In addition, the semantic cross-cultural validation of the translated items showed some inconsistencies and there was evidence of translated items not being linguistically or conceptually equivalent (Ercikan, 1998; Ercikan & Pellegrino, 2017; Osborn, 2004; Upsing & Rittberger, 2018). Therefore, researchers conducting secondary analyses are encouraged to examine construct validity and to assess item content, wording, and translations when conceptualising latent constructs.

## Contribution and implications

This doctoral project contributes empirical knowledge to the research field of teaching quality and secondary analysis of ILSA data. Findings from the three empirical studies advance our understanding of the relationships between teaching quality aspects and students' mathematics achievement and mathematics confidence. Given that previous studies generally include mathematics achievement as an outcome variable, the inclusion of mathematics confidence provides a broader picture of the relationships in the situated classroom context, as outlined in the *conceptual model of teaching quality and the situated classroom context* (see Fig. 5). This proposed conceptual model visualises the complex and dynamic relationships that contribute to student learning outcomes and offers a lens for understanding these relations, the situated classroom context, and classroom composition effect.

These findings bring some implications for both practising teachers and teacher education. The results indicate that students in well-managed, calm classrooms are more likely to be provided with high-quality instruction, appropriate scaffolding, support, and greater learning opportunities. This highlights the need to strengthen teachers' general pedagogical knowledge, particularly regarding how to manage classroom diversity and support students in multilingual and lower-SES classrooms. Strengthening teachers' general pedagogical knowledge may give them effective strategies and tools needed to manage more diverse classrooms while building positive teacher-student interactions. Subsequently, such improvements could be assumed to increase students' mathematics confidence, as well as their academic achievement, and promote educational equity. In short, teachers equipped with these strategies and competencies may be able to provide more effective teaching within a supportive environment, potentially contributing to narrowing the achievement gap.

The findings also suggest that Swedish teacher education could better prepare future teachers to meet the needs of an increasingly diverse student population and the large variations between schools. Currently, students enrolled in Swedish teacher education typically complete their practical teacher training at the same school throughout their studies. While this approach gives continuity and a deeper understanding of the contextual factors from a single school, it may limit their exposure to a broader variety of schools. Given the increasingly diverse student population, it could be argued that students enrolled in teacher education would gain more pedagogical knowledge if they experienced a wider range of school contexts during their training.

These findings may also carry implications for policymakers. First, there is a need to offer working teachers professional development opportunities aimed at strengthening their general pedagogical knowledge and relational competencies. Second, results from this thesis highlight the presence of an achievement gap in Sweden, as reflected in the classroom-level differences related to SES and language spoken at home. Some classrooms are more diverse, characterised by low SES and highly multilingual groups, while other classrooms are more homogeneous, with high SES and most of the students speaking Swedish at home. Consequently, teachers would benefit from professional development opportunities focused on strategies for teaching diverse student groups, particularly in supporting second language learners. This is particularly important given the low number of certified teachers in 'Swedish as second language' (Skolverket, 2022b). These apparent differences across schools and classrooms highlight structural educational inequalities and a school selection which is influenced, to some extent, by students' SES or linguistic background. Subsequently, educational inequalities contribute to widening the achievement gap. Considering that education is a basic human right, a cornerstone of critical thinking, the agency of citizenship, democratic participation, and ultimately, democracy (Nations, 2021), these complex problems need to be addressed by policymakers and politicians.

## Limitations and future research

There are limitations to this doctoral thesis that warrant discussion. First, it is limited to two subdimensions of teaching quality - classroom management and instructional clarity - included in the student questionnaire of TIMSS 2019 Grade 4. Therefore, there could be a construct-under representation, as the teaching quality construct is not as broad as in previous studies (e.g., Bellens et al., 2019). Future studies could include additional dimensions of teaching quality aspects, such as those included in the expanded TIMSS 2023 framework of teaching quality aligning with the TBD framework. Second, this thesis is limited by the self-reported measures in the student questionnaire which could introduce mono-method bias. Further studies could include items from both student and teacher questionnaires to broaden the construct of teaching quality and improve the validity of the findings. Third, the cross-sectional design of data such as TIMSS does not allow for causal conclusions but only correlational inferences.

In light of the findings and the limitations discussed, further studies on teaching quality could include other age groups, academic subjects, or other cycles of ILSA studies. The expanded TIMSS 2023 framework of teaching quality provides important information for an extended follow-up study. Further secondary studies could be enriched by access to students' grades and performances on national assessments if provided by the Swedish National Agency for Education. To develop a more comprehensive understanding of teaching quality, additional studies could adopt a mixed-methods approach by combining quantitative ILSA data with a qualitative approach with classroom observations or interviews with teachers and students. Such triangulation would strengthen the validity of the findings.

## Conclusion

Although this thesis focused specifically on the mathematics classroom, the findings may well be relevant across other subject areas, given that the teaching quality aspects examined can be considered as part of general pedagogical knowledge rather than subject-specific knowledge. By leaning on the *conceptual model of teaching quality and the situated classroom context*, this thesis contributes empirical information on the relationships between teaching quality aspects and student learning outcomes, highlighting the situated classroom context. Considering the generalisability of results, although the Nordic countries share educational values and structures, there are differences in how education is organised and implemented. Therefore, cross-country comparisons must be done carefully, as the significant findings in this thesis could be different in other countries or cultures.

# Swedish summary

## Bakgrund

Utbildning är en grundläggande mänsklig rättighet som bidrar till individens självständighet och välmående vilket främjar samhällets utveckling (UNESCO, 2019). Vidare möjliggör utbildning att människor kan delta aktivt och ansvarsfullt i samhället, vilket är en garant för jämlikhet och demokrati (Nations, 2021; OECD, 2015). Enligt §4 i Skollagen (SFS 2010:800) ska skolan främja elevers utveckling och lärande samt förankra mänskliga rättigheter och demokratiska värderingar hos elever. Utbildning ska enligt Sveriges Skollag (SFS 2010:800) vara likvärdig, oberoende av elevers socioekonomiska förhållanden, etnicitet, skola, eller bostadsort. Skolan ska därför kompensera för skillnader i elevers olika förutsättningar så att alla elever ska erbjudas möjlighet till likvärdig utbildning.

Det är välkänt att elevers socioekonomiska status (SES) påverkar deras lärande och skolprestationer (Coleman, 1988; Sirin, 2005). Forskning visar att elevers SES har ökat i betydelse för hur de presterar akademiskt, vilket har bidragit till ökade kunskapsskillnader mellan elever, så väl i Sverige som i andra länder (Skolverket, 2024b; SOU 2019:40; Yang Hansen & Gustafsson, 2019). Idag lämnar allt fler elever i Sverige grundskolan utan behörighet till gymnasiet. Vårterminen 2024 saknade 16.3% av elever i årskurs 9 fullständiga betyg och gymnasial behörighet (Skolverket, 2024a). Vidare finns det idag en geografisk segregation, där elever med bäst förutsättningar återfinns i större storstadsområden på skolor med högre andel legitimerade lärare, medan elever med sämre förutsättningar återfinns på skolor på landsbygden och mindre städer, där det dessutom råder brist på utbildade och legitimerade lärare (SOU 2019:40).

En central fråga är hur den svenska skolan ska lyckas med det kompensatoriska uppdraget som formuleras i Skollagen och vilka faktorer som kan bidra till att öka alla elevers lärande och minska de ökande kunskapsskillnaderna mellan elever. Tidigare forskning har lyft fram läraren och undervisningens kvalitet som viktiga faktorer (t.ex. Brophy & Good, 1984; Hattie, 2009). Forskning om lärares betydelse för elevers lärande kan delas in i två huvudsakliga dimensioner, lärarkvalitet (teacher quality) och undervisningskvalitet (teaching quality) (Darling-Hammond, 2021). Lärarkvalitet avser lärarens formella utbildning, fortbildning, yrkeserfarenhet, ämneskompetens samt didaktiska förmågor. Undervisningskvalitet syftar på lärarens observerbara handlingar och agerande i klassrummet, så som hur aktiviteter organiseras och struktureras i klassrummet samt hur läraren ger instruktioner och interagerar med eleverna (Pianta & Hamre, 2009). Det är rimligt att anta att ett välorganiserat klassrum, med tydlig struktur och goda rutiner, skapar goda förutsättningar för studiero och ett positivt klassrumsklimat. Detta utgör i sin tur en grund för undervisningen av hög kvalitet, vilket kan främja elevers lärande och utveckling (Goe, 2007; Hattie & Yates, 2013). Läraren behöver god ledarskapsförmåga för att bygga positiva och förtroendefulla relationer med elever samt att skapa ett positivt klassrumsklimat (Aspelin, 2020). Samtidigt är det svårt att mäta undervisningskvalitet eftersom det är ett teoretiskt och mångfacetterat begrepp och tidigare forskning visar på varierande samband mellan undervisningskvalitet och elevers kognitiva och affektiva lärande (Blömeke et al., 2022; Klieme & Nilsen, 2022). Forskning om yngre elevers uppfattningar om undervisningskvalitet och hur dessa samvarierar med deras lärande är begränsad. Denna doktorsavhandling syftar därför till att bidra med analyser av empiriska samband mellan undervisningskvalitet och yngre elevers lärande, såväl kognitivt som affektivt.

## Syfte

Denna avhandling utgår från att undervisning och lärande är sociala och kulturella processer som sker i samspel mellan människor och som påverkas av det sammanhang de äger rum i, till exempel i klassrummet (Eccles & Wigfield, 2023; Vygotsky, 1978). För att undersöka dessa processer används data från den internationella storskaliga kunskapsmätningen Trends in Mathematics and Science Study (TIMSS) från 2019 för årskurs 4, där elever bland annat har skattat kvaliteten på matematikundervisningen i klassrummet. Med undervisningskvalitet avses här lärarens observerbara handlingar i klassrummet vilket är ett multidimensionellt begrepp som i tidigare forskning har beskrivits och konceptualiserats på åtskilliga sätt (Bellens et al., 2019; Senden et al., 2022). I avhandlingen undersöks matematikundervisning eftersom det i första hand är ett skolämne som undervisas av en lärare i en klassrumssituation, till skillnad från exempelvis engelska (Sundqvist & Sylvén, 2016). Trots att det finns ämnesspecifika strategier läraren använder för att ge matematikinstruktioner (Cohen et al., 2003; Lester, 2007), kan det antas att dessa strategier är av mer generell karaktär i årskurs 4 jämfört med årskurs 8. I denna avhandling har undervisningskvalitet pragmatiskt operationaliserats utifrån de två dimensioner av elevers upplevda undervisningskvalitet som mäts i TIMSS 2019:s elevenkäter. Dessa två dimensioner är, dels lärarens organisation i klassrummet och hur studiero skapas i undervisningen (Organisering), dels lärarens tydliga instruktioner, i form av förklaringar, stöd som ges, framåtsyftande återkopplingar och hur lagom svåra utmaningar erbjuds i undervisningen (Tydliga instruktioner). Vidare analyseras hur elevernas upplevda undervisningskvalitet relaterar till deras självförtroende (Självförtroende) och prestation i matematik (Prestation).

Denna sammanläggningsavhandling består av en kappa, som är en sammanfattande och förklarande text, samt tre empiriska studier. Datamaterialet utgörs av TIMSS 2019, för årskurs 4, i Sverige och Norden. **Studie I** innefattar TIMSS data för elever i Sverige och syftar till att undersöka validitet och reliabilitet för de tre latenta variablerna (konstrukt) som används i avhandling (Organisering, Tydliga instruktioner och Självförtroende). Ett latent konstrukt avser en teoretisk, icke-observerbar företeelse, som till exempel motivation eller rädsla. Man kan mäta ett latent konstrukt genom att kombinera flera enkätfrågor (indikatorer) som formar ett teoretiskt begrepp. Fokus i Studie I låg på att analysera de nio indikatorer som ingår i skalan för självförtroende i matematik, varav fyra är positivt formulerade och fem negativt formulerade. Syftet var att undersöka eventuella metodeffekter eller svarsbias kopplade till skalans blandade format. Skalan bestod av indikatorer med såväl positivt formulerade påståenden (exempelvis *Matematik är roligt*) som dubbelt negerande påståenden (exempelvis *Matematik är inte tråkigt*). Genom konfirmatorisk faktoranalys (CFA) jämfördes olika mätmodeller och operationaliseringar av det latenta konstruktet. Analyserna gav på en valid mätmodell av konstruktet självförtroende i matematik (Självförtroende). Med hjälp av strukturell ekvationsmodellering (SEM) kopplades därefter flera mätmodeller samman för att undersöka sambanden mellan två dimensioner av undervisningskvalitet (Organisering och Tydliga instruktioner) och prestation i matematik (Prestation) samt om dessa samband medierades via självförtroende i matematik (Självförtroende)

**Studie II** undersökte sambanden mellan elevers upplevelse av organisation i klassrummet och studiero, lärares tydliga instruktioner och två utfallsvariabler: självförtroende i matematik och prestation i matematik. Analyserna baserades på TIMSS 2019 för årskurs 4 i Sverige. Resultaten visade på betydande skillnader mellan klassrum och givet datas hierarkiska struktur genomfördes flernivåanalys med hjälp av strukturell ekvationsmodellering (MSEM). Genom att använda flernivåanalys delas variationen upp i två delar: individer och klassrummet. Detta möjliggör analyser som hanterar skillnader mellan individer såväl som skillnader mellan klassrum (lärares undervisningskvalitet).

**Studie III** vidgades till att omfatta de fyra nordiska länder som deltog i TIMSS 2019 med årskurs 4: Danmark, Finland, Norge och Sverige. Först analyserades huruvida de latenta konstrukten (Organisering, Tydliga instruktioner och Självförtroende) uppfattades på liknande sätt av eleverna i de olika länderna och om det därmed var möjligt att göra jämförelser mellan länderna. Dessa undersökningar genomfördes med hjälp av mätvariationsanalyser (Measurement invariance, MI) med multi-grupp konfirmatorisk faktoranalys (Multigroup confirmatory factor analysis, MGCFA). Med hjälp av flernivå SEM (MSEM), undersöktes därefter, separat för varje land, sambanden mellan variablerna organisering av klassrummet och studiero och lärares tydliga instruktioner och två utfallsvariabler, självförtroende i matematik och prestation i matematik.

De övergripande forskningsfrågorna för avhandlingen var:

1. Hur väl speglar TIMSS enkätfrågor de latenta konstrukten självförtroende i matematik och undervisningskvalitet genom dimensionerna organisering av undervisning och tydliga instruktioner?
2. Hur ser sambanden ut mellan undervisningskvalitet och elevers självförtroende och prestation i matematik och finns det skillnader för olika elevgrupper med avseende på socioekonomisk bakgrund, hemspråk och kön?
3. I vilken utsträckning finns det klassrumsskillnader beträffande sambanden mellan undervisningskvalitet och självförtroende och prestation i matematik?
4. I vilken utsträckning kan undervisning av hög kvalitet minska en negativ påverkan från elevers bakgrundsfaktorer, med avseende på socioekonomisk bakgrund, hemspråk och kön, och således minska kunskapsskillnaderna mellan elevgrupper?

## Teoretiskt ramverk

Denna avhandling återfinns inom forskningsfältet *educational effectiveness research* (EER). Inom detta forskningsfält beskrivs utbildning som hierarkiskt organiserad över flera nivåer (individ, klassrum, skola och nation) och man undersöker bland annat faktorer som kan förklara elevers lärande och prestationer (Kyriakides et al., 2020). Flera teoretiska ramverk beskriver det hierarkiska utbildningssystemet och faktorer som kan påverka elevers kognitiva och affektiva lärande, däribland *dynamic model* (Creemers & Kyriakides, 2007). Enligt detta ramverk är individnivån och klassrumsnivån de mest avgörande nivåerna för elevers lärande (Baumert et al., 2010). På klassrumsnivån återfinns faktorer som lärarkvalitet (exempelvis utbildning och erfarenhet) samt lärarens undervisningskvalitet, till exempel lärarens observerbara handlingar i klassrummet (se Figur 1). Undervisningskvalitet är ett teoretiskt, multidimensionellt begrepp som består av exempelvis organisering av aktiviteter i klassrummet och studiero, lärarens tydliga instruktioner, kognitiv aktivering av eleverna samt ett stöttande klassrumsklimat (Praetorius et al., 2018). Som tidigare nämnts inkluderar elevenkäten i TIMSS 2019 två dimensioner för att spegla undervisningskvalitet: organisering i klassrummet och studiero samt lärares tydliga instruktioner. Således görs en pragmatisk operationalisering av begreppet undervisningskvalitet i denna avhandling. För att beakta den individuella variationen inkluderas bakgrundsfaktorer som SES, hemspråk och kön i analyserna.

Lärande är en sociokulturell process som sker i en klassrumskontext och denna närmiljö påverkar elevers upplevelser, tolkningar, motivation och lärande (Eccles & Wigfield, 2023). Med hjälp av ramverket *situated expectancy-value* (SEVT) undersöker denna avhandling den situerade lärmiljön i klassrummet vilken påverkar elevers affektiva och kognitiva lärande. I SEVT-ramverket beskriver Eccles and Wigfield (2023) de intrikata sambanden mellan olika faktorer som påverkar elevers motivation, lärande och prestation. Elever befinner sig i klassrum som utgör en kulturell miljö där de tolkar händelser och interaktioner utifrån sin bakgrund, förmåga och erfarenhet. Dessa tolkningar kan påverka deras självförtroende och tilltro till den egna förmågan i olika grad. Elevers lärande påverkas således inte enbart av deras individuella förutsättningar och intressen, utan även klassrumsmiljön påverkar både enskilda elevers och hela klassens beteenden, attityder och upplevd förmåga. Vidare bedömer elever de uppgifter som ska utföras utifrån hur viktiga och intressanta dessa anses vara samt hur troligt det är att de ska lyckas med uppgifterna.

Under mitt doktorandprojekt har jag dock saknat ett teoretiskt ramverk som integrerar den situerade klassrumskontexten med lärarens undervisningskvalitet, i syfta att undersöka och förklara faktorer som relaterar till elevers lärande. Därför presenteras i avhandlingen en konceptuell model, *conceptual model of teaching quality and the situated classroom context,* (se Fig. 5) som inkluderar både undervisningskvalitet och den situerade klassrums-kontexten. Denna modell utgår från ramverken *dynamic model* och TBD (Creemers & Kyriakides, 2010; Nilsen et al., 2016;

Praetorius et al., 2018) samt SEVT (Eccles & Wigfield, 2023), och är tänkt att erbjuda en struktur för att analysera sambanden mellan olika faktorerna som kan förklara elevers lärande.

## Material och metod

TIMSS är en internationell storskalig kunskapsmätning (ILSA) som administreras av International Association for the Evaluation of Educational Achievement (IEA). Sedan 1995 genomförs TIMSS vart fjärde år och testar elevers kunskaper i matematik och naturorienterande ämnen i årskurs 4 och 8. Dessutom besvarar elever, lärare och rektorer enkäter för att samla in kontextuell information som kan användas för att förstå hur olika bakgrundsfaktorer kan förklara undervisning och lärande. Enkäter ges även till föräldrar i årskurs 4. För att säkerställa ett randomiserat och representativt urval väljs skolor ut och sedan en till två klasser på skolan där alla elever i klassrummet deltar (Mullis & Martin, 2017). Eftersom TIMSS inkluderar alla elever i ett klassrum möjliggörs analyser där såväl provresultat och enkätdata kan aggregeras till klasssrumsnivå så att det är möjligt att undersöka klassrummets betydelse för elevers lärande (Scherer & Gustafsson, 2015). Urvalet till Studie I och II baseras på TIMSS 2019, årskurs 4 i Sverige ($N$ = 3 965). I studie III ingår även Danmark ($N$ = 3 227), Finland ($N$ = 4 696) och Norge ($N$ = 3 951) med årskurs 4.

### Variabler

Avhandlingen baseras på data från TIMSS 2019 årskurs 4 som består av enkätfrågor om elevers bakgrund, deras lärarende och motivation samt elevers prestation i matematik. De enkätsvar som används i avhandlingen reflekterar elevers rapporterade uppfattning om a) organisering av aktiviteter i klassrummet och studiero, b) lärarens tydliga instruktioner, samt c) självförtroende i matematik. Följande variabler användes i de tre studierna:

- **Organisering** avser elevers uppfattningar om organisering av klassrummet, studiero, störande moment och ordning under matematiklektionen. Skalan består av sex indikatorer, exempelvis *Eleverna lyssnar inte på vad läraren säger* och *Det är för stökigt för att kunna arbeta bra*. Eleverna anger i vilken utsträckning de instämmer i respektive påstående på en fyrgradig Likertskala (1–4) (Varje eller nästan varje lektion - Ungefär hälften av lektionerna - Vissa lektioner - Aldrig). Höga värden indikerar bättre upplevd organisering och studiero.

- **Tydliga instruktioner** avser elevernas uppfattningar om lärarens tydlighet i sina instruktioner, förmåga att tydligt förklara och omformulera sig och framåtsyftande återkoppling för att öka elevernas förståelse under matematiklektionen. Skalan består av fem indikatorer, exempelvis *Det är lätt att förstå min* lärare och *Min lärare gör olika saker för att hjälpa oss att förstå*. Eleverna anger i vilken utsträckning de instämmer i respektive påstående på en fyrgradig Likertskala (1–4) (stämmer precis - stämmer ganska bra - stämmer inte så bra - stämmer inte alls). Skalan har reverserats så att högre värden motsvarar mer positiva svar.
- **Självförtroende i matematik** avser elevers uppfattningar om sin egen matematikförmåga och i vilken grad de tror sig kunna lösa matematikuppgifter. Skalan omfattar nio indikatorer med både positivt och negativt formulerade påståenden. Resultat från Studie I visade på metodeffekt och svarsbias till följd av de negativt formulerade indikatorerna. Därför användes ett reviderat latent konstrukt bestående av tre positivt formulerade indikatorer, såsom: *Jag lär mig snabbt i matematik, Jag är bra på att lösa svåra matematikuppgifter*. Eleverna anger i vilken utsträckning de instämmer i respektive påstående på en fyrgradig Likertskala (1–4) (stämmer precis - stämmer ganska bra - stämmer inte så bra - stämmer inte alls). De positivt formulerade påståendena har reverserats så att högre värden motsvarar positiva svar.
- **Prestation i matematik** mäts genom elevers prestation på TIMSS matematikuppgifter inom områdena taluppfattning och aritmetik, mätningar och geometri samt statistik. Elevernas prestation representeras av fem plausibla värden, vilka samtliga ingår i de statistiska analyserna genom imputation i Mplus.
- **Socioekonomisk status** (SES) avser föräldrars eller vårdnadshavares sociala och ekonomiska status och utbildningsnivå. TIMSS innehåller olika mått för att mäta SES. I denna avhandling används ett proxy-mått baserat på det antal böcker eleverna uppger att de har i hemmet (BOOKS). Detta mått har visat sig vara reliabelt och av hög validitet i västerländsk kontext (Nilsen & Gustafsson, 2016; Rolfe, 2021). Antalet böcker elever uppger mäts på en femgradig Likertskala: 1 = 0–10; 2 = 11–25; 3 = 26–100; 4 = 101–200; 5 = mer än 200.
- **Hemspråk** avser elevernas språkliga bakgrund och i vilken mån eleverna talar provspråket i hemmiljön (LANG). Måttet har visat sig vara en reliabel proxy-variabel för att mäta immigrationsbakgrund (Rutkowski, L., &

Rutkowski, 2016). Det bygger på en fyrgradig Likertskala: 1= aldrig; 2 = ibland; 3 = nästan alltid; 4 = alltid.

- **Kön** består av en binär variabel (dummy kodad) där elevens kön kodats 0 för pojkar och 1 för flickor.

## Analysmetod

Avhandlingen bygger på kvantitativa metoder och utgör en sekundäranalys av befintliga ILSA data från TIMSS 2019. Data förbereddes med hjälp av IDB Analyzer som IEA tillhandahåller på den offentliga webbplattformen (IEA, nd). Därefter användes SPSS (version 27) för inledande deskriptiv statistik samt reliabilitetsanalyser. För mer avancerade analysmetoder överfördes data till Mplus version 8.6 (Muthén & Muthén, 2017).

Med hjälp av konfirmatorisk faktoranalys (CFA) i Mplus skapas mätmodeller för att analysera faktorstrukturen för indikatorerna i de tre latenta konstrukten Organisering, Tydliga instruktioner och Självförtroende i matematik. I en CFA får varje indikator en faktorladdning mellan -1 och 1 som visar hur starkt den associerar med det latenta konstruktet. En faktorladdning över 0.70 (eller -.70) indikerar en god koppling mellan indikatorn och det latenta konstruktet. Ett faktorladdning på 0 visar att indikatorn inte mäter någon del av det underliggande konstruktet, medan ett faktorladdning på 1 indikerar en fullständig överlappning mellan indikatorn och konstruktet (Brown, 2015). Konfirmatorisk faktoranalys används för att säkerställa att det latenta konstruktet har hög validitet och reliabilitet vid operationalisering, det vill säga vid ”byggandet” av konstruktet med indikatorer från enkätdata. Vid jämförande analyser mellan länder, där latenta konstrukt ingår, behövs analyser för att kontrollera att konstrukten uppfattas likvärdigt av respondenter och är jämförbara mellan länder. Detta undersöks med hjälp av analyser av mätvariation (Measurement Invariance, MI) med multigrupp konfirmatorisk faktoranalys (Multi group confirmatory factor analysis, MGCFA).

De mätmodeller som har validerats med konfirmatorisk faktoranalys kombineras sedan med utfalls- och bakgrundsvariabler i en strukturell ekvationsmodellering (SEM). I Studie I användes strukturell ekvationsmodellering (SEM) för att undersöka om självförtroende i matematik fungerar som en medierande faktor i sambandet mellan matematikprestation och två dimensioner av undervisningskvalitet, organisering av klassrummet och studiero i den ena modell och i andra modellen lärarens tydliga instruktioner (Brown, 2015; Kline, 2016). För att beakta den hierarkiska datastrukturen där elever är samlade i klasser, samt de höga interkorrelationerna mellan indikatorerna (ICC), tillämpades flernivå-SEM (MSEM) i Studie II och III (Hox, 2002).

# Resultat

## Studie I

Asp, L. Validating the mixed-worded mathematics confidence scale in TIMSS 2019: Examining its relationships to teaching quality and mathematics achievement. (inskickad till *International Journal of Educational Research,* under granskning).

Studie I undersöker validiteten och reliabiliteten av skalan som mäter elevers självförtroende i matematik i TIMSS 2019. Skalan består av nio indikatorer av blandat format, fyra är positivt formulerade och fem negativt. Tidigare forskning som undersökt elevers självförtroende i matematik med TIMSS data har operationaliserat det latenta konstruktet på olika sätt. Vissa studier har inkluderat samtliga indikatorer från TIMSS skala (Ivanova & Michaelides, 2022; Lee & Chen, 2019) och andra studier har delat upp skalan i positiva och negativa faktorer (Roszkowski & Soven, 2010; van Sonderen et al., 2013). Det finns även studier som har använt en tregradig skala som TIMSS tillhandahåller (e.g., Berger et al., 2020). Vidare har andra studier använt fyra positiva indikatorer (Chen, 2022; Chen & Lu, 2022; Eklof, 2007; Min et al., 2016), eller enbart en enskild indikator utav flera i skalan (Kim & Sidney, 2024; Stankov et al., 2014). Få studier redogör för empiriska eller teoretiska överväganden som ligger till grund för hur det latenta konstruktet operationaliserats. Därför analyserar denna studie konstruktvaliditet för det latenta konstruktet självförtroende i matematik och bidrar med empirisk information till forskningsfältet.

Resultaten visade på metodeffekt kopplade till de negativt formulerade indikatorerna. En möjlig förklaring kan vara att dessa indikatorer kan uppfattas av elever som mer krävande och medföra högre kognitiv belastning. De översatta indikatorerna undersöktes i en semantisk valideringsprocess. Analyserna visade att översättningen från engelska till svenska för fyra av indikatorerna inte var lingvistiskt ekvivalenta. Detta skulle kunna leda till att elever tolkar indikatorerna på olika sätt mellan de två kulturella kontexterna (Ercikan, 1998).

Studie I använde en bi-faktor ESEM-analys (exploratory structural equation modelling) där flera indikatorer hade faktorladdningar under 0.4, vilket är ett tecken på låg reliabilitet och validitet. Baserat på dessa resultat operationaliserades det latenta konstruktet självförtroende i matematik med tre positivt formulerade indikatorer. Denna faktor användes sedan i de empiriska studierna.

Vidare undersöktes i Studie I relationerna mellan organisering i klassrummet och studiero, lärarens tydliga instruktioner och prestation i matematik med hjälp av en SEM-analys och där självförtroende i matematik inkluderades som en medierande variabel. I linje med tidigare forskning var relationen mellan självförtroende i matematik och prestation i matematik stark (Lee & Stankov, 2018; Marsh & Craven, 2006). Organisering i klassrummet och studiero visade ett positivt och signifikant samband med prestation i matematik, såväl direkt som indirekt. Lärarens tydliga instruktioner uppvisade däremot två kontrasterande förhållanden till prestation i matematik, en signifikant negativ direkt effekt och en signifikant positiv indirekt effekt via självförtroende i matematik. Observera att termen 'effekt' i detta sammanhang inte avser kausala samband, utan används deskriptivt inom ramen för medierande SEM modeller som enbart analyserar samband . Denna kontrasterande relation indikerar en så kallad suppressoreffekt. Suppressoreffekt uppstår när två variabler interagerar och påverkar relationen mellan den ena variabeln och en utfallsvariabel (Kline, 2016; MacKinnon et al., 2000). I den initiala modellen utan medieringsvariabeln var sambandet signifikant positivt mellan lärares tydliga instruktioner och prestation i matematik. Däremot, när självförtroende i matematik inkluderades i modellen som medierande variabel utövades en reciprok suppressoreffekt, vilket ökade den prediktiva validiteten av lärares tydliga instruktioner (Watson et al., 2013). Förklaringen kan vara att den medierande variabeln kontrollerar för irrelevant varians i variabeln för lärares tydliga instruktioner, vilket skulle tydliggöra att det finns ett starkare underliggande samband. Utan den medierande faktorn skulle det faktiska sambandet mellan lärares tydliga instruktioner och självförtroende i matematik varit dolt och felaktiga tolkningar skulle kunna dras.

## Studie II

Asp, L., Klapp, A., and Rolfe, V. The associations between student-perceived teaching quality and students' mathematics confidence and mathematics achievement: A study of Swedish Grade 4 TIMSS 2019. Inskickad till *Instructional Science* (under slutgranskning inför publicering).

Med teoretisk utgångspunkt i ramverket TBD undersöks i Studie II sambanden mellan undervisningskvalitet och elevers kognitiva och affektiva lärande. Analyserna genomfördes med hjälp av flernivå-SEM (MSEM) givet den hierarkiska strukturen i data och de höga intraklass koefficienterna (ICC) som visades i variablerna för Prestation i matematik, Hemspråk, Organisation, och Tydliga instruktioner. Med hjälp av MSEM kan variationen på klassrumsnivå skiljas från variationen på elevnivå. I två olika MSEM inkluderades först självförtroende i matematik som utfallsvariabel och därefter prestation i matematik som utfallsvariabel. I MSEM modellerna ingick även bakgrundsfaktorer för SES, hemspråk samt kön.

Resultaten visade att lärares tydliga instruktioner var starkast relaterat med elevers självförtroende i matematik, både på elevnivå och på klassrumsnivå. Bakgrundsfaktorerna visade enbart signifikanta samband på elevnivån. När prestation i matematik var utfallsvariabel uppvisade organisering av klassrum och studiero på signifikanta samband både på elevnivå och klassrumsnivå. Dessa resultat visar på stora skillnader mellan klassrum beträffande prestation i matematik, ju lugnare klassrumsmiljön upplevdes av eleverna i klassen, desto högre prestation i matematik för klassen. Det starkaste sambandet med prestation i matematik, både på elevnivå och klasssrumsnivå, visades för SES. I den initiala modellen hade SES ett starkt signifikant samband med prestationen i matematik med en standardiserad regressions koefficient på .81 på klassrumsnivå. Däremot i den sista modellen som inkluderade alla faktorer (Organisation, Tydliga instruktioner, Självförtroende och Hemspråk), sjönk den standardiserade regressions koefficienten för sambandet mellan SES och prestation i matematik till .66 på klassrumsnivå. Detta resultat visade att SES minskade i betydelse för elevers prestation i matematik. En tolkning kan vara att lärarens undervisningskvalitet kompenserade för elevers lägre SES. Även hemspråk visade på ett signifikant samband med prestation i matematik, både på elev- och klassrumsnivå. Resultaten visade att klassrum som är mer språkligt homogena, det vill säga där en majoritet av eleverna talar svenska i hemmet, tenderade att ha högre prestation i matematik. En jämförelse mellan den initiala modellen med hemspråk och prestation i matematik på klassrumsnivå, visar att den standardiserad regressions koefficienten var .65 och i den sista modellen, med alla faktorer inkluderade, var den standardiserad regressions koefficient .30 mellan hemspråk prestationen i matematik. Detta resultat visar att betydelsen av att tala svenska i hemmet minskade. En möjlig tolkning är att lärarens undervisningskvalitet kompenserade för elever som har ett annat hemspråk.

## Studie III

Asp, L., Klapp, A., and Rosén, M. Does Teaching Quality Matter for Nordic Primary School Students' Mathematics Confidence and Mathematics Achievement? A Multilevel Structural Equation Analysis of Nordic TIMSS 2019 Grade 4 Data. *Large-Scale Assessments in Education*, 2025, Vol. 13, Article 7, https://doi.org/10.1186/s40536-025-00238-x

I Studie III vidgas fokus till att inkludera de fyra nordiska länderna Danmark, Finland, Norge, och Sverige som deltog med årskurs 4 i TIMSS 2019. Dessa länder är lämpande för jämförande analyser eftersom de har liknande skolsystem och en gemensam syn på utbildning och lärande, ofta benämnd som den nordiska skolmodellen (Blossing et al., 2014).

Resultat från analyserna av mätvariation (MI) visade att eleverna i de fyra nordiska länderna uppfattade enkätfrågorna på likvärdigt sätt, vilket möjliggjorde jämförande analyser mellan länderna. Med hjälp av MSEM undersöktes sambanden mellan de två undervisningskvalitetsfaktorerna (organisering av klassrummet och studiero och lärarens tydliga instruktioner) och två utfallsvariabler: självförtroende i matematik och prestationen i matematik på individ- och klassrumsnivå separat för varje land.

Resultaten visade att sambandet mellan lärares tydliga instruktioner och självförtroende i matematik var signifikanta på båda nivåer i alla fyra länder. Ju tydligare undervisning elever upplevde i klassrummet, desto högre var klassens aggregerade självförtroende i matematik. SES relaterade signifikant positivt till självförtroende i matematik på elevnivå i alla länder. Däremot visade enbart Finland på ett signifikant samband på klassrumsnivå mellan SES och självförtroende i matematik. Detta kan förklaras av tidigare forskning som har pekat på att skillnader mellan klassrum beträffande matematikprestation kan relatera till en dold tidig selektion där elever i Finland väljer specialiserade klasser redan i årskurs 1 (Kosunen et al., 2020).

Med prestation i matematik som utfallsvariabel var resultaten varierande mellan de nordiska länderna. För alla länder utom Danmark fanns signifikanta samband mellan organisering av klassrummet och studiero och prestation i matematik på elevnivå. På klassrumsnivå hade enbart Danmark och Sverige signifikanta samband mellan dessa faktorer. Detta resultat visade att klassens prestation i matematik relaterade till hur välorganiserat klassrummet uppleves av hela klassen. Ju mindre störande moment, desto högre prestation i matematik av hela klassen i

Sverige och Danmark. Lärares tydliga instruktioner relaterade signifikant till prestation i matematik i alla länder utom Sverige på elevnivå. Vidare var klassens prestation i matematik relaterad till hur språkligt homogent klassrummet var i alla länder utom Danmark. SES hade ett starkt samband till prestation i matematik på både elevnivå och klassrumsnivå i alla länder.

## Diskussion och slutsatser

Syftet med föreliggande avhandling har varit att undersöka huruvida lärarens undervisningskvalitet, reflekterad genom elevers upplevelser av lärares organisering av klassrummet och lärarens tydliga instruktioner, relaterar till elevers självförtroende och prestation i matematik. Ett reliabelt sätt att undersöka kvaliteten på lärarens undervisning och praktik är att aggregera elevers uppfattningar om lärarens observerbara handlingar i klassrummet.

Tre övergripande teman kan urskiljas utifrån studierna i avhandlingen. Det första gäller den situerade klassrumskontexten och kompositionseffekter relaterade till denna. Resultaten visar på signifikanta samband mellan elevers bakgrundsfaktorer (SES och hemspråk) och prestation i matematik. I de nordiska länder som undersöktes visade resultaten att utbildningen inte var likvärdig mellan klassrum. En betydande del av klassens prestation i matematik kunde förklaras av SES, hemspråk och hur elever upplever att läraren organiserar klassrummet och möjliggör studiero. Ju högre klass-SES, desto högre prestation i matematik för klassen. Denna relation var särskilt påtaglig i Finland. En förklaring kan finnas i den delvis dolda differentieringen då elever kan välja specialklasser i matematik, språk, instrument eller idrott (Kosunen et al., 2020). Överraskande visade resultaten för Sverige att undervisningskvalitet försvagade sambandet mellan SES och prestation i matematik på klassrumsnivå (se Studie II). Vidare visade resultaten att klassens SES och hemspråk relaterade till klassens prestation i matematik i betydligt högre utsträckning än till självförtroende i matematik. I flera av de nordiska länderna visade resultaten på starka och signifikanta samband mellan högt klass-SES och prestation i matematik på elevnivå men framför allt på klassrumsnivå, likaså för hemspråk. Dessa resultat bekräftar tidigare forskning som visat på att undervisningskvalitet bidrar till att skillnader i kunskaper ökar mellan elevgrupper i flera av de nordiska länderna (Atlay et al., 2019; Hansson, 2012; Nilsen & Bergem, 2020). Dock visade resultaten för Sverige på att lärares undervisningskvalitet kunde bidra till att minska sambandet mellan SES och hemspråk och prestation i matematik på klassrumsnivå. Vidare visade resultaten att i klassrum med högre självförtroende i matematik undervisade lärare som upplevdes av elever som mest tydliga i sina instruktioner och praktiker. Tidigare forskning har visat att högre självförtroende i matematik bidrar till högre prestation i matematik (Stankov & Lee, 2017). Därför borde det rimligtvis kunna antas att om klassens självförtroende i matematik skulle öka, skulle hela klassens prestation i matematik i sin tur påverkas positivt.

Det andra temat som framkommit i avhandlingsarbetet är att undervisningskvalitet spelar roll för elevers kognitiva och affektiva lärande. Ju högre upplevd organisering av klassrummet och studiero desto högre prestation i matematik på både elev- och klassrumsnivå. Det är rimligt att anta att ett välorganiserat och lugnare klassrum med högre studiero möjliggör att läraren kan fokusera på undervisningen med stöttande och utmanande uppgifter som är anpassade till elevers behov, vilket i sin tur tycks leda till bättre prestationer i matematik. I linje med tidigare forskning visar resultaten på ett positivt samband mellan lärares tydliga instruktioner (rapporterade av eleverna) och elevers självförtroende i matematik, såväl för den enskilde eleven som för hela klassens (Hattie, 2009; Titsworth et al., 2015). Emellertid har tidigare forskning även rapporterat om negativa samband mellan lärares tydliga instruktioner och prestation i matematik (Chen & Lu, 2022; Titsworth et al., 2015). De tre empiriska studierna i avhandlingen visade på kontrasterande resultat. I Studie I, när självförtroende i matematik ingick som en medierande variabel, fanns ett direkt negativt samband mellan lärares tydliga instruktioner och prestation i matematik men ett indirekt positivt medierat samband. Däremot, i Studie II och III, då elevers bakgrundsvariabler inkluderades, fanns enbart positiva samband mellan lärares tydliga instruktioner och matematikprestation.

Det tredje temat berör metodeffekter relaterade till de indikatorer som används för att operationalisera latenta konstrukt. Analyserna i Studie I visar på metodeffekter och svarsbias relaterade till de negativt formulerad indikatorerna i skalan för självförtroende i matematik. Dessa resultat bekräftar tidigare forskning om ökad lingvistiska komplexitet för negativt formulerade svarsalternativ (Michaelides, 2019; Reynolds et al., 2022). Framför allt är det tveksamt om elever i årskurs 4 hinner läsa och förstå indikatorer med skiftande svarsformuleringar då dessa bidrar till ökad kognitiv belastning. Studier vilka tillämpar sekundäranalyser av ILSA data rekommenderas att validera de latenta konstrukt som består av indikatorer av blandat format. Vidare, i överensstämmelse med tidigare forskning, visar resultaten att de översatta indikatorerna inte alltid är lingvistiskt ekvivalenta. Detta kan leda till att respondenter från olika språkliga och kulturella kontexter kan uppfatta och besvara enkätfrågor på olika sätt, vilket påverkar jämförbarheten mellan länder (Ercikan, 1998; Ercikan & Pellegrino, 2017; Osborn, 2004; Upsing & Rittberger, 2018).

# Implikationer

Även om denna avhandling har fokuserat på matematikklassrummet i årskurs 4, kan resultaten ha relevans även för andra ämnesområden. Detta eftersom de undersökta dimensionerna av undervisningskvalitet (organisering av klassrummet och studiero samt lärarens tydliga instruktioner) kan till stor del betraktas som generiska och till en mindre grad som ämnesspecifika. Resultat från studierna visar på fenomen som bör vara relevanta för olika aktörer inom det svenska utbildningssystemet att ta del av men resultaten kan även vara av värde för aktörer inom andra nationella kontexter så som de nordiska länderna.

Avhandlingen visar att det finns stora skillnader mellan klassrum beträffande prestation i matematik, socioekonomisk status (SES), immigrationsstatus (mätt genom elevens hemspråk) och undervisningskvalitet. Till exempel visar resultaten att klassrum där elever presterar väl i matematik, också kännetecknas av en hög genomsnittlig SES och av att eleverna upplever en hög kvalitet på lärarens undervisning. En möjlig åtgärd som avhandlingen lyfter fram för att minska kunskapsklyftorna mellan olika elevgrupper är vikten av lärarens allmänna pedagogiska förmåga att leda, strukturera och organisera klassrummet. Ökad allmänpedagogisk kompetens hos läraren skulle kunna ge elever bättre möjligheter att arbeta i ett välorganiserat och lugnt klassrum, vilket i sin tur skapar utrymme för läraren att fokusera på undervisning och instruktion av hög kvalitet, exempelvis tydliga, stöttande och stimulerande undervisningsaktiviteter med framåtsyftande återkoppling. Detta är av särskild vikt i klassrum med flerspråkiga elever och i klassrum där elever har lägre SES. Beslutsfattare på skol- och politisk nivå skulle kunna erbjuda yrkesverksamma lärare utvecklingsprogram för att stärka deras allmänna pedagogiska kunskaper och relationella färdigheter i syfte att öka kvaliteten på organisering av klassrumsaktiviteter. Vidare bör lärare erbjudas fortbildning och kunskap om hur de på lämpligt sätt kan stötta elever i flerspråkiga klassrum.

I en nyligen presenterad statlig offentlig utredning kring en reformerad lärarutbildning, betonas vikten av lärarens förmåga att skapa trygghet och studiero i klassrummet samt hur lärarstudenterna under lärarutbildningen bör träna på praktiska färdigheter inom sociala relationer, ledarskap och konflikthantering

(SOU 2024:81). Framtida lärare kommer att verka i klassrum med elever som uppvisar stor variation i förutsättningar, förmågor, motivation och språkkunskaper. För att arbeta i dessa alltmer komplexa klassrum bör lärarutbildningen ge blivande lärare kunskaper i pedagogiskt ledarskap samt relationell kompetens, med särskilt fokus på strategier för att organisera klassrumsaktiviteter som främjar studiero och som stöttar alla elevers lärande. Genom att stärka framtida lärares ledarskapsförmåga, allmänna pedagogiska kunskaper, och relationella färdigheter, kommer de att ha bättre förutsättningar för att framgångsrikt undervisa en alltmer diversifierade elevgrupp. Detta skulle exempelvis kunna ske inom den verksamhetsförlagda utbildningen (VFU) som lärarstudenter genomför på en skola. VFU-utbildningen kompletterar den universitetsförlagda utbildningen och ger den lärarstudenter möjlighet att lära känna en verksamhet på djupet och utveckla yrkeskompetens. VFU omfattar 20 veckor som är fördelade över lärarutbildningens terminer och målet är att den ska genomföras på samma skola. Detta upplägg ger visserligen fördjupade kunskaper om den aktuella skolans verksamhet och kultur, men innebär samtidigt att studenternas erfarenhet av den variation som finns mellan skolor begränsas. För att ge lärarstudenten en bredare inblick i variationen mellan klassrum och skolor och därmed förse lärarstudenten med fler möjligheter att utveckla en bredare yrkeskompetens, kan det vara värdefullt att genomföra VFU på olika skolor. Sedan 2022 ansvar svenska lärosäten för att etablera övningsskolor där lärarstudenterna får handledning av särskilt utbildade handledare som undervisar på övningsskolorna. Syftet är att höja kvaliteten i lärarutbildningen och samtidigt stärka skolornas kompetens-försörjning. Emellertid finns en risk att detta innebär att lärarstudenter inte ges tillfälle för VFU på skolor med varierande elevunderlag.

## Avslutande ord

Avslutningsvis kan det konstateras att denna doktorsavhandling, genom en sekundäranalys av ILSA-data, bidrar med empirisk information om sambanden mellan lärares undervisningskvalitet och elevers lärande. Avhandlingen lyfter fram betydelsen av lärarens roll i klassrummet för elevers lärande. För det första bidrar var och en av de tre studierna till en fördjupad förståelse av sambanden mellan undervisningskvalitet och elevers prestationer och självförtroende i matematik. Resultaten visar att ju lugnare klassrum och ju mer studiero eleverna upplever, desto högre blir deras prestation i matematik. Vidare framkommer att ju tydligare instruktioner från läraren som eleverna upplever, desto högre blir deras självförtroende i matematik. Dessutom belyser resultaten hur den situerade klassrumskontexten bidrar till hela klassens lärande. Elevsammansättningen i klassrummet påverkar såväl undervisningens kvalitet som elevers lärande och välbefinnande. Avhandlingen bidrar även med information beträffande betydelsen av konstruktvaliditet vid användandet av ILSA-skalor med blandat frågeformat. Dessutom presenteras en utvidgad konceptuell modell, *the conceptual model of teaching quality and the situated classroom context* (se Fig. 5), för att analysera hur lärarens undervisningskvalitet och den situerade klassrumskontexten kan förklara elevers kognitiva och affektiva lärande. Denna konceptuella model kan användas i studier för att undersöka hur den situerade klassrumskontexten kan förklara lärarande, för såväl den enskilde eleven som för hela klassen.

# References


Aditomo, A., & Köhler, C. (2020). Do student ratings provide reliable and valid information about teaching quality at the school level? Evaluating measures of science teaching in PISA 2015. *Educational Assessment, Evaluation and Accountability*, *32*(3), 275-310. https://doi.org/10.1007/s11092-020-09328-6

AERA. (2014). *Standards for educational and psychological testing*. American Educational Research Association.

Andersen, F. Ø. (2010). Danish and Finnish PISA results in a comparative, qualitative perspective: How can the stable and distinct differences between the Danish and Finnish PISA results be explained? *Educational Assessment, Evaluation and Accountability*, *22*(2), 159-175. https://doi.org/10.1007/s11092-010-9095-x

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling*, *16*(3), 397-438. https://doi.org/10.1080/10705510903008204

Aspelin, J. (2020). Teaching as a way of bonding: a contribution to the relational theory of teaching. *Educational Philosophy and Theory*, *53*(6), 588-596. https://doi.org/10.1080/00131857.2020.1798758

Atlay, C., Tieben, N., Hillmert, S., & Fauth, B. (2019). Instructional quality and achievement inequality: How effective is teaching in closing the social achievement gap? *Learning and Instruction*, *63*, Article 101211. https://doi.org/10.1016/j.learninstruc.2019.05.008

Balkis, M. (2018). Academic amotivation and intention to school dropout: the mediation role of academic achievement and absenteeism. *Asia Pacific journal of education*, *38*(2), 257-270. https://doi.org/10.1080/02188791.2018.1460258

Bandura, A. (1997). *Self-efficacy: the exercise of control*. W. H. Freeman.

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

Bellens, K., Van Damme, J., Van Den Noortgate, W., Wendt, H., & Nilsen, T. (2019). Instructional quality: catalyst or pitfall in educational systems' aim for high achievement and equity? An answer based on multilevel SEM analyses of TIMSS 2015 data in Flanders (Belgium), Germany, and Norway. *Large-Scale Assessments in Education*, *7*(1), 1-27. https://doi.org/10.1186/s40536-019-0069-2

Berger, N., Mackenzie, E., & Holmes, K. (2020). Positive attitudes towards mathematics and science are mutually beneficial for student achievement: a latent profile analysis of TIMSS 2015. *The Australian Educational Researcher*, *47*(3), 409-444. https://doi.org/10.1007/s13384-020-00379-8

Bernelius, V., & Vaattovaara, M. (2016). Choice and segregation in the 'most egalitarian' schools: Cumulative decline in urban schools and neighbourhoods of Helsinki, Finland. *Urban studies (Edinburgh, Scotland)*, *53*(15), 3155-3171. https://doi.org/10.1177/0042098015621441

Blikstad-Balas, M., Klette, K., & Tengberg, M. (Eds.). (2021). *Ways of analyzing teaching quality: Potentials and pitfalls*. https://doi.org/DOI: 10.18261/9788215045054-2021.

Blossing, U., Imsen, G., & Moos, L. (2014). *The Nordic education model: "A school for all" encounters neo-liberal policy* (2014 ed., Vol. 1). Dordrecht: Springer. https://doi.org/10.1007/978-94-007-7125-3

Blömeke, S., Hsieh, F.-J., Kaiser, G., & Schmidt, W. H. (2013). *International perspectives on teacher knowledge, beliefs and opportunities to learn: TEDS-M results* (2014 ed., Vol. 6). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-6437-8

Blömeke, S., Jentsch, A., Ross, N., Kaiser, G., & König, J. (2022). Opening up the black box: Teacher competence, instructional quality, and students' learning progress. *Learning and Instruction*, *79*, Article 101600. https://doi.org/10.1016/j.learninstruc.2022.101600
Blömeke, S., & Olsen, R. V. (2019). Consistency of results regarding teacher effects across subjects, school levels, outcomes and countries. *Teaching & Teacher Education*, *77*, 170-182. https://doi.org/10.1016/j.tate.2018.09.018
Blömeke, S., Olsen, R. V., & Ute, S. (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 outcome. Relationships across countries, cohorts and time.* (pp. 21–50). Springer. https://doi.org/doi: 10.1007/978-3-319-41252-8_2
Bourdieu, P. (1986). The forms of capital  In J. Richardson (Ed.), *Handbook of theory and research for the sociology of education* (pp. 241–258 ). Greenwood Press.
Brophy, J. (2000). *Teaching* (Vol. 1). International Academy of Education. https://files.eric.ed.gov/fulltext/ED440066.pdf
Brophy, J., & Good, T. (1984). *Teacher behavior and student achievement.* Institute for Research on Teaching, Michigan State University.
Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (Second edition. ed.). The Guilford Press.
Bruner, J. S. (1996). *The culture of education*. Cambridge, Mass. : Harvard University Press.
Buchholtz, N., Stuart, A., & Frønes, T. S. (2020). Equity, equality and diversity-putting educational justice in the Nordic model to a test. In T. S. Frønes, A. Pettersen, J. Radišić, & N. Buchholtz (Eds.), *Equity, Equality and Diversity in the Nordic Model of Education*. Springer. https://doi.org/https://doi.org/10.1007/978-3-030-61648-9_2
Burke, M. A., & Sass, T. R. (2013). Classroom peer effects and student achievement. *Journal of Labor Economics*, *31*(1), 51-82. https://doi.org/10.1086/666653
Calsyn, R. J., & Kenny, D. A. (1977). Self-concept of ability and perceived evaluation of others: Cause or effect of academic achievement? *Journal of Educational Psychology*, *69*(2), 136-145. https://doi.org/10.1037/0022-0663.69.2.136
Carroll, J. B. (1963). A model of school learning. *Teachers college record*, *64*(8), 1-9.
Charalambous, C. Y., & Praetorius, A.-K. (2022). Synthesizing collaborative reflections on classroom observation frameworks and reflecting on the necessity of synthesized frameworks. *Studies in Educational Evaluation*, *75*, 101202. https://doi.org/10.1016/j.stueduc.2022.101202
Chen, M., & Hastedt, D. (2022). The paradoxical relationship between students' non-cognitive factors and mathematics & science achievement using TIMSS 2015 dataset. *Studies in Educational Evaluation*, *73*, Article 101145. https://doi.org/10.1016/j.stueduc.2022.101145
Chen, X. (2022). The effects of individual- and class-level achievement on attitudes towards mathematics: An analysis of Hong Kong students using TIMSS 2019. *Studies in Educational Evaluation*, *72*, Article 101113. https://doi.org/10.1016/j.stueduc.2021.101113
Chen, X., & Lu, L. (2022). How classroom management and instructional clarity relate to students' academic emotions in Hong Kong and England: A multi-group analysis based on the control-value theory. *Learning and Individual Differences*, *98*, Article 102183. https://doi.org/10.1016/j.lindif.2022.102183
Chiu, M.-S. (2008). Achievements and self-concepts in a comparison of math and science: exploring the internal/external frame of reference model across 28 countries. *Educational research and evaluation*, *14*(3), 235-254. https://doi.org/10.1080/13803610802048858
Chyung, S. Y., Kennedy, M., & Campbell, I. (2018). Evidence-based survey design: The use of ascending or descending order of Likert-type response options. *Performance Improvement (International Society for Performance Improvement)*, *57*(9), 9-16. https://doi.org/10.1002/pfi.21800
Cohen, D. K., Raudenbush, S. W., & Ball, D. L. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis*, *25*(2), 119-142. https://doi.org/10.3102/01623737025002119

Coleman, J. S. (1988). Social capital in the creation of human capital. *The American Journal of Sociology*, *94*(1988), 95-120. https://doi.org/10.1086/228943

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of Educational Opportunity*. National Center for Educational Statistics.

Colston, H. L. (1999). "Not good" is "bad," but "not bad" is not "good": An analysis of three accounts of negation asymmetry. *Discourse Processes*, *28*(3), 237-256. https://doi.org/10.1080/01638539909545083

Creemers, B., & Kyriakides, L. (2007). *The Dynamics of educational effectiveness: A contribution to policy, practice and theory in contemporary schools*. Routledge. https://doi.org/10.4324/9780203939185

Creemers, B. P. M., & Kyriakides, L. (2010). Using the Dynamic Model to develop an evidence-based and theory-driven approach to school improvement. *Irish Educational Studies*, *29*(1), 5-23. https://doi.org/10.1080/03323310903522669

Danielson, C. (2007). *Enhancing professional practice: A framework for teaching* (2nd Edition ed.). Association for Supervision and Curriculum Development. https://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED509067&site=ehost-live

Darling-Hammond, L. (2000). Teacher quality and student achievement. A review of state policy evidence. *Education Policy Analysis Archives*, *8*(1), 8:1. https://epaa.asu.edu/index.php/epaa/article/view/392/515

Darling-Hammond, L. (2021). Defining teaching quality around the world. *European Journal of Teacher Education*, *44*(3), 295-308. https://doi.org/10.1080/02619768.2021.1919080

Darling-Hammond, L., Flook, L., Cook-Harvey, C., Barron, B., & Osher, D. (2020). Implications for educational practice of the science of learning and development. *Applied Developmental Science*, *24*(2), 97-140. https://doi.org/10.1080/10888691.2018.1537791

DiStefano, C., & Motl, R. W. (2006). Further investigating method effects associated with negatively worded items on self-report surveys. *Structural Equation Modeling*, *13*(3), 440-464. https://doi.org/10.1207/s15328007sem1303_6

Doyle, W. (2013). Ecological approaches to classroom management In C. M. Evertson & C. S. Weinstein (Eds.), *Handbook of Classroom Management* (pp. 107–136). Routledge. https://www.researchgate.net/publication/243771420_Classroom_organization_and_management

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, *53*(1), 109-132. https://doi.org/10.1146/annurev.psych.53.100901.135153

Eccles, J. S., & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology*, *61*, Article 101859. https://doi.org/10.1016/j.cedpsych.2020.101859

Eccles, J. S., & Wigfield, A. (2023). Expectancy-value theory to situated expectancy-value theory: Reflections on the legacy of 40+ years of working together. *Motivation Science*, *9*(1), 1-12. https://doi.org/10.1037/mot0000275

Eklof, H. (2006). Development and validation of scores from an instrument measuring student test-taking motivation. *Educational and Psychological Measurement*, *66*(4), 643-656. https://doi.org/10.1177/0013164405278574

Eklof, H. (2007). Self-concept and valuing of mathematics in TIMSS 2003: Scale structure and relation to performance in a Swedish setting. *Scandinavian Journal of Educational Research*, *51*(3), 297-313. https://doi.org/10.1080/00313830701356141

Elliott, J., Stankov, L., Lee, J., & Beckmann, J. F. (2019). What did PISA and TIMSS ever do for us? The potential of large scale datasets for understanding and improving educational practice. *Comparative Education*, *55*(1), 133-155. https://doi.org/10.1080/03050068.2018.1545386

Emmer, E. T., & Stough, L. M. (2001). Classroom management: A critical part of educational psychology, with implications for teacher education. *Educational Psychologist*, *36*(2), 103-112. https://doi.org/10.1207/S15326985EP3602_5

Ercikan, K. (1998). Translation effects in international assessments. *International Journal of Educational Research*, *29*, 543-553. https://doi.org/10.1016/S0883-0355(98)00047-0

Ercikan, K., & Pellegrino, J. W. (2017). *Validation of score meaning for the next generation of assessments: The use of response processes* (1st edition. ed.). London: Routledge. https://doi.org/10.4324/9781315708591

Ercikan, K., Roth, W.-M., & Asil, M. (2015). Cautions about inferences from international assessments: The case of PISA 2009. *Teachers College Record (1970)*, *117*(1), 1-28. https://doi.org/10.1177/016146811511700102

Fan, W., & Williams, C. (2018). The mediating role of student motivation in the linking of perceived school climate and achievement in reading and mathematics. *Frontiers in Education*, *3*, Article 50. https://doi.org/10.3389/feduc.2018.00050

Fauth, B., Decristan, J., Rieser, S., Klieme, E., & Büttner, G. (2014). Student ratings of teaching quality in primary school: Dimensions and prediction of student outcomes. *Learning and Instruction*, *29*, 1-9. https://doi.org/10.1016/j.learninstruc.2013.07.001

Fauth, B., Wagner, W., Bertram, C., Göllner, R., Roloff, J., Lüdtke, O., Polikoff, M. S., Klusmann, U., & Trautwein, U. (2020). Don't blame the teacher? The need to account for classroom characteristics in evaluations of teaching quality. *Journal of Educational Psychology*, *112*(6), 1284-1302. https://doi.org/10.1037/edu0000416

Fenn, J., Tan, C.-S., & George, S. (2020). Development, validation and translation of psychological tests. *BJPsych advances*, *26*(5), 306-315. https://doi.org/10.1192/bja.2020.33

Field, A. P. (2018). *Discovering statistics using IBM SPSS statistics* (5. ed.). Sage Publications.

Fishbein, B., Foy, P., & Yin, L. (2021). *TIMSS 2019 user guide for the international database (2nd ed.)* Retrieved from Boston College, TIMSS & PIRLS International Study Center website: https://timssandpirls.bc.edu/timss2019/international-database/

Fjellman, A.-M. (2019). *School choice, space and the geography of marketization : analyses of educational restructuring in upper secondary education in Sweden* Göteborgs universitet. Acta universitatis Gothoburgensis. https://gupea.ub.gu.se/handle/2077/58025

Foy, P., Fishbein, B., von Davier, M., & Yin, L. (2020). Implementing the TIMSS 2019 scaling methodology. In M. O. Martin, M. von Davier, & I. V. S. Mullis (Eds.), *Methods and Procedures: TIMSS 2019 Technical Report*. Boston College. https://timssandpirls.bc.edu/timss2019/methods/chapter-11.html

Frønes, T. S., Pettersen, A., Radišić, J., & Buchholtz, N. (Eds.). (2021). *Equity, equality and diversity in the Nordic Model of education*. Cham : Springer. https://doi.org/10.1007/978-3-030-61648-9.

Gaertner, H., & Brunner, M. (2018). Once good teaching, always good teaching? The differential stability of student perceptions of teaching quality. *Educational Assessment, Evaluation and Accountability*, *30*(2), 159-182. https://doi.org/10.1007/s11092-018-9277-5

Goe, L. (2007). *The link between teacher quality and student outcomes: A research synthesis*. https://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED521219&site=ehost-live

Gomes, A. M., Martins, M. C., Farinha, M., Silva, B., Ferreira, E., Caldas, A. C., & Brandão, T. (2020). Bullying's negative effect on academic achievement. *International Journal of Educational Psychology*, *9*(3), 243-268. https://doi.org/10.17583/ijep.2020.4812

Grek, S. (2009). Governing by numbers: the PISA 'effect' in Europe. *Journal of education policy*, *24*(1), 23-37. https://doi.org/10.1080/02680930802412669

Gren Landell, M. (2021). *School attendance problems: a research update and where to go* (Första upplagan ed.). Jerringfonden.

Guerriero, S. (2017). Teachers' pedagogical knowledge: What it is and how it functions. In S. Guerriero (Ed.), *Pedagogical Knowledge and the Changing Nature of the Teaching Profession* (pp. 99-118). OECD Publishing.

Gustafsson, J.-E. (2016). Lasting effects of quality of schooling: Evidence from PISA and PIAAC. *Intelligence*, *57*, 66-72. https://doi.org/10.1016/j.intell.2016.05.004

Gustafsson, J.-E. (2018). International large scale assessments: current status and ways forward. *Scandinavian Journal of Educational Research*, *62*(3), 328-332. https://doi.org/10.1080/00313831.2018.1443573

Gustafsson, J.-E., Sörlin, S., & Vlachos, J. (2016). *Policyidéer för svensk skola*. Studieförbundet Näringsliv och, samhälle: SNS.

Gustafsson, J.-E., Yang Hansen, K., & Rosén, M. (2013). Effects of home background on student achievement in reading, mathematics, and science at the fourth grade. . In M. O. Martin & I. V. S. Mullis (Eds.), *TIMSS and PIRLS 2011: Relationships among reading, mathematics, and science achievement at the fourth grade – implications for early learning.* (pp. 181-285). TIMSS & PIRLS International Study Center, Boston College.

Hafsteinsdóttir, E., Appelgren, A., Bergman, E., Bergman, M., Melin, C., Lilja, A., Samuelsson, M., & Skolforskningsinstitutet. (2021). *Främja studiero i klassrummet: lärares ledarskap*. Skolforskningsinstitutet. https://www.skolfi.se/wp-content/uploads/2021/09/Främja-studiero-i-klassrummet-—-lärares-ledarskap-fulltext.pdf

Hansson, A. (2012). The meaning of mathematics instruction in multilingual classrooms: analyzing the importance of responsibility for learning. *Educational Studies in Mathematics*, *81*(1), 103-125. https://doi.org/10.1007/s10649-012-9385-y

Hansson, Å., & Gustafsson, J.-E. (2016). Pedagogisk segregation: Lärarkompetens i den svenska grundskolan ur ett likvärdighetsperspektiv. *Pedagogisk forskning i Sverige 21*(1/2), 56-128. blob:https://artikelsok-ip-btj-se.ezproxy.ub.gu.se/eb365f01-5bea-44f8-925d-95004f511bde

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

Hattie, J., & Yates, G. (2013). *Visible learning and the science of how we learn*. Routledge.

He, J., Barrera-Pedemonte, F., & Buchholz, J. (2019). Cross-cultural comparability of noncognitive constructs in TIMSS and PISA. *Assessment in Education: Principles, Policy & Practice*, *26*(4), 369-385. https://doi.org/10.1080/0969594X.2018.1469467

He, J., Buchholz, J., & Fischer, J. (2022). Cross-cultural comparability of latent constructs in ILSAs. In T. Nilsen, A. Stancel-Piątak, & J.-E. Gustafsson (Eds.), *International Handbook of Comparative Large-Scale Studies in Education. Perspectives, Methods and Findings* (pp. 845-870). Springer International Handbooks of Education. https://doi.org/https://doi.org/10.1007/978-3-030-88178-8_58

Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning: A Project of the National Council of teachers of mathematics* (pp. 371–404). Information Age Publishing, Incorporated. https://ebookcentral.proquest.com/lib/gu/detail.action?docID=3315201

Holmlund, H., Häggblom, J., Lindahl, E., Martinson, S., Sjögren, A., Vikman, U., & Öckert, B. (2014). *Decentralisering, skolval och fristående skolor: resultat och likvärdighet i svensk skola* (Report 2014:25). https://www.ifau.se/Forskning/Publikationer/Rapporter/2014/Decentralisering-skolval-och-fristaende-skolor-resultat-och-likvardighet-i-svensk-skola/

Hooper, M. (2022). Dilemmas in developing context questionnaires for International Large-Scale Assessments. In T. Nilsen, A. Stancel-Piątak, & J.-E. Gustafsson (Eds.), *International handbook of comparative large-scale studies in education. Perspectives, methods and findings* (pp. 721-748). Springer International Handbooks of Education. https://doi.org/https://doi.org/10.1007/978-3-030-88178-8_29

Hooper, M., Arora, A., Martin, M. O., & Mullis, I. V. S. (2013). *Examining the behavior of "reverse directional" items in the TIMSS 2011 context questionnaire scales* the 5th IEA International Research Conference (IRC-2013), Singapore. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.639.1034&rep=rep1&type=pdf

Hooper, M., Mullis, I. V. S., Martin, M. O., & Fishbein, B. (2017). TIMSS 2019 context questionnaire framework. In I. V. S. Mullis & M. O. Martin (Eds.), *TIMSS 2019 assessment frameworks* (pp. 59-78). Boston College, TIMSS & PIRLS International Study Center. http://timssandpirls.bc.edu/timss2019/frameworks/

Hox, J. J. (2002). *Multilevel analysis: Techniques and applications*. Lawrence Erlbaum Associates. https://web-p-ebscohost-com.ezproxy.ub.gu.se/ehost/ebookviewer/ebook/bmxlYmtfXzY5MDYyX19BTg2?sid=6643c4cd-ae68-4976-9fc4-ed99618465f3@redis&vid=0&format=EB&rid=1

Hox, J. J., Roberts, J. K., Taylor, & Francis. (2011). *Handbook of advanced multilevel analysis*. New York: Routledge.

IEA. (nd). *IEA IDB Analyzer*. IEA. https://www.iea.nl/data-tools/tools

Ivanova, M., & Michaelides, M. P. (2022). Motivational components in TIMSS 2015 and their effects on engaging teaching practices and mathematics performance. *Studies in Educational Evaluation*, *74*, Article Article 101173. https://doi.org/10.1016/j.stueduc.2022.101173

Johansson, S. (2016). International large-scale assessments: What uses, what consequences? *Educational Research*, *58*(2), 139-148. https://doi.org/10.1080/00131881.2016.1165559

Kane, M. T. (2006). Validation. In R. L. Brennan (Ed.), *Educational Measurement* (4th ed., Vol. Forth Edition, pp. 17-64). American Council on Education/Praeger Publishers.

Kavli, A.-B. (2018). TIMSS and PISA in the Nordic countries In N. C. o. Ministers (Ed.), *Northern Lights on TIMSS and PISA 2018* (pp. 11-30). Nordic Council of Ministers. https://doi.org/10.6027/TN2018-524

Kearney, C. A., Gonzálvez, C., Graczyk, P. A., & Fornander, M. J. (2019). Reconciling contemporary approaches to school attendance and school absenteeism: Toward promotion and nimble response, global policy review and implementation, and future adaptability (Part 2). *Frontiers in Psychology*, *10*, 2605-2605. https://doi.org/10.3389/fpsyg.2019.02605

Kelly, D. L., Centurino, V. A. S., Martin, M. O., & Mullis, I. V. S. (2020). *TIMSS 2019 encyclopedia: Education policy and curriculum in mathematics and science.* Boston College, TIMSS & PIRLS International Study Center website. https://timssandpirls.bc.edu/timss2019/encyclopedia/

Khine, M. S., Ping, L. C., & Cunningham, D. (2013). *Application of structural equation modeling in educational research and practice*. Birkhäuser Boston. https://doi.org/10.1007/978-94-6209-332-4

Kim, M. H., & Sidney, P. G. (2024). Do teacher instructional practices shape children's academic self-concept and interest in mathematics and science? Evidence from TIMSS 2015. *Infant and Child Development*, *33*(2), Article e242. https://doi.org/10.1002/icd.2429

Klapp, T., Gustafsson, J.-E., & Johansson, S. (2025). Big-Fish-Little-Pond effects on ninth-grade students' mathematics and language self-concepts: The moderating role of cognitive ability. *Journal of Educational Psychology*, *117*(2), 257-272. https://doi.org/10.1037/edu0000919

Klieme, E., & Nilsen, T. (2022). *Teaching quality and student outcomes in TIMSS and PISA*. Springer International Publishing. https://doi.org/10.1007/978-3-030-88178-8_37

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (Fourth edition. ed.). Guilford Press. https://ebookcentral.proquest.com/lib/GU/detail.action?docID=4000663

Koivuhovi, S., Vainikainen, M.-P., & Kalalahti, M. (2022). The effect of studying in selective classes on the change in pupils' action-control beliefs during lower secondary school in Finland. *Scandinavian Journal of Educational Research*, *66*(1), 105-118. https://doi.org/10.1080/00313831.2020.1833246

Konstantinidou, E., & Kyriakides, L. (2022). Instructional engagement and student learning outcomes: Direct and indirect effects based on country-specific contingencies. *Studies in Educational Evaluation*, *73*, Article 101144. https://doi.org/10.1016/j.stueduc.2022.101144

Korpershoek, H., Harms, T., de Boer, H., van Kuijk, M., & Doolaard, S. (2016). A meta-analysis of the effects of classroom management strategies and classroom management programs on students' academic, behavioral, emotional, and motivational outcomes. *Review of Educational Research*, *86*(3), 643-680. https://doi.org/10.3102/0034654315626799

Kosunen, S., Bernelius, V., Seppänen, P., & Porkka, M. (2020). School choice to lower secondary schools and mechanisms of segregation in urban Finland. *Urban Education*, *55*(10), 1461-1488. https://doi.org/10.1177/0042085916666933

Kunter, M., & Baumert, J. (2006). Who is the expert?: Construct and criteria validity of student and teacher ratings of instruction. *Learning Environments Research*, *9*(3), 231-251. https://doi.org/10.1007/s10984-006-9015-7

Kunter, M., Baumert, J., & Köller, O. (2007). Effective classroom management and the development of subject-related interest. *Learning and Instruction*, *17*(5), 494-509. https://doi.org/10.1016/j.learninstruc.2007.09.002

Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T., & Hachfeld, A. (2013). Professional competence of reachers: Effects on instructional quality and student development. *Journal of Educational Psychology*, *105*(3), 805-820. https://doi.org/10.1037/a0032583

Kyriakides, L., Charalambous, C., Charalambous, E., & Blömeke, S. (2022). Using ILSAs to promote quality and equity in education: The contribution of the dynamic model of educational effectiveness. In T. Nilsen, A. Stancel-Piątak, & J.-E. Gustafsson (Eds.), *International handbook of comparative large-scale studies in education: Perspectives, methods and findings* (pp. 253-276). Cham: Springer International Publishing. https://doi.org/https://doi.org/10.1007/978-3-030-88178-8_13

Kyriakides, L., Christoforou, C., & Charalambous, C. Y. (2013). What matters for student learning outcomes: A meta-analysis of studies exploring factors of effective teaching. *Teaching and Teacher Education*, *36*, 143-152. https://doi.org/10.1016/j.tate.2013.07.010

Kyriakides, L., Creemers, B., & Panayiotou, A. (2020). Developing and testing theories of educational effectiveness addressing the dynamic nature of education. In (pp. 33-69). Springer International Publishing. https://doi.org/10.1007/978-3-030-44810-3_3

Kyriakides, L., & Creemers, B. P. M. (2008). Using a multidimensional approach to measure the impact of classroom-level factors upon student achievement: a study testing the validity of the dynamic model. *School Effectiveness and School Improvement*, *19*(2), 183-205. https://doi.org/10.1080/09243450802047873

König, J., & Blömeke, S. (2012). Future teachers' general pedagogical knowledge from a comparative perspective: does school experience matter? *ZDM*, *44*(3), 341-354. https://doi.org/10.1007/s11858-012-0394-1

Lazarides, R., & Buchholz, J. (2019). Student-perceived teaching quality: How is it related to different achievement emotions in mathematics classrooms? *Learning & Instruction*, *61*, 45-59. https://doi.org/10.1016/j.learninstruc.2019.01.001

Lazarides, R., & Ittel, A. (2012). Instructional quality and attitudes toward mathematics: Do self-concept and interest differ across students' patterns of perceived instructional quality in mathematics classrooms? *Child Development Research*, 1-11. https://doi.org/10.1155/2012/813920

Lazarides, R., & Raufelder, D. (2021). Control-value theory in the context of teaching: does teaching quality moderate relations between academic self-concept and achievement emotions? *British Journal of Educational Psychology*, *91*(1), 127-147. https://doi.org/10.1111/bjep.12352

Lee, J., & Chen, M. (2019). Cross-country predictive validities of non-cognitive variables for mathematics achievement : Evidence based on TIMSS 2015. *Eurasia Journal of Mathematics, Science and Technology Education*, *15*(8), Article em1725. https://doi.org/10.29333/ejmste/106230

Lee, J., & Stankov, L. (2018). Non-cognitive predictors of academic achievement: Evidence from TIMSS and PISA. *Learning and Individual Differences*, *65*, 50-64. https://doi.org/10.1016/j.lindif.2018.05.009

Legault, L., Green-Demers, I., & Pelletier, L. (2006). Why do high school students lack motivation in the classroom? Toward an understanding of academic amotivation and the role of social support. *Journal of Educational Psychology*, *98*(3), 567-582. https://doi.org/10.1037/0022-0663.98.3.567

Leijen, Ä., Malva, L., Pedaste, M., & Mikser, R. (2022). What constitutes teachers' general pedagogical knowledge and how it can be assessed: A literature review. *Teachers and Teaching, Theory and Practice*, *28*(2), 206-225. https://doi.org/10.1080/13540602.2022.2062710

Lester, F. K. (2007). *Second handbook of research on mathematics teaching and learning a project of the national council of teachers of mathematics*. Information Age Publishing, Inc.

Lindström, M., Johansson, S., & Borger, L. (2024). Does formal teacher competence matter for students' mathematics achievement? Results from Swedish TIMSS 2019. *Educational Research and Evaluation*, 1-30. https://doi.org/10.1080/13803611.2024.2367486

Linn, R. (2010). Educational measurement: Overview. *International Encyclopedia of Education*, *4*, 45–49. https://doi.org/10.1016/B978-0-08-044894-7.00243-8

Liu, S., & Meng, L. (2010). Re-examining factor structure of the attitudinal items from TIMSS 2003 in cross-cultural study of mathematics self-concept. *Educational Psychology*, *30*(6), 699-712. https://doi.org/10.1080/01443410.2010.501102

Liu, X., Hansen, K. Y., Valcke, M., & De Neve, J. (2024). A decade of PISA: student perceived instructional quality and mathematics achievement across European countries. *ZDM*, *56*(5), 859-891. https://doi.org/10.1007/s11858-024-01630-7

Lundahl, L. (2016). Equality, inclusion and marketization of Nordic education: Introductory notes. *Research in Comparative and International Education*, *11*(1), 3-12. https://doi.org/10.1177/1745499916631059

Lüdtke, O., Marsh, H. W., Robitzsch, A., Trautwein, U., Asparouhov, T., & Muthén, B. (2008). The multilevel latent covariate model: A new, more reliable approach to group-level effects in contextual studies. *Psychological Methods*, *13*(3), 203-229. https://doi.org/10.1037/a0012869

Lüdtke, O., Robitzsch, A., Trautwein, U., & Kunter, M. (2009). Assessing the impact of learning environments: How to use student ratings of classroom or school characteristics in multilevel modelling. *Contemporary Educational Psychology*, *34*(2), 120-131. https://doi.org/10.1016/j.cedpsych.2008.12.001

MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effect. *Prevention Science*, *1*(4), 173-181. https://doi.org/10.1023/A:1026595011371

MacKinnon, D. P., Lockwood, C. M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research*, *39*(1), 99-128. https://doi.org/10.1207/s15327906mbr3901_4

Marsh, H. W. (1994). Using the national longitudinal dtudy of 1988 to evaluate theoretical models of self-concept: The self-description questionnaire. *Journal of Educational Psychology*, *86*(3), 439-456. https://doi.org/10.1037/0022-0663.86.3.439

Marsh, H. W., Abduljabbar, A. S., Abu-Hilal, M. M., Morin, A. J. S., Abdelfattah, F., Leung, K. C., Xu, M. K., Nagengast, B., & Parker, P. (2013). Factorial, convergent, and discriminant validity of TIMSS math and science motivation measures: A comparison of Arab and Anglo-Saxon countries. *Journal of Educational Psychology*, *105*(1), 108-128. https://doi.org/10.1037/a0029907

Marsh, H. W., & Craven, R. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective: Beyond seductive pleasure and unidimensional perspectives. *Perspectives on Psychological Science*, *1*(2), 133-163. https://doi.org/10.1111/j.1745-6916.2006.00010.x

Marsh, H. W., Lüdtke, O., Nagengast, B., Trautwein, U., Morin, A. J. S., Abduljabbar, A. S., & Köller, O. (2012). Classroom climate and contextual effects: Conceptual and methodological issues in the evaluation of group-level effects. *Educational Psychologist*, *47*(2), 106-124. https://doi.org/10.1080/00461520.2012.670488

Marsh, H. W., & Martin, A. J. (2011). Academic self-concept and academic achievement: Relations and causal ordering. *British Journal of Educational Psychology*, *81*(1), 59-77. https://doi.org/10.1348/000709910X503501

Marsh, H. W., Parker, P. D., & Pekrun, R. (2019). Three paradoxical effects on academic self-concept across countries, schools, and students: Frame-of-reference as a unifying theoretical explanation. *European Psychologist*, *24*(3), 231-242. https://doi.org/10.1027/1016-9040/a000332

Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades and standardized test scores: Reciprocal effects models of causal ordering. *Child Development*, *76*(2), 397-416. https://doi.org/10.1111/j.1467-8624.2005.00853.x

Martin, M. O., Mullis, I. V., & Foy, P. (2017). TIMSS 2019 assessment design. In I. V. S. Mullis & M. O. Martin (Eds.), *IMSS 2019 assessment frameworks*. Boston College. http://timssandpirls.bc.edu/timss2019/frameworks/

Martin, M. O., von Davier, M., & Mullis, I. V. S. (2020). *Methods and procedures: TIMSS 2019 technical report*. Boston College and International Association for the Evaluation of Educational Achievement (IEA). https://timssandpirls.bc.edu/timss2019/methods

Marzano, R. J., Marzano, J. S., & Pickering, D. (2003). Classroom management that works: Research-based strategies for every teacher. In *Educational Leadership* (Vol. 61, pp. 6-13). Association for Supervision and Curriculum Development. https://ascd.org/el/articles/the-key-to-classroom-management

Meinck, S., Gonzalez, E., & Wagemaker, H. (2020). Consequential validity: Data access, data use, analytical support, and training. In H. Wagemaker (Ed.), R*eliability and validity of international large-scale assessment: Understanding IEA's comparative studies of student achievement* (pp. 231-244). Springer Cham. https://doi.org/https://doi.org/10.1007/978-3-030-53081-5

Merritt, S. M. (2012). The two-factor solution to Allen and Meyer's (1990) affective commitment scale: Effects of negatively worded items. *Journal of Business and Psychology*, *27*(4), 421-436. https://doi.org/10.1007/s10869-011-9252-3

Messick, S. (1984). The psychology of educational measurement. *Journal of Educational Measurement*, *21*(3), 215-237. https://doi.org/10.1111/j.1745-3984.1984.tb01030.x

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational Measurement* (pp. 13-104). Macmillan Publishing Company.

Michaelides, M. P. (2019). Negative keying effects in the factor structure of TIMSS 2011 motivation scales and associations with reading achievement. *Applied Measurement in Education*, *32*(4), 365-378. https://doi.org/10.1080/08957347.2019.1660349

Min, I., Cortina, K. S., & Miller, K. F. (2016). Modesty bias and the attitude-achievement paradox across nations: A reanalysis of TIMSS. *Learning and Individual Differences*, *51*, 359-366. https://doi.org/10.1016/j.lindif.2016.09.008

Mittal, O., Nilsen, T., & Björnsson, J. K. (2021). Measuring equity across the Nordic education systems-conceptual and methodological choices as implications for educational policies. In T. S. Frønes, A. Pettersen, J. Radišic, & N. Buchholtz (Eds.), *Equity, Equality and Diversity in the Nordic Model of Education* (pp. 43-71). https://doi.org/https://doi.org/10.1007/978-3-030-61648-9_3

Muijs, D., Kyriakides, L., van der Werf, G., Creemers, B., Timperley, H., & Earl, L. (2014). State of the art - teacher effectiveness and professional learning. *School Effectiveness and School Improvement*, *25*(2), 231-256. https://doi.org/10.1080/09243453.2014.885451

Mullis, I. V. S., Martin, M., Foy, P., Kelly, D., & Fischbein, B. (2020). *TIMSS 2019 international results in mathematics and science*. TIMSS & PIRLS International Study Center, IEA, Boston College. https://ur.bc.edu/islandora/timss-2019-international-results-mathematics-and-science

Mullis, I. V. S., & Martin, M. O. (2017). TIMSS 2019 assessment frameworks. http://timssandpirls.bc.edu/timss2019/frameworks/

Muthén, L. K., & Muthén, B. O. (2017). *Mplus version 8 user's guide*. Muthén & Muthén.

Nations, U. (2021). *Our common agenda: Report of the Secretary-General*. https://digitallibrary.un.org/record/3896455?v=pdf#files

Newton, P. E., & Shaw, S. D. (2014). *Validity in educational and psychological assessment*. SAGE. https://doi.org/https://doi.org/10.4135/9781446288856

Nilsen, & Gustafsson. (2016). *Teacher quality, instructional quality and student outcomes: Relationships across countries, cohorts and time* (Vol. 2). Springer Open. https://doi.org/10.1007/978-3-319-41252-8

Nilsen, T., & Bergem, O. K. (2020). Teacher competence and equity in the Nordic countries. Mediation and moderation of the relation between SES and achievement. *Acta Didactica Norden*, *14*(1), 1-26. https://doi.org/10.5617/adno.7946

Nilsen, T., Gustafsson, J.-E., & Blömeke, S. (2016). Conceptual framework and methodology of this report. In T. Nilsen & J.-E. Gustafsson (Eds.), *Teacher quality, instructional quality and* *student outcomes: Relationships across countries, cohorts and time* (pp. 1 - 20). Springer Open. https://doi.org/10.1007/978-3-319-41252-8

Nilsen, T., Scherer, R., & Blömeke, S. (2018). The relation of science teachers' quality and instruction to student motivation and achievement in the 4th and 8th grade: A Nordic perspective. In *Northern Lights on TIMSS and PISA 2018* (Vol. Nordic Council of Ministers, pp. pp. 61-94). https://doi.org/https://doi.org/10.6027/TN2018-524

Nilsen, T., Scherer, R., Gustafsson, J. E., Teig, N., & Kaarstein, H. (2021). Teachers' role in enhancing equity-A multilevel structural equation modelling with mediated moderation. In T. S. Frønes, A. Pettersen, J. Radišic, & N. Buchholtz (Eds.), *Equity, equality and diversity in the Nordic Model of education* (pp. 173-196). https://doi.org/10.1007/978-3-030-61648-9_7

Nordic Co-operation. (nd). Retrieved from https://www.norden.org/en/information/education-research-and-digitalisation-nordic-region

OECD. (2015). *Universal basic skills* (E. A. Hanushek & L. Woessmann, Eds.). https://doi.org/doi:https://doi.org/10.1787/9789264234833-en

OECD. (2017). *Pedagogical knowledge and the changing nature of the teaching profession*. OECD Publishing. https://doi.org/doi:https://doi.org/10.1787/9789264270695-en

OECD. (2023). *PISA 2022 Results (Volume I).* https://doi.org/doi:https://doi.org/10.1787/53f23881-en

Orth, U., Robins, R. W., & Widaman, K. F. (2012). Life-span development of self-esteem and its effects on important life outcomes. *Journal of Personality and Social Psychology*, *102*(6), 1271-1288. https://doi.org/10.1037/a0025558

Osborn, M. (2004). New methodologies for comparative research? Establishing 'constants' and 'contexts' in educational experience. *Oxford Review of Education*, *30*(2), 265-285. https://doi.org/10.1080/0305498042000215566

Panayiotou, A., Herbert, B., Sammons, P., & Kyriakides, L. (2021). Conceptualizing and exploring the quality of teaching using generic frameworks: A way forward. *Studies in Educational Evaluation*, *70*, Article 101028. https://doi.org/10.1016/j.stueduc.2021.101028

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher*, *38*(2), 109-119. https://doi.org/10.3102/0013189x09332374

Praetorius, A.-K., Klieme, E., Herbert, B., & Pinger, P. (2018). Generic dimensions of teaching quality: the German framework of Three Basic Dimensions. *ZDM*, *50*(3), 407-426. https://doi.org/10.1007/s11858-018-0918-4

Praetorius, A.-K., Klieme, E., Kleickmann, T., Brunner, E., Lindmeier, A., Taut, S., & Charalambous, C. Y. (2020). Towards developing a theory of generic teaching quality. Origin, current status, and necessary next steps regarding the Three Basic Dimensions Model. *Zeitschrift für Pädagogik*, *Beiheft, 66*, 15–36. https://doi.org/10.25656/01:25861

Ramazan, O., Danielson, R. W., Rougee, A., Ardasheva, Y., & Austin, B. W. (2023). Effects of classroom and school climate on language minority students' PISA mathematics self-concept and achievement scores. *Large-Scale Assessments in Education*, *11*(1), 11-30. https://doi.org/10.1186/s40536-023-00156-w

Reimer, D., Sortkear, B., Oskarsson, M., & Nilsen, T. (2018). *Northern Lights on TIMSS and PISA 2018*. Nordic Council of Ministers, Nordic Council of Ministers Secretariat. https://doi.org/10.6027/TN2018-524

Reynolds, D., Sammons, P., De Fraine, B., Van Damme, J., Townsend, T., Teddlie, C., & Stringfield, S. (2014). Educational effectiveness research (EER): a state-of-the-art review. *School Effectiveness and School Improvement*, *25*(2), 197-230. https://doi.org/10.1080/09243453.2014.885450

Reynolds, K., Khorramdel, L., & von Davier, M. (2022). Can students' attitudes towards mathematics and science be compared across countries? Evidence from measurement invariance modeling in TIMSS 2019. *Studies in Educational Evaluation*, *74*, Article 101169. https://doi.org/10.1016/j.stueduc.2022.101169

Rolfe, V. (2021). Tailoring a measurement model of socioeconomic status: Applying the alignment optimization method to 15 years of PISA. *International Journal of Educational Research*, *106*, Article 101723. https://doi.org/10.1016/j.ijer.2020.101723

Rosén, M., & Nilsen, T. (2024). Are changes in content coverage related to changes in achievement over Time? In N. Teig, T. Nilsen, & K. Yang Hansen (Eds.), *Effective and equitable teacher practice in mathematics and science education,* (Vol. IEA Research for Education, pp. 187-205). Springer, Cham. https://doi.org/https://doi.org/10.1007/978-3-031-49580-9_6

Rosenberg, M., Schooler, C., Schoenbach, C., & Rosenberg, F. (1995). Global self-esteem and specific self-esteem: Different concepts, different outcomes. *American Sociological Review*, *60*(1), 141-156. https://doi.org/10.2307/2096350

Roszkowski, M. J., & Soven, M. (2010). Shifting gears: consequences of including two negatively worded items in the middle of a positively worded questionnaire. *Assessment and Evaluation in Higher Education*, *35*(1), 113-130. https://doi.org/10.1080/02602930802618344

Rutkowski, D., & Rutkowski, L. (2013). Measuring socioeconomic background in PISA: One size might not fit all. *Research in Comparative and International Education*, *8*(3), 259-278. https://doi.org/10.2304/rcie.2013.8.3.259

Rutkowski, L., Gonzalez, E., Joncas, M., & von Davier, M. (2010). International large-scale assessment data: Issues in secondary analysis and reporting. *Educational Researcher*, *39*(2), 142-151. https://doi.org/10.3102/0013189X10363170

Rutkowski, L., & Rutkowski, D. (2016). The relation between students' perceptions of instructional quality and bullying victimization. In T. Nilsen & J.-E. Gustafsson (Eds.), *Teacher quality, instructional quality and student outcomes relationships across countries, cohorts and time* (Vol. 2, pp. 115-134). Springer. https://link.springer.com/book/10.1007/978-3-319-41252-8

Rutkowski, L., & Svetina, D. (2014). Assessing the hypothesis of measurement invariance in the context of large-scale international surveys. *Educational and Psychological Measurement*, *74*(1), 31-57. https://doi.org/10.1177/0013164413498257

Rutkowski, L., von Davier, M., & Rutkowski, D. (2013). *Handbook of international large-scale assessment: Background, technical issues, and methods of data analysis* (1 ed., Vol. 13). Chapman and Hall/CRC. https://doi.org/https://doi.org/10.1201/b16061

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *The American Psychologist*, *55*(1), 68-78. https://doi.org/10.1037/0003-066X.55.1.68

Scheerens, J. (2015). Theories on educational effectiveness and ineffectiveness. *School Effectiveness and School Improvement*, *26*(1), 10-31. https://doi.org/10.1080/09243453.2013.858754

Scheerens, J., & Blömeke, S. (2016). Integrating teacher education effectiveness research into educational effectiveness models. *Educational Research Review*, *18*, 70-87. https://doi.org/10.1016/j.edurev.2016.03.002

Scherer, R., & Gustafsson, J.-E. (2015). Student assessment of teaching as a source of information about aspects of teaching quality in multiple subject domains: An application of multilevel bifactor structural equation modeling. *Frontiers in Psychology*, *6*(6), Article1550. https://doi.org/10.3389/fpsyg.2015.01550

Scherer, R., & Nilsen, T. (Eds.). (2016). *The relations among school climate, instructional quality, and achievement motivation in mathematics* (Vol. 2). Springer. https://doi.org/10.1007/978-3-319-41252-8_3.

Scherer, R., Nilsen, T., & Jansen, M. (2016). Evaluating individual students' perceptions of instructional quality. An investigation of their factor structure, measurement invariance, and relations to educational outcomes. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00110

Senden, B., Nilsen, T., & Blömeke, S. (2022). Instructional quality: A review of conceptualizations, measurement approaches, and research findings. In M. Blikstad-Balas, K. Klette, & M. Tengberg (Eds.), *Ways of analyzing teaching quality* (pp. 140-172). Scandinavian University Press. https://doi.org/10.18261/9788215045054-2021-05

Skollag (Education Act). https://rkrattsbaser.gov.se/sfst?bet=2010:800

Shadish, W. R., Cook, T. D., & Campbell, D. T. C. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin. https://doi.org/https://doi.org/10.4135/9781446288856

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research*, *46*(3), 407-441. https://doi.org/10.3102/00346543046003407

Shen, C., & Pedulla, J. J. (2000). The relationship between students' achievement and their self-perception of competence and rigour of mathematics and science: a cross-national analysis. *Assessment in Education: Principles, Policy & Practice*, *7*(2), 237-253. https://doi.org/10.1080/713613335

Shen, C., & Tam, H. P. (2008). The paradoxical relationship between student achievement and self-perception: a cross-national analysis based on three waves of TIMSS data. *Educational Research and Evaluation*, *14*(1), 87-100. https://doi.org/10.1080/13803610801896653

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Research*, *57*, 1-22. https://doi.org/https://doi.org/10.17763/haer.57.1.j463w79r56455411

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, *75*(3), 417-453. https://doi.org/10.3102/00346543075003417

Skolinspektionen. (2016). *Omfattande frånvaro. En granskning av skolors arbete med omfattande frånvaro* (Diarienummer 2015:2855). https://www.skolinspektionen.se/beslut-rapporter/publikationer/kvalitetsgranskning/2016/omfattande-franvaro/

Skolverket. (2014). *Privata aktörer inom förskola och skola. En nationell kartläggning av enskilda huvudmän och ägare* (Rapport 410). https://www.skolverket.se/getFile?file=3307

Skolverket. (2018). *Analyser av familjebakgrundens betydelse för skolresultaten och skillnader mellan skolor. En kvantitativ studie av utvecklingen över tid i slutet av grundskolan* (Rapport 467). https://www.skolverket.se/download/18.6bfaca41169863e6a65d200/1553967875648/pdf3927.pdf

Skolverket. (2019). *Lärarprognos 2019. Redovisning av uppdrag att ta fram återkommande prognoser öer behovet av förskollärare och olika lärarkategorier* (Diarienummer 5.1.3-2018:1500). Retrieved from https://www.skolverket.se/getFile?file=5394

Skolverket. (2020). *TIMSS 2019 Svenska grundskoleelevers kunskaper i matematik och naturvetenskap i ett internationellt perspektiv* (2020:8). https://www.skolverket.se/publikationsserier/rapporter/2020/timss-2019.-svenska-grundskoleelevers-kunskaper-i-matematik-och-naturvetenskap-i-ett-internationellt-perspektiv.

Skolverket. (2021). *Nationell kartläggning av elevfrånvaro. De obligatoriska skolformerna samt gymnasie- och gymnasiesärskolan*. (5.1.3-2019:1103). Retrieved from https://www.skolverket.se/publikationsserier/rapporter/2021/nationell-kartlaggning-av-elevfranvaro

Skolverket. (2022a). *Elever och skolenheter i grundskolan. Läsåret 2021/22* (Diarienummer: 2022:344). https://www.skolverket.se/download/18.68c99c081804c5929ea58cf/1655292646380/pdf9573.pdf

Skolverket. (2022b). *Pedagogisk personal i skola och vuxenutbildning. Läsåret 2021/22* (Diarienummen 2022:276). https://www.skolverket.se/getFile?file=9547

Skolverket. (2023a). *Förskole- och skolenheter/fritidshem. Barn/elever läsåret 2022/2023.* https://www.skolverket.se/skolutveckling/statistik/sok-statistik-om-forskola-skola-och-vuxenutbildning?sok=SokC&omrade=Skolor%20och%20elever&lasar=2023/24&run=1

Skolverket. (2023b). *Skolverkets bedömning av läget i skolväsendet 2023.* (Rapport 2023:9). https://www.skolverket.se/getFile?file=11810

Skolverket. (2024a). *Slutbetyg i grundskolan. Våren 2024* (Diarienummer 2024:2199). Skolverket. https://www.skolverket.se/getFile?file=13149

Skolverket. (2024b). *TIMSS 2023. Svenska grundskoleelevers kunskaper i matematik och naturvetenskap i ett internationellt perspektiv* (2024:13). (Internationella studier). https://www.skolverket.se/publikationsserier/rapporter/2024/timss-2023

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

SOU 2016:94. *Saknad! Uppmärksamma elevers frånvaro och agera: Betänkande av Att vända frånvaro till närvaro – en utredning om problematisk elevfrånvaro.* https://www.regeringen.se/rattsliga-dokument/statens-offentliga-utredningar/2017/01/sou-201694/

SOU 2017:35. *Samling för skolan. Nationell strategi för kunskap och likvärdighet.* https://www.regeringen.se/contentassets/e94a1c61289142bfbcfdf54a44377507/samling-for-skolan---nationell-strategi-for-kunskap-och-likvardighet-sou-201735.pdf

SOU 2019:40. *Jämlikhet i möjligheter och utfall i den svenska skolan*. Stockholm. https://www.regeringen.se/contentassets/23c13d7ae0ef48e4bed43b68917573d3/jamlikhet-i-mojligheter-och-utfall-i-den-svenska-skolan-sou-201940.pdf

SOU 2024:81. *Ämneskunskaper och lärarskicklighet - en reformerad lärarutbildning*. Stockholm: Statens Offentliga Utredningar. https://data.riksdagen.se/fil/84025A34-48E2-41A4-969D-0B54147C36D4

Spector, P. E. (1992). *Summated rating scale construction: an introduction* (Vol. 82). Sage Publications. https://doi.org/https://doi.org/10.4135/9781412986038

Stankov, L., & Lee, J. (2017). Self-beliefs: Strong correlates of mathematics achievement and intelligence. *Intelligence 61*, 11-16. https://doi.org/10.1016/j.intell.2016.12.001

Stankov, L., Morony, S., & Lee, Y. P. (2014). Confidence: the best non-cognitive predictor of academic achievement? *Educational Psychology 34*(1), 9-28. https://doi.org/10.1080/01443410.2013.814194

Steinmann, I., Sánchez, D., van Laar, S., & Braeken, J. (2022). The impact of inconsistent responders to mixed-worded scales on inferences in international large-scale assessments. *Assessment in Education: Principles, Policy and Practice*, *29*(1), 5-26. https://doi.org/10.1080/0969594X.2021.2005302

Steinmann, I., Strietholt, R., & Braeken, J. (2022). A constrained factor mixture analysis model for consistent and Inconsistent respondents to mixed-worded scales. *Psychological Methods*, *27*(4), 667-702. https://doi.org/10.1037/met0000392

Sundberg, D. (2005). *Skolreformernas dilemman: En läroplansteoretisk studie av kampen om tid i den svenska obligatoriska skolan* (Vol. 61). University Press. https://www.diva-portal.org/smash/get/diva2:206699/FULLTEXT01.pdf

Sundberg, D. (2021). *Svenska läroplaner : läroplansteori för de pedagogiska professionerna* (Upplaga 1 ed.). Studentlitteratur AB.

Sundqvist, P., & Sylvén, L. K. (2016). *Extramural English in teaching and learning. From theory and research to practice*. Palgrave Macmillan. https://doi.org/https://doi.org/10.1057/978-1-137-46048-6

Suri, H. (2020). Ethical considerations of conducting systematic reviews in educational research. In O. Zawacki-Richter, M. Kerres, S. Bedenlier, M. Bond, & K. Buntins (Eds.), *Systematic reviews in educational research methodology, perspectives and application* (pp. 41-54). Springer. https://doi.org/https://doi.org/10.1007/978-3-658-27602-7_3

Teig, N., Nilsen, T., & Yang Hansen, K. (Eds.). (2024). *Effective and equitable teacher practice in mathematics and science education: A Nordic perspective across time and groups of students* (Vol. 14). Springer Cham. https://doi.org/https://doi.org/10.1007/978-3-031-49580-9.

Titsworth, S., Mazer, J. P., Goodboy, A. K., Bolkan, S., & Myers, S. A. (2015). Two meta-analyses exploring the relationship between teacher clarity and student learning. *Communication Education*, *64*(4), 385-418. https://doi.org/10.1080/03634523.2015.1041998

Toropova, A., Johansson, S., & Myrberg, E. (2019). The role of teacher characteristics for student achievement in mathematics and student perceptions of instructional quality. *Education Inquiry*, *10*(4), 275-299. https://doi.org/http://dx.doi.org/10.1080/20004508.2019.1591844

UNESCO. (2019). *Right to education handbook*. Right to Education Initiative. https://unesdoc.unesco.org/ark:/48223/pf0000366556.locale=en

Upsing, B., & Rittberger, M. (2018). The translator's perspective on translation quality control processes for international large-scale assessment studies. *The International Journal of Translation and Interpreting Research*, *10*(2), 55-72. https://doi.org/10.12807/ti.110202.2018.a05

Wagemaker, H. (2020). *Reliability and validity of international large-scale assessment: Understanding IEA's comparative studies of student achievement*. Springer International Publishing AG. https://doi.org/10.1007/978-3-030-53081-5

Wagner, W., Göllner, R., Helmke, A., Trautwein, U., & Lüdtke, O. (2013). Construct validity of student perceptions of instructional quality is high, but not perfect: Dimensionality and generalizability of domain-independent assessments. *Learning and Instruction*, *28*, 1-11. https://doi.org/10.1016/j.learninstruc.2013.03.003

Wagner, W., Göllner, R., Werth, S., Voss, T., Schmitz, B., & Trautwein, U. (2016). Student and teacher ratings of instructional quality: Consistency of ratings over time, agreement, and predictive power. *Journal of Educational Psychology*, *108*(5), 705-721. https://doi.org/http://dx.doi.org/10.1037/edu0000075

Walberg, H. J., & Haertel, G. D. (1980). Evaluation reflections: Validity and use of educational environment assessments. *Studies in Educational Evaluation*, *6*(3), 225-238. https://doi.org/10.1016/0191-491X(80)90026-7

Van Dam, N. T., Hobkirk, A. L., Danoff-Burg, S., & Earleywine, M. (2012). Mind your words: Positive and negative items create method effects on the five facet mindfulness questionnaire. *Assessment*, *19*(2), 198-204. https://doi.org/10.1177/1073191112438743

van Sonderen, E., Sanderman, R., & Coyne, J. C. (2013). Ineffectiveness of reverse wording of questionnaire items: Let's learn from cows in the rain. *PloS one*, *8*(7), e68967-e68967. https://doi.org/10.1371/journal.pone.0068967

van Zyl, L. E., & Klooster, P. M. (2022). Exploratory structural equation modeling: Practical guidelines and tutorial with a convenient online tool for Mplus. *Frontiers in Psychiatry*, *12*, Article 795672. https://doi.org/10.3389/fpsyt.2021.795672

Wang, M. C., Haertel, G. D., & Walberg, H. J. (1993). Toward a knowledge base for school learning. *Review of Educational Research*, *63*(3), 249-294. https://doi.org/10.2307/1170546

Wang, Y., Kim, E. S., Dedrick, R. F., Ferron, J. M., & Tan, T. (2018). A multilevel bifactor approach to construct validation of mixed-format scales. *Educational and Psychological Measurement*, *78*(2), 253-271. https://doi.org/10.1177/0013164417690858

Vansteenkiste, M., & Ryan, R. M. (2013). On psychological growth and vulnerability: Basic psychological need satisfaction and need frustration as a unifying principle. *Journal of Psychotherapy Integration*, *23*(3), 263-280. https://doi.org/10.1037/a0032359

Watson, D., Clark, L. A., Chmielewski, M., Kotov, R., & Reynolds, C. R. (2013). The value of suppressor effects in explicating the construct validity of symptom measures. *Psychological Assessment*, *25*(3), 929-941. https://doi.org/10.1037/a0032781

Vetenskapsrådet. (2024). *God forskningssed 2024*. https://www.vr.se/analys/rapporter/vara-rapporter/2024-10-02-god-forskningssed-2024.html

Volmari, K. (2019). *Basic education in the Nordic region. Similar values, different policies* (Reports and surveys 2019:4). F. N. A. f. Education. https://www.oph.fi/sites/default/files/documents/30955454_basic_education_in_the_nordic_region_netti_19_04_08_0.pdf

von Davier, M. (2020). TIMSS 2019 scaling methodology: Item response theory, population models, and linking across modes. In M. von Davier, M. O. Martin, & I. V. S. Mullis (Eds.), *Methods and procedures: TIMSS 2019 technical report*. Boston College. https://timssandpirls.bc.edu/timss2019/methods/index.html

von Davier, M., Gonzalez, E., & Mislevy, R. J. (2009). What are plausible values and why are they useful? In M. von Davier & D. Hastedt (Eds.), *IERI monograph series issues and methodologies in large-scale assessments* (Vol. 2). IEA-ETS Research Institute.

https://www.ierinstitute.org/fileadmin/Documents/IERI_Monograph/IERI_Monograph_Volume_02.pdf

Wubbels, T. (2011). An international perspective on classroom management: What should prospective teachers learn? *Teaching Education*, *22*(2), 113-131. https://doi.org/10.1080/10476210.2011.567838

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (1st ed.). Harvard University Press. https://doi.org/https://doi-org.ezproxy.ub.gu.se/10.2307/j.ctvjf9vz4

Yang Hansen, K., & Gustafsson, J.-E. (2019). Identifying the key source of deteriorating educational equity in Sweden between 1998 and 2014. *International Journal of Educational Research*, *93*, 79-90. https://doi.org/10.1016/j.ijer.2018.09.012

Yang Hansen, K., Gustafsson, J.-E., Rosén, M., Sulkunen, S., Nissinen, K., Kupari, P., Ólafsson, R. F., Björnsson, J. K., Grønmo, L. S., Rønberg, L., Mejding, J., Borge, I. C., & Hole, A. (Eds.). (2014). *Northern lights on TIMSS and PIRLS 2011: Differences and similarities in the Nordic countries*. Nordic Council of Ministers. https://doi.org/http://dx.doi.org/10.6027/TN2014-528.

Ye, F., & Wallace, T. L. (2014). Psychological sense of school membership scale: Method effects associated with negatively worded items. *Journal of Psychoeducational Assessment*, *32*(3), 202-215. https://doi.org/10.1177/0734282913504816

Yin, & Fishbein. (2020). Creating and interpreting the TIMSS 2019 context questionnaire scales. In M. O. Martin, M. v. Davier, & I. V. S. Mullis (Eds.), *Methods and procedures: TIMSS 2019 technical report* (pp. 16.11-16.331). https://timssandpirls.bc.edu/timss2019/methods/chapter-16.html

# APPENDIX A

Table A1 Confirmatory factor analysis with standardised factor loadings for the Swedish sample

| | | |
|---|---|---|
| **Classroom management** | | |
| Model fit: $X^2$ = 139.827(9); RMSEA = .061; CFI = .983; TLI = .972; SRMR = .025 | | |
| Item | Indicator | Factor loading |
| 1 | Students don't listen to what the teacher says | .644 |
| 2 | There is disruptive noise | .750 |
| 3 | It is too disorderly for students to work well | .761 |
| 4 | My teacher has to wait a long time for students to quiet down | .824 |
| 5 | Students interrupt the teacher | .765 |
| 6 | My teacher has to keep telling us to follow the classroom rules | .629 |
| **Instructional clarity** | | |
| Model fit: $X^2$ = 127.729(9); RMSEA = .058; CFI = .970; TLI = .950; SRMR = .030 | | |
| 1 | I know what my teacher expects me to do | .378 |
| 2 | My teacher is easy to understand | .717 |
| 3 | My teacher has clear answers to my questions | .778 |
| 4 | My teacher is good at explaining mathematics | .790 |
| 5 | My teacher does a variety of things to help us learn | .732 |
| 6 | My teacher explains a topic again when we don't understand | .593 |
| **Mathematics confidence** | | |
| Model fit: $X^2$ = 922.839(27); RMSEA = .092; CFI = .887; TLI = .850; SRMR = .053 | | |
| 1 | **I** usually do well in mathematics | .743 |
| 2 | Mathematics is harder for me than for many of my classmates | -.626 |
| 3 | I am just not good at mathematics | -.717 |
| 4 | I learn things quickly in mathematics | .742 |
| 5 | Mathematics makes me nervous | -.454 |
| 6 | I am good at working out difficult mathematics problems | .709 |
| 7 | My teacher tells me I am good at mathematics | .394 |
| 8 | Mathematics is harder for me than any other subject | .745 |
| 9 | Mathematics makes me confused | .616 |

*Note*. All estimates are significant on the 0.001 level. For more information see Studies I – III.

# Studies I – III

Study I.
Asp, L. (under revision). Validating the mixed-worded mathematics confidence scale in TIMSS 2019: Examining its relationships to teaching quality and mathematics achievement

Study II.
Asp, L., Klapp, A., & Rolfe, V. (under revision). The associations between student-perceived teaching quality and mathematics confidence and mathematics achievement: A study of Swedish Grade 4 TIMSS 2019.

Study III.
Asp, L., Klapp, A., & Rosén M. (2025). Does teaching quality matter for Nordic primary school students' mathematics confidence and mathematics achievement? A multilevel structural equation analysis of Nordic TIMSS 2019 grade 4 data *Large-Scale Assessments in Education*, *13*(1), 1–31. https://doi.org/10.1186/s40536-025-00238-x

Previous publications:

Editors: Kjell Härnqvist and Karl-Gustaf Stukát

1. KARL-GUSTAF STUKÁT *Lekskolans inverkan på barns utveckling.* Stockholm 1966

2. URBAN DAHLLÖF *Skoldifferentiering och undervisningsförlopp.* Stockholm 1967

3. ERIK WALLIN *Spelling. Factorial and experimental studies.* Stockholm 1967

4. BENGT-ERIK ANDERSSON *Studies in adolescent behaviour. Project Yg, Youth in Göteborg.* Stockholm 1969

5. FERENCE MARTON *Structural dynamics of learning.* Stockholm 1970

6. ALLAN SVENSSON *Relative achievement. School performance in relation to intelligence, sex and home environment.* Stockholm 1971

7. GUNNI KÄRRBY *Child rearing and the development of moral structure.* Stockholm 1971

Editors: Urban Dahllöf, Kjell Härnqvist and Karl-Gustaf Stukát

8. ULF P. LUNDGREN *Frame factors and the teaching process. A contribution to curriculum theory and theory on teaching.* Stockholm 1972

9. LENNART LEVIN *Comparative studies in foreign-language teaching.* Stockholm 1972

10. RODNEY ÅSBERG *Primary education and national development.* Stockholm 1973

11. BJÖRN SANDGREN *Kreativ utveckling.* Stockholm 1974

12. CHRISTER BRUSLING *Microteaching - A concept in development.* Stockholm 1974

13. KJELL RUBENSON *Rekrytering till vuxenutbildning. En studie av kortutbildade yngre män.* Göteborg 1975

14. ROGER SÄLJÖ *Qualitative differences in learning as a function of the learner's conception of the task.* Göteborg 1975

15. LARS OWE DAHLGREN *Qualitative differences in learning as a function of content-oriented guidance.* Göteborg 1975

16. MARIE MÅNSSON *Samarbete och samarbetsförmåga. En kritisk granskning.* Lund 1975

17. JAN-ERIC GUSTAFSSON *Verbal and figural aptitudes in relation to instructional methods. Studies in aptitude - treatment interactions.* Göteborg 1976

18. MATS EKHOLM *Social utveckling i skolan. Studier och diskussion.* Göteborg 1976

19. LENNART SVENSSON *Study skill and learning.* Göteborg 1976

20. BJÖRN ANDERSSON *Science teaching and the development of thinking.* Göteborg 1976

21. JAN-ERIK PERNEMAN *Medvetenhet genom utbildning.* Göteborg 1977

Editors: Kjell Härnqvist, Ference Marton and Karl-Gustaf Stukát

22. INGA WERNERSSON *Könsdifferentiering i grundskolan.* Göteborg 1977

23. BERT AGGESTEDT & ULLA TEBELIUS *Barns upplevelser av idrott.* Göteborg 1977

24. ANDERS FRANSSON *Att rädas prov och att vilja veta.* Göteborg 1978

25. ROLAND BJÖRKBERG *Föreställningar om arbete, utveckling och livsrytm.* Göteborg 1978

26. GUNILLA SVINGBY *Läroplaner som styrmedel för svensk obligatorisk skola. Teoretisk analys och ett empiriskt bidrag.* Göteborg 1978

27. INGA ANDERSSON *Tankestilar och hemmiljö.* Göteborg 1979

28. GUNNAR STANGVIK *Self-concept and school segregation.* Göteborg 1979

29. MARGARETA KRISTIANSSON *Matematikkunskaper Lgr 62, Lgr 69.* Göteborg 1979

30. BRITT JOHANSSON *Kunskapsbehov i omvårdnadsarbete och kunskapskrav i vårdutbildning.* Göteborg 1979

31. GÖRAN PATRIKSSON *Socialisation och involvering i idrott.* Göteborg 1979

32. PETER GILL *Moral judgments of violence among Irish and Swedish adolescents.* Göteborg 1979

33. TAGE LJUNGBLAD *Förskola - grundskola i samverkan. Förutsättningar och hinder.* Göteborg 1980

34. BERNER LINDSTRÖM *Forms of representation, content and learning.* Göteborg 1980

35. CLAES-GÖRAN WENESTAM *Qualitative differences in retention.* Göteborg 1980

*36. BRITT JOHANSSON* *Pedagogiska samtal i vårdutbildning. Innehåll och språkbruk.* Göteborg 1981

37. LEIF LYBECK *Arkimedes i klassen. En ämnespedagogisk berättelse.* Göteborg 1981

38. BIÖRN HASSELGREN *Ways of apprehending children at play. A study of pre-school student teachers' development.* Göteborg 1981

39. LENNART NILSSON *Yrkesutbildning i nutidshistoriskt perspektiv. Yrkesutbildningens utveckling från skråväsendets upphörande 1846 till 1980-talet samt tankar om framtida inriktning.* Göteborg 1981

40. GUDRUN BALKE-AURELL *Changes in ability as related to educational and occupational experience.* Göteborg 1982

41. ROGER SÄLJÖ *Learning and understanding. A study of differences in constructing meaning from a text.* Göteborg 1982

42. ULLA MARKLUND *Droger och påverkan. Elevanalys som utgångspunkt för drogundervisning.* Göteborg 1983

43. SVEN SETTERLIND *Avslappningsträning i skolan. Forskningsöversikt och empiriska studier.* Göteborg 1983

44. EGIL ANDERSSON & MARIA LAWENIUS *Lärares uppfattning av undervisning.* Göteborg 1983

45. JAN THEMAN *Uppfattningar av politisk makt.* Göteborg 1983

46. INGRID PRAMLING *The child's conception of learning.* Göteborg 1983

47. PER OLOF THÅNG *Vuxenlärarens förhållningssätt till deltagarerfarenheter. En studie inom AMU.* Göteborg 1984

48. INGE JOHANSSON *Fritidspedagog på fritidshem. En yrkesgrupps syn på sitt arbete.* Göteborg 1984

49. GUNILLA SVANBERG *Medansvar i undervisning. Metoder för observation och kvalitativ analys.* Göteborg 1984

50. SVEN-ERIC REUTERBERG *Studiemedel och rekrytering till högskolan.* Göteborg 1984

51. GÖSTA DAHLGREN & LARS-ERIK OLSSON *Läsning i barnperspektiv.* Göteborg 1985

52. CHRISTINA KÄRRQVIST *Kunskapsutveckling genom experimentcentrerade dialoger i ellära.* Göteborg 1985

53. CLAES ALEXANDERSSON *Stabilitet och förändring. En empirisk studie av förhållandet mellan skolkunskap och vardagsvetande.* Göteborg 1985

54. LILLEMOR JERNQVIST *Speech regulation of motor acts as used by cerebral palsied children. Observational and experimental studies of a key feature of conductive education.* Göteborg 1985

55. SOLVEIG HÄGGLUND *Sex-typing and development in an ecological perspective.* Göteborg 1986

56. INGRID CARLGREN *Lokalt utvecklingsarbete.* Göteborg 1986

57. LARSSON, ALEXANDERSSON, HELMSTAD & THÅNG *Arbetsupplevelse och utbildningssyn hos icke facklärda.* Göteborg 1986

58. ELVI WALLDAL *Studerande vid gymnasieskolans vårdlinje. Förväntad yrkesposition, rollpåverkan, självuppfattning.* Göteborg 1986

Editors: Jan-Eric Gustafsson, Ference Marton and Karl-Gustaf Stukát

59. EIE ERICSSON *Foreign language teaching from the point of view of certain student activities.* Göteborg 1986

60. JAN HOLMER *Högre utbildning för lågutbildade i industrin.* Göteborg 1987

61. ANDERS HILL & TULLIE RABE *Psykiskt utvecklingsstörda i kommunal förskola.* Göteborg 1987

62. DAGMAR NEUMAN *The origin of arithmetic skills. A phenomenographic approach.* Göteborg 1987

63. TOMAS KROKSMARK *Fenomenografisk didaktik.* Göteborg 1987

64. ROLF LANDER *Utvärderingsforskning - till vilken nytta?* Göteborg 1987

65. TORGNY OTTOSSON *Map-reading and wayfinding.* Göteborg 1987

66. MAC MURRAY *Utbildningsexpansion, jämlikhet och avlänkning.* Göteborg 1988

67. ALBERTO NAGLE CAJES *Studievalet ur den väljandes perspektiv.* Göteborg 1988

68. GÖRAN LASSBO *Mamma - (Pappa) - barn. En utvecklingsekologisk studie av socialisation i olika familjetyper.* Göteborg 1988

69. LENA RENSTRÖM *Conceptions of matter. A phenomenographic approach.* Göteborg 1988

70. INGRID PRAMLING *Att lära barn lära.* Göteborg 1988

71. LARS FREDHOLM *Praktik som bärare av undervisnings innehåll och form. En förklaringsmodell för uppkomst av undervisningshandlingar inom en totalförsvarsorganisation.* Göteborg 1988

72. OLOF F. LUNDQUIST *Studiestöd för vuxna. Utveckling, utnyttjande, utfall.* Göteborg 1989

73. BO DAHLIN *Religionen, själen och livets mening. En fenomenografisk och existensfilosofisk studie av religionsundervisningens villkor.* Göteborg 1989

74. SUSANNE BJÖRKDAHL ORDELL *Socialarbetare. Bakgrund, utbildning och yrkesliv.* Göteborg 1990

75. EVA BJÖRCK-ÅKESSON *Measuring Sensation Seeking.* Göteborg 1990

76. ULLA-BRITT BLADINI *Från hjälpskolelärare till förändringsagent. Svensk speciallärarutbildning 1921-1981 relaterad till specialundervisningens utveckling och förändringar i speciallärarens yrkesuppgifter.* Göteborg 1990

77. ELISABET ÖHRN *Könsmönster i klassrumsinteraktion. En observations- och intervjustudie av högstadieelevers lärarkontakter.* Göteborg 1991

78. TOMAS KROKSMARK *Pedagogikens vägar till dess första svenska professur.* Göteborg 1991

Editors: Ingemar Emanuelsson, Jan-Eric Gustafsson and Ference Marton

79. ELVI WALLDAL *Problembaserad inlärning. Utvärdering av påbyggnadslinjen Utbildning i öppen hälso- och sjukvård.* Göteborg 1991

80. ULLA AXNER *Visuella perceptionssvårigheter i skolperspektiv. En longitudinell studie.* Göteborg 1991

81. BIRGITTA KULLBERG *Learning to learn to read.* Göteborg 1991

82. CLAES ANNERSTEDT *Idrottslärarna och idrottsämnet. Utveckling, mål, kompetens - ett didaktiskt perspektiv.* Göteborg 1991

83. EWA PILHAMMAR ANDERSSON *Det är vi som är dom. Sjuksköterskestuderandes föreställningar och perspektiv under utbildningstiden.* Göteborg 1991

84. ELSA NORDIN *Kunskaper och uppfattningar om maten och dess funktioner i kroppen. Kombinerad enkät- och intervjustudie i grundskolans årskurser 3, 6 och 9.* Göteborg 1992

85. VALENTIN GONZÁLEZ *On human attitudes. Root metaphors in theoretical conceptions.* Göteborg 1992

86. JAN-ERIK JOHANSSON *Metodikämnet i förskollärarutbildningen. Bidrag till en traditionsbestämning.* Göteborg 1992

87. ANN AHLBERG *Att möta matematiska problem. En belysning av barns lärande.* Göteborg 1992

88. ELLA DANIELSON *Omvårdnad och dess psykosociala inslag. Sjuksköterskestuderandes uppfattningar av centrala termer och reaktioner inför en omvårdnadssituation.* Göteborg 1992

89. SHIRLEY BOOTH *Learning to program. A phenomenographic perspective.* Göteborg 1992

90. EVA BJÖRCK-ÅKESON *Samspel mellan små barn med rörelsehinder och talhandikapp och deras föräldrar - en longitudinell studie.* Göteborg 1992

91. KARIN DAHLBERG *Helhetssyn i vården. En uppgift för sjuksköterskeutbildningen.* 1992

92. RIGMOR ERIKSSON *Teaching Language Learning. In-service training for communicative teaching and self directed learning in English as a foreign language.* 1993

93. KJELL HÄRENSTAM *Skolboks-islam. Analys av bilden av islam i läroböcker i religionskunskap.* Göteborg 1993.

94. INGRID PRAMLING *Kunnandets grunder. Prövning av en fenomenografisk ansats till att utveckla barns sätt att uppfatta sin omvärld.* Göteborg 1994.

95. MARIANNE HANSSON SCHERMAN *Att vägra vara sjuk. En longitudinell studie av förhållningssätt till astma/allergi.* Göteborg 1994

96. MIKAEL ALEXANDERSSON *Metod och medvetande.* Göteborg 1994

97. GUN UNENGE *Pappor i föräldrakooperativa daghem. En deskriptiv studie av pappors medverkan.* Göteborg 1994

98. BJÖRN SJÖSTRÖM *Assessing acute postoperative pain. Assessment strategies and quality in relation to clinical experience and professional role.* Göteborg 1995

99. MAJ ARVIDSSON *Lärares orsaks- och åtgärdstankar om elever med svårigheter.* Göteborg 1995

100. DENNIS BEACH *Making sense of the problems of change: An ethnographic study of a teacher education reform.* Göteborg 1995.

101. WOLMAR CHRISTENSSON *Subjektiv bedömning - som besluts och handlingsunderlag.* Göteborg 1995

102. SONJA KIHLSTRÖM *Att vara förskollärare. Om yrkets pedagogiska innebörder.* Göteborg 1995

103. MARITA LINDAHL *Inlärning och erfarande. Ettåringars möte med förskolans värld.* Göteborg. 1996

104. GÖRAN FOLKESTAD *Computer Based Creative Music Making - Young Peoples´ Music in the Digital Age.* Göteborg 1996

105. EVA EKEBLAD *Children • Learning • Numbers. A phenomenographic excursion into first-grade children's arithmetic.* Göteborg 1996

106. HELGE STRÖMDAHL *On mole and amount of substance. A study of the dynamics of concept formation and concept attainment.* Göteborg 1996

107. MARGARETA HAMMARSTRÖM *Varför inte högskola? En longitudinell studie av olika faktorers betydelse för studiebegåvade ungdomars utbildningskarriär.* Göteborg 1996

108. BJÖRN MÅRDÉN *Rektorers tänkande. En kritisk betraktelse av skolledarskap.* Göteborg 1996

109. GLORIA DALL'ALBA & BIÖRN HASSELGREN (EDS) *Reflections on Phenomenography - Toward a Methodology?* Göteborg 1996

110. ELISABETH HESSLEFORS ARKTOFT *I ord och handling. Innebörder av "att anknyta till elevers erfarenheter", uttryckta av lärare.* Göteborg 1996

111. BARBRO STRÖMBERG *Professionellt förhållningssätt hos läkare och sjuksköterskor. En studie av uppfattningar.* Göteborg 1997

112. HARRIET AXELSSON *Våga lära. Om lärare som förändrar sin miljöundervisning.* Göteborg 1997

113. ANN AHLBERG *Children's ways of handling and experiencing numbers.* Göteborg 1997

114. HUGO WIKSTRÖM *Att förstå förändring. Modellbyggande, simulering och gymnasieelevers lärande.* Göteborg 1997

115. DORIS AXELSEN *Listening to recorded music. Habits and motivation among high-school students.* Göteborg 1997.

116. EWA PILHAMMAR ANDERSSON *Handledning av sjuksköterskestuderande i klinisk praktik.* Göteborg 1997

117. OWE STRÅHLMAN *Elitidrott, karriär och avslutning.* Göteborg 1997

118. AINA TULLBERG T*eaching the 'mole'. A phenomenographic inquiry into the didactics of chemistry.* Göteborg 1997.

119. DENNIS BEACH *Symbolic Control and Power Relay Learning in Higher Professional Education.* Göteborg 1997

120. HANS-ÅKE SCHERP *Utmanande eller utmanat ledarskap. Rektor, organisationen och förändrat undervisningsmönster i gymnasieskolan.* Göteborg 1998

121. STAFFAN STUKÁT *Lärares planering under och efter utbildningen.* Göteborg 1998

122. BIRGIT LENDAHLS ROSENDAHL *Examensarbetets innebörder. En studie av blivande lärares utsagor.* Göteborg 1998

123. ANN AHLBERG *Meeting Mathematics. Educational studies with young children.* Göteborg 1998

124. MONICA ROSÉN *Gender Differences in Patterns of Knowledge.* Göteborg 1998.

125. HANS BIRNIK L*ärare- elevrelationen. Ett relationistiskt perspektiv.* Göteborg 1998

126. MARGRETH HILL *Kompetent för "det nya arbetslivet"? Tre gymnasieklasser reflekterar över och diskuterar yrkesförberedande studier.* Göteborg 1998

127. LISBETH ÅBERG-BENGTSSON *Entering a Graphicate Society. Young Children Learning Graphs and Charts.* Göteborg 1998

128. MELVIN FEFFER *The Conflict of Equals: A Constructionist View of Personality Development.* Göteborg 1999

129. ULLA RUNESSON *Variationens pedagogik. Skilda sätt att behandla ett matematiskt innehåll.* Göteborg 1999

130. SILWA CLAESSON *"Hur tänker du då?" Empiriska studier om relationen mellan forskning om elevuppfattningar och lärares undervisning.* Göteborg 1999

131. MONICA HANSEN *Yrkeskulturer i möte. Läraren, fritidspedagogen och samverkan.* Göteborg 1999

132. JAN THELIANDER *Att studera arbetets förändring under kapitalismen. Ure och Taylor i pedagogiskt perspektiv.* Göteborg 1999

133. TOMAS SAAR *Musikens dimensioner - en studie av unga musikers lärande.* Göteborg 1999

134. GLEN HELMSTAD *Understandings of understanding. An inquiry concerning experiential conditions for developmental learning.* Göteborg 1999

135. MARGARETA HOLMEGAARD *Språkmedvetenhet och ordinlärning. Lärare och inlärare reflekterar kring en betydelsefältsövning i svenska som andraspråk.* Göteborg 1999

136. ALYSON MCGEE *Investigating Language Anxiety through Action Inquiry: Developing Good Research Practices.* Göteborg 1999

137. EVA GANNERUD *Genusperspektiv på lärargärning. Om kvinnliga klasslärares liv och arbete.* Göteborg 1999

138. TELLERVO KOPARE *Att rida stormen ut. Förlossningsberättelser i Finnmark och Sápmi.* Göteborg 1999

139. MAJA SÖDERBÄCK *Encountering Parents. Professional Action Styles among Nurses in Pediatric Care.* Göteborg 1999

140. AIRI ROVIO - JOHANSSON *Being Good at Teaching. Exploring different ways of handling the same subject in Higher Education.* Göteborg 1999

141. EVA JOHANSSON *Etik i små barns värld. Om värden och normer bland de yngsta barnen i förskolan.* Göteborg 1999

142. KENNERT ORLENIUS *Förståelsens paradox. Yrkeserfarenhetens betydelse när förskollärare blir grundskollärare.* Göteborg 1999.

143. BJÖRN MÅRDÉN *De nya hälsomissionärerna – rörelser i korsvägen mellan pedagogik och hälsopromotion.* Göteborg 1999

144. MARGARETA CARLÉN *Kunskapslyft eller avbytarbänk? Möten med industriarbetare om utbildning för arbete.* Göteborg 1999

145. MARIA NYSTRÖM *Allvarligt psykiskt störda människors vardagliga tillvaro.* Göteborg 1999

146. ANN-KATRIN JAKOBSSON *Motivation och inlärning ur genusperspektiv. En studie av gymnasieelever på teoretiska linjer/program.* Göteborg 2000

147. JOANNA GIOTA *Adolescents' perceptions of school and reasons for learning.* Göteborg 2000

148. BERIT CARLSTEDT *Cognitive abilities – aspects of structure, process and measurement.* Göteborg 2000

149. MONICA REICHENBERG *Röst och kausalitet i lärobokstexter. En studie av elevers förståelse av olika textversioner.* Göteborg 2000

150. HELENA ÅBERG *Sustainable waste management in households – from international policy to everyday practice. Experiences from two Swedish field studies.* Göteborg 2000

151. BJÖRN SJÖSTRÖM & BRITT JOHANSSON *Ambulanssjukvård. Ambulanssjukvårdares och läkares perspektiv.* Göteborg 2000

152. AGNETA NILSSON *Omvårdnadskompetens inom hemsjukvård – en deskriptiv studie.* Göteborg 2001

153. ULLA LÖFSTEDT *Förskolan som lärandekontext för barns bildskapande.* Göteborg 2001

154. JÖRGEN DIMENÄS *Innehåll och interaktion. Om elevers lärande i naturvetenskaplig undervisning.* Göteborg 2001

155. BRITT MARIE APELGREN *Foreign Language Teachers' Voices. Personal Theories and Experiences of Change in Teaching English as a Foreign Language in Sweden.* Göteborg 2001

156. CHRISTINA CLIFFORDSON *Assessing empathy: Measurement characteristics and interviewer effects.* Göteborg 2001

157. INGER BERGGREN *Identitet, kön och klass. Hur arbetarflickor formar sin identitet.* Göteborg 2001

158. CARINA FURÅKER *Styrning och visioner – sjuksköterskeutbildning i förändring.* Göteborg 2001

159. INGER BERNDTSSON *Förskjutna horisonter. Livsförändring och lärande i samband med synnedsättning eller blindhet.* Göteborg 2001

160. SONJA SHERIDAN *Pedagogical Quality in Preschool. An issue of perspectives.* Göteborg 2001

161. JAN BAHLENBERG *Den otroliga verkligheten sätter spår. Om Carlo Derkerts liv och konstpedagogiska gärning.* Göteborg 2001

162. FRANK BACH *Om ljuset i tillvaron. Ett undervisningsexperiment inom optik.* Göteborg 2001

163. PIA WILLIAMS *Barn lär av varandra. Samlärande i förskola och skola.* Göteborg 2001

164. VIGDIS GRANUM *Studentenes forestillinger om sykepleie som fag og funksjon.* Göteborg 2001

165. MARIT ALVESTAD *Den komplekse planlegginga. Førskolelærarar om pedagogisk planlegging og praksis.* Göteborg 2001

166. GIRMA BERHANU *Learning-In-Context. An Ethnographic Investigation of Mediated Learning Experiences among Ethiopian Jews in Israel.* Göteborg 2001.

167. OLLE ESKILSSON *En longitudinell studie av 10 – 12-åringars förståelse av materiens förändringar.* Göteborg 2001

168. JONAS EMANUELSSON *En fråga om frågor. Hur lärares frågor i klassrummet gör det möjligt att få reda på elevernas sätt att förstå det som undervisningen behandlar i matematik och naturvetenskap.* Göteborg 2001

169. BIRGITTA GEDDA *Den offentliga hemligheten. En studie om sjuksköterskans pedagogiska funktion och kompetens i folkhälsoarbetet.* Göteborg 2001

170. FEBE FRIBERG *Pedagogiska möten mellan patienter och sjuksköterskor på en medicinsk vårdavdelning. Mot en vårddidaktik på livsvärldsgrund.* Göteborg 2001

171. MADELEINE BERGH *Medvetenhet om bemötande. En studie om sjuksköterskans pedagogiska funktion och kompetens i närståendeundervisning.* Göteborg 2002

172. HENRIK ERIKSSON *Den diplomatiska punkten – maskulinitet som kroppsligt identitetsskapande projekt i svensk sjuksköterskeutbildning.* Göteborg 2002

173. SOLVEIG LUNDGREN *I spåren av en bemanningsförändring. En studie av sjuksköterskors arbete på en kirurgisk vårdavdelning.* Göteborg 2002

174. BIRGITTA DAVIDSSON *Mellan soffan och katedern. En studie av hur förskollärare och grundskollärare utvecklar pedagogisk integration mellan förskola och skola.* Göteborg 2002

175. KARI SØNDENÅ *Tradisjon og Transcendens – ein fenomenologisk studie av refleksjon i norsk førskulelærarutdanning.* Göteborg 2002

176. CHRISTINE BENTLEY *The Roots of Variation of English-Teaching. A Phenomenographic Study Founded on an Alternative Basic Assumption.* Göteborg 2002

177. ÅSA MÄKITALO *Categorizing Work: Knowing, Arguing, and Social Dilemmas in Vocational Guidance.* Göteborg 2002

178. MARITA LINDAHL *VÅRDA – VÄGLEDA – LÄRA. Effektstudie av ett interventionsprogram för pedagogers lärande i förskolemiljön.* Göteborg 2002

179. CHRISTINA BERG *Influences on schoolchildren's dietary selection. Focus on fat and fibre at breakfast.* Göteborg 2002

180. MARGARETA ASP *Vila och lärande om vila. En studie på livsvärldsfenomenologisk grund.* Göteborg 2002

181. FERENCE MARTON & PAUL MORRIS (EDS) *What matters? Discovering critical contitions of classroom learning.* Göteborg 2002

182. ROLAND SEVERIN *Dom vet vad dom talar om. En intervjustudie om elevers uppfattningar av begreppen makt och samhällsförändring.* Göteborg 2002

Editors: Björn Andersson, Jan Holmer and
Ingrid Pramling Samuelsson

183. MARLÉNE JOHANSSON *Slöjdpraktik i skolan – hand, tanke, kommunikation och andra medierande redskap.* Göteborg 2002

184. INGRID SANDEROTH *Om lust att lära i skolan: En analys av dokument och klass 8y*. Göteborg 2002

185. INGA-LILL JAKOBSSON *Diagnos i skolan. En studie av skolsituationer för elever med syndromdiagnos*. Göteborg 2002

186. EVA-CARIN LINDGREN *Empowering Young Female Athletes – A Possible Challenge to the Male Hegemony in Sport. A Descriptive and Interventional Study*. Göteborg 2002

187. HANS RYSTEDT *Bridging practices. Simulations in education for the health-care professions*. Göteborg 2002

188. MARGARETA EKBORG *Naturvetenskaplig utbildning för hållbar utveckling? En longitudinell studie av hur studenter på grunskollärarprogrammet utvecklar för miljöundervisning relevanta kunskaper i naturkunskap*. Göteborg 2002

189. ANETTE SANDBERG *Vuxnas lekvärld. En studie om vuxnas erfarenheter av lek*. Göteborg 2002

190. GUNLÖG BREDÄNGE *Gränslös pedagog. Fyra studier om utländska lärare i svensk skola*. Göteborg 2003

191. PER-OLOF BENTLEY *Mathematics Teachers and Their Teaching. A Survey Study*. Göteborg 2003

192. KERSTIN NILSSON *MANDAT – MAKT – MANAGEMENT. En studie av hur vårdenhetschefers ledarskap konstrueras*. Göteborg 2003

193. YANG YANG *Measuring Socioeconomic Status and its Effects at Individual and Collective Levels: A Cross-Country Comparison*. Göteborg 2003

194. KNUT VOLDEN *Mediekunnskap som mediekritikk*. Göteborg 2003.

195. LOTTA LAGER-NYQVIST *Att göra det man kan – en longitudinell studie av hur sju lärarstudenter utvecklar sin undervisning och formar sin lärarroll i naturvetenskap*. Göteborg 2003

196. BRITT LINDAHL *Lust att lära naturvetenskap och teknik? En longitudinell studie om vägen till gymnasiet*. Göteborg 2003

197. ANN ZETTERQVIST *Ämnesdidaktisk kompetens i evolutionsbiologi. En intervjuundersökning med no/biologilärare*. Göteborg 2003

198. ELSIE ANDERBERG *Språkanvändningens funktion vid utveckling av kunskap om objekt*. Göteborg 2003.

199. JAN GUSTAFSSON *Integration som text, diskursiv och social praktik. En policyetnografisk fallstudie av mötet mellan skolan och förskoleklassen*. Göteborg 2003.

200. EVELYN HERMANSSON *Akademisering och professionalisering – barnmorskans utbildning i förändring*. Göteborg 2003

201. KERSTIN VON BRÖMSSEN *Tolkningar, förhandlingar och tystnader. Elevers tal om religion i det mångkulturella och postkoloniala rummet*. Göteborg 2003

202. MARIANNE LINDBLAD FRIDH *Från allmänsjuksköterska till specialistsjuksköterska inom intensivvård. En studie av erfarenheter från specialistutbildningen och från den första yrkesverksamma tiden inom intensivvården*. Göteborg 2003

203. BARBRO CARLI *The Making and Breaking of a Female Culture: The History of Swedish Physical Education 'in a Different Voice'*. Göteborg 2003

204. ELISABETH DAHLBORG-LYCKHAGE *"Systers" konstruktion och mumifiering – i TV-serier och i studenters föreställningar*. Göteborg 2003

205. ULLA HELLSTRÖM MUHLI *Att överbrygga perspektiv. En studie av behovsbedömningssamtal inom äldreinriktat socialt arbete*. Göteborg 2003

206. KRISTINA AHLBERG *Synvändor. Universitetsstudenters berättelser om kvalitativa förändringar av sätt att erfara situationers mening under utbildningspraktik*. Göteborg 2004

207. JONAS IVARSSON *Renderings & Reasoning: Studying artifacts in human knowing*. Göteborg 2004

208. MADELEINE LÖWING *Matematikundervisningens konkreta gestaltning. En studie av kommunikationen lärare – elev och matematiklektionens didaktiska ramar*. Göteborg 2004

209. PIJA EKSTRÖM *Makten att definiera. En studie av hur beslutsfattare formulerar villkor för specialpedagogisk verksamhet*. Göteborg 2004

210. CARIN ROOS *Skriftspråkande döva barn. En studie om skriftspråkligt lärande i förskola och skola*. Göteborg 2004

211. JONAS LINDEROTH *Datorspelandets mening. Bortom idén om den interaktiva illusionen*. Göteborg 2004

212. ANITA WALLIN *Evolutionsteorin i klassrummet. På väg mot en ämnesdidaktisk teori för undervisning i biologisk evolution*. Göteborg 2004

213. EVA HJÖRNE *Excluding for inclusion? Negotiating school careers and identities in pupil welfare settings in the Swedish school*. Göteborg 2004

214. MARIE BLIDING *Inneslutandets och uteslutandets praktik. En studie av barns relationsarbete i skolan*. Göteborg 2004

215. LARS-ERIK.JONSSON *Appropriating Technologies in Educational Practices. Studies in the Contexts of Compulsory Education, Higher Education, and Fighter Pilot Training*. Göteborg 2004

216. MIA KARLSSON *An ITiS Teacher Team as a Community of Practice*. Göteborg 2004

217. SILWA CLAESSON *Lärares levda kunskap*. Göteborg 2004

218. GUN-BRITT WÄRVIK *Ambitioner att förändra och artefakters verkan. Gränsskapande och stabiliserande praktiker på produktionsgolvet*. Göteborg 2004

219. KARIN LUMSDEN WASS *Vuxenutbildning i omvandling. Kunskapslyftet som ett sätt att organisera förnyelse.* Göteborg 2004

220. LENA DAHL *Amningspraktikens villkor. En intervjustudie av en grupp kvinnors föreställningar på och erfarenheter av amning.* Göteborg 2004

221. ULRIC BJÖRCK *Distributed Problem-Based Learning. Studies of a Pedagogical Model in Practice.* Göteborg 2004

222. ANNEKA KNUTSSON *"To the best of your knowledge and for the good of your neighbour". A study of traditional birth attendants in Addis Ababa, Ethiopia.* Göteborg 2004

223. MARIANNE DOVEMARK *Ansvar – flexibilitet – valfrihet. En etnografisk studie om en skola i förändring.* Göteborg 2004

224. BJÖRN HAGLUND *Traditioner i möte. En kvalitativ studie av fritidspedagogers arbete med samlingar i skolan.* Göteborg 2004

225. ANN-CHARLOTTE MÅRDSJÖ *Lärandets skiftande innebörder – uttryckta av förskollärare i vidareutbildning.* Göteborg 2005

226. INGRID GRUNDÉN *Att återerövra kroppen. En studie av livet efter en ryggmärgsskada.* Göteborg 2005

227. KARIN GUSTAFSSON & ELISABETH MELLGREN *Barns skriftspråkande – att bli en skrivande och läsande person.* Göteborg 2005

228. GUNNAR NILSSON *Att äga π. Praxisnära studier av lärarstudenters arbete med geometrilaborationer.* Göteborg 2005.

229. BENGT LINDGREN *Bild, visualitet och vetande. Diskussion om bild som ett kunskapsfält inom utbildning.* Göteborg 2005

230. PETRA ANGERVALL *Jämställdhetsarbetets pedagogik. Dilemman och paradoxer i arbetet med jämställdhet på ett företag och ett universitet.* Göteborg 2005

231. LENNART MAGNUSSON *Designing a responsive support service for family carers of frail older people using ICT.* Göteborg 2005

232. MONICA REICHENBERG *Gymnasieelever samtalar kring facktexter. En studie av textsamtal med goda och svaga läsare.* Göteborg 2005

233. ULRIKA WOLFF *Characteristics and varieties of poor readers.* Göteborg 2005

234. CECILIA NIELSEN *Mellan fakticitet och projekt. Läs- och skrivsvårigheter och strävan att övervinna dem.* Göteborg 2005.

235. BERITH HEDBERG *Decision Making and Communication in Nursing Practice. Aspects of Nursing Competence.* Göteborg 2005

236. MONICA ROSÉN, EVA MYRBERG & JAN-ERIC GUSTAFSSON *Läskompetens i skolår 3 och 4. Nationell rapport från PIRLS 2001 i Sverige. The IEA Progress in International Reading Literacy Study.* Göteborg 2005

237. INGRID HENNING LOEB *Utveckling och förändring i kommunal vuxenutbildning. En yrkeshistorisk ingång med berättelser om lärarbanor.* Göteborg 2006.

238. NIKLAS PRAMLING *Minding metaphors: Using figurative language in learning to represent.* Göteborg 2006

239. KONSTANTIN KOUGIOUMTZIS *Lärarkulturer och professionskoder. En komparativ studie av idrottslärare i Sverige och Grekland.* Göteborg 2006

240. STEN BÅTH *Kvalifikation och medborgarfostran. En analys av reformtexter avseende gymnasieskolans samhällsuppdrag.* Göteborg 2006.

241. EVA MYRBERG *Fristående skolor i Sverige – Effekter på 9-10-åriga elevers läsförmåga.* Göteborg 2006

242. MARY-ANNE HOLFVE-SABEL *Attitudes towards Swedish comprehensive school. Comparisons over time and between classrooms in grade 6.* Göteborg 2006

243. CAROLINE BERGGREN *Entering Higher Education – Gender and Class Perspectives.* Göteborg 2006

244. CRISTINA THORNELL & CARL OLIVESTAM *Kulturmöte i centralafrikansk kontext med kyrkan som arena.* Göteborg 2006

245. ARVID TREEKREM *Att leda som man lär. En arbetsmiljöpedagogisk studie av toppledares ideologier om ledarskapets taktiska potentialer.* Göteborg 2006

246. EVA GANNERUD & KARIN RÖNNERMAN *Innehåll och innebörd i lärares arbete i förskola och skola – en fallstudie ur ett genusperspektiv.* Göteborg 2006

247. JOHANNES LUNNEBLAD *Förskolan och mångfalden – en etnografisk studie på en förskola i ett multietniskt område.* Göteborg 2006

248. LISA ASP-ONSJÖ *Åtgärdsprogram – dokument eller verktyg? En fallstudie i en kommun.* Göteborg 2006

249. EVA JOHANSSON & INGRID PRAMLING SAMUELSSON *Lek och läroplan. Möten mellan barn och lärare i förskola och skola.* Göteborg 2006

250. INGER BJÖRNELOO *Innebörder av hållbar utveckling. En studie av lärares utsagor om undervisning.* Göteborg 2006

251. EVA JOHANSSON *Etiska överenskommelser i förskolebarns världar.* Göteborg 2006

252. MONICA PETERSSON *Att genuszappa på säker eller osäker mark. Hem- och konsumentkunskap ur ett könsperspektiv.* Göteborg 2007

253. INGELA OLSSON *Handlingskompetens eller inlärd hjälplöshet? Lärandeprocesser hos verkstadsindustriarbetare.* Göteborg 2007

254. HELENA PEDERSEN *The School and the Animal Other. An Ethnography of human-animal relations in education.* Göteborg 2007

255. ELIN ERIKSEN ØDEGAARD *Meningsskaping i barnehagen. Innhold og bruk av barns og voksnes samtalefortellinger.* Göteborg 2007

256. ANNA KLERFELT *Barns multimediala berättande. En länk mellan mediakultur och pedagogisk praktik.* Göteborg 2007

257. PETER ERLANDSON *Docile bodies and imaginary minds: on Schön's reflection-in-action.* Göteborg 2007

258. SONJA SHERIDAN OCH PIA WILLIAMS *Dimensioner av konstruktiv konkurrens. Konstruktiva konkurrensformer i förskola, skola och gymnasium.* Göteborg 2007

259. INGELA ANDREASSON *Elevplanen som text - om identitet, genus, makt och styrning i skolans elevdokumentation.* Göteborg 2007

Editors: Jan-Eric Gustafsson, Annika Härenstam and Ingrid Pramling Samuelsson

260. ANN-SOFIE HOLM *Relationer i skolan. En studie av femininiteter och maskuliniteter i år 9.* Göteborg 2008

261. LARS-ERIK NILSSON *But can't you see they are lying: Student moral positions and ethical practices in the wake of technological change.* Göteborg 2008

262. JOHAN HÄGGSTRÖM *Teaching systems of linear equations in Sweden and China: What is made possible to learn?* Göteborg 2008

263. GUNILLA GRANATH *Milda makter! Utvecklingssamtal och loggböcker som disciplineringstekniker.* Göteborg 2008

264. KARIN GRAHN *Flickor och pojkar i idrottens läromedel. Konstruktioner av genus i ungdomstränarutbildningen.* Göteborg 2008.

265. PER-OLOF BENTLEY *Mathematics Teachers and Their Conceptual Models. A New Field of Research.* Göteborg 2008

266. SUSANNE GUSTAVSSON *Motstånd och mening. Innebörd i blivande lärares seminariesamtal.* Göteborg 2008

267. ANITA MATTSSON *Flexibel utbildning i praktiken. En fallstudie av pedagogiska processer i en distansutbildning med en öppen design för samarbetslärande.* Göteborg 2008

268. ANETTE EMILSON *Det önskvärda barnet. Fostran uttryckt i vardagliga kommunikationshandlingar mellan lärare och barn i förskolan.* Göteborg 2008

269. ALLI KLAPP LEKHOLM *Grades and grade assignment: effects of student and school charachterisitcs.* Göteborg 2008

270. ELISABETH BJÖRKLUND *Att erövra litteracitet. Små barns kommunikativa möten med berättande, bilder, text och tecken i förskolan.* Göteborg 2008

271. EVA NYBERG *Om livets kontinuitet. Undervisning och lärande om växters och djurs livscykler - en fallstudie i årskurs 5.* Göteborg 2008

272. CANCELLED

273. ANITA NORLUND *Kritisk sakprosaläsning i gymnasieskolan. Didaktiska perspektiv på läroböcker, lärare och nationella prov.* Göteborg 2009

274. AGNETA SIMEONSDOTTER SVENSSON *Den pedagogiska samlingen i förskoleklasen. Barns olika sätt att erfara och hantera svårigheter.* Göteborg 2009

275. ANITA ERIKSSON *Om teori och praktik i lärarutbildningen. En etnografisk och diskursanalytisk studie.* Göteborg 2009

276. MARIA HJALMARSSON *Lärarprofessionens genusordning. En studie av lärares uppfattningar om arbetsuppgifter, kompetens och förväntningar.* Göteborg 2009.

277. ANNE DRAGEMARK OSCARSON *Self-Assessement of Writing in Learning English as a Foreign Language. A Study at the Upper Secondary School Level.* Göteborg 2009

278. ANNIKA LANTZ-ANDERSSON *Framing in Educational Practices. Learning Activity, Digital Technology and the Logic of Situated Action.* Göteborg 2009

279. RAUNI KARLSSON *Demokratiska värden i förskolebarns vardag.* Göteborg 2009

280. ELISABETH FRANK *Läsförmågan bland 9-10-åringar. Betydelsen av skolklimat, hem- och skolsamverkan, lärarkompetens och elevers hembakgrund.* Göteborg 2009

281. MONICA JOHANSSON *Anpassning och motstånd. En etnografisk studie av gymnasieelevers institutionella identitetsskapande.* Göteborg 2009

282. MONA NILSEN *Food for Thought. Communication and the transformation of work experience in web-based in-service training.* Göteborg 2009

283. INGA WERNERSSON (RED) *Genus i förskola och skola. Förändringar i policy, perspektiv och praktik.* Göteborg 2009

284. SONJA SHERIDAN, INGRID PRAMLING SAMUELSSON & EVA JOHANSSON (RED) *Barns tidiga lärande. En tvärsnittsstudie om förskolan som miljö för barns lärande.* Göteborg 2009

285. MARIE HJALMARSSON *Lojalitet och motstånd - anställdas agerande i ett föränderligt hemtjänstarbete.* Göteborg 2009.

286. ANETTE OLIN *Skolans mötespraktik - en studie om skolutveckling genom yrkesverksammas förståelse.* Göteborg 2009

287. MIRELLA FORSBERG AHLCRONA *Handdockans kommunikativa potential som medierande redskap i förskolan.* Göteborg 2009

288. CLAS OLANDER *Towards an interlanguage of biological evolution: Exploring students´ talk and writing as* an arena for sense-making. Göteborg 2010

Editors: Jan-Eric Gustafsson, Åke Ingerman and Ingrid Pramling Samuelsson

289. PETER HASSELSKOG *Slöjdlärares förhållningssätt i undervisningen.* Göteborg 2010

290. HILLEVI PRELL *Promoting dietary change. Intervening in school and recognizing health messages in commercials.* Göteborg 2010

291. DAVOUD MASOUMI *Quality Within E-learning in a Cultural Context. The case of Iran.* Göteborg 2010

292. YLVA ODENBRING *Kramar, kategoriseringar och hjälpfröknar. Könskonstruktioner i interaktion i förskola, förskoleklass och skolår ett.* Göteborg 2010

293. ANGELIKA KULLBERG *What is taught and what is learned. Professional insights gained and shared by teachers of mathematics.* Göteborg 2010

294. TORGEIR ALVESTAD *Barnehagens relasjonelle verden - små barn som kompetente aktører i produktive forhandlinger.* Göteborg 2010

295. SYLVI VIGMO *New spaces for Language Learning. A study of student interaction in media production in English.* Göteborg 2010

296. CAROLINE RUNESDOTTER *I otakt med tiden? Folkhögskolorna i ett föränderligt fält.* Göteborg 2010

297. BIRGITTA KULLBERG *En etnografisk studie i en thailändsk grundskola på en ö i södra Thailand. I sökandet efter en framtid då nuet har nog av sitt.* Göteborg 2010

298. GUSTAV LYMER *The work of critique in architectural education.* Göteborg 2010

299. ANETTE HELLMAN *Kan Batman vara rosa? Förhandlingar om pojkighet och normalitet på en förskola.* Göteborg 2010

300. ANNIKA BERGVIKEN-RENSFELDT *Opening higher education. Discursive transformations of distance and higher education government.* Göteborg 2010

301. GETAHUN YACOB ABRAHAM *Education for Democracy? Life Orientation: Lessons on Leadership Qualities and Voting in South African Comprehensive Schools.* Göteborg 2010

302. LENA SJÖBERG *Bäst i klassen? Lärare och elever i svenska och europeiska policytexter.* Göteborg 2011

303. ANNA POST *Nordic stakeholders and sustainable catering.* Göteborg 2011

304. CECILIA KILHAMN *Making Sense of Negative Numbers.* Göteborg 2011

305. ALLAN SVENSSON (RED) *Utvärdering Genom Uppföljning. Longitudinell individforskning under ett halvsekel.* Göteborg 2011

306. NADJA CARLSSON *I kamp med skriftspråket. Vuxenstuderande med läs- och skrivsvårigheter i ett livsvärldsperspektiv.* Göteborg 2011

307. AUD TORILL MELAND *Ansvar for egen læring. Intensjoner og realiteter ved en norsk videregående skole.* Göteborg 2011

308. EVA NYBERG *Folkbildning för demokrati. Colombianska kvinnors perspektiv på kunskap som förändringskraft.* Göteborg 2011

309. SUSANNE THULIN *Lärares tal och barns nyfikenhet. Kommunikation om naturvetenskapliga innehåll i förskolan.* Göteborg 2011

310. LENA FRIDLUND *Interkulturell undervisning – ett pedagogiskt dilemma. Talet om undervisning i svenska som andraspråk och i förberedelseklass.* Göteborg 2011

311. TARJA ALATALO *Skicklig läs- och skrivundervisning i åk 1-3. Om lärares möjligheter och hinder.* Göteborg 2011

312. LISE-LOTTE BJERVÅS *Samtal om barn och pedagogisk dokumentation som bedömningspraktik i förskolan. En diskursanalys.* Göteborg 2011

313. ÅSE HANSSON *Ansvar för matematiklärande. Effekter av undervisningsansvar i det flerspråkiga klassrummet.* Göteborg 2011

314. MARIA REIS *Att ordna, från ordning till ordning. Yngre förskolebarns matematiserande.* Göteborg 2011

315. BENIAMIN KNUTSSON *Curriculum in the Era of Global Development – Historical Legacies and Contemporary Approaches.* Göteborg 2011

316. EVA WEST *Undervisning och lärande i naturvetenskap. Elevers lärande i relation till en forskningsbaserad undervisning om ljud, hörsel och hälsa.* Göteborg 2011

317. SIGNILD RISENFORS *Gymnasieungdomars livstolkande.* Göteborg 2011

318. EVA JOHANSSON & DONNA BERTHELSEN (Ed.) *Spaces for Solidarity and Individualism in Educational Contexts.* Göteborg 2012

319. ALASTAIR HENRY *L3 Motivation.* Göteborg 2012

320. ANN PARINDER *Ungdomars matval – erfarenheter, visioner och miljöargument i eget hushåll.* Göteborg 2012

321. ANNE KULTTI *Flerspråkiga barn i förskolan: Villkor för deltagande och lärande.* Göteborg 2012

322. BO-LENNART EKSTRÖM *Kontroversen om DAMP. En kontroversstudie av vetenskapligt gränsarbete och översättning mellan olika kunskapsparadigm.* Göteborg 2012

323. MUN LING LO *Variation Theory and the Improvement of Teaching and Learning.* Göteborg 2012

324. ULLA ANDRÉN *Self-awareness and self-knowledge in professions. Something we are or a skill we learn.* Göteborg 2012

325. KERSTIN SIGNERT *Variation och invarians i Maria Montessoris sinnestränande materiel.* Göteborg 2012

326. INGEMAR GERRBO *Idén om en skola för alla och specialpedagogisk organisering i praktiken.* Göteborg 2012

327. PATRIK LILJA *Contextualizing inquiry. Negotiations of tasks, tools and actions in an upper secondary classroom.* Göteborg 2012

328. STEFAN JOHANSSON *On the Validity of Reading Assessments: Relationships Between Teacher Judgements, External Tests and Pupil Self-assessments.* Göteborg 2013

329. STEFAN PETTERSSON *Nutrition in Olympic Combat Sports. Elite athletes' dietary intake, hydration status and experiences of weight regulation.* Göteborg 2013

330. LINDA BRADLEY *Language learning and technology – student activities in web-based environments.* Göteborg 2013

331. KALLE JONASSON *Sport Has Never Been Modern.* Göteborg 2013

332. MONICA HARALDSSON STRÄNG *Yngre elevers lärande om natur. En studie av kommunikation om modeller i institutionella kontexter.* Göteborg 2013

333. ANN VALENTIN KVIST *Immigrant Groups and Cognitive Tests – Validity Issues in Relation to Vocational Training.* Göteborg 2013

334. ULRIKA BENNERSTEDT *Knowledge at play. Studies of games as members' matters.* Göteborg 2013

335. EVA ÄRLEMALM-HAGSÉR *Engagerade i världens bästa? Lärande för hållbarhet i förskolan.* Göteborg 2013

336. ANNA-KARIN WYNDHAMN *Tänka fritt, tänka rätt. En studie om värdeöverföring och kritiskt tänkande i gymnasieskolans undervisning.* Göteborg 2013

337. LENA TYRÈN *"Vi får ju inte riktigt förutsättningarna för att genomföra det som vi vill." En studie om lärares möjligheter och hinder till förändring och förbättring i praktiken.* Göteborg 2013

Editors: Jan-Eric Gustafsson, Åke Ingerman and Pia Williams

338. ANNIKA LILJA *Förtroendefulla relationer mellan lärare och elev.* Göteborg 2013

339. MAGNUS LEVINSSON *Evidens och existens. Evidensbaserad undervisning i ljuset av lärares erfarenheter.* Göteborg 2013

340. ANNELI SCHWARTZ *Pedagogik, plats och prestationer. En etnografisk studie om en skola i förorten.* Göteborg 2013

341. ELISABET ÖHRN och LISBETH LUNDAHL (red) *Kön och karriär i akademin. En studie inom det utbildningsvetenskapliga fältet.* Göteborg 2013

342. RICHARD BALDWIN *Changing practice by reform. The recontextualisation of the Bologna process in teacher education.* Göteborg 2013

343. AGNETA JONSSON *Att skapa läroplan för de yngsta barnen i förskolan. Barns perspektiv och nuets didaktik.* Göteborg 2013

344. MARIA MAGNUSSON *Skylta med kunskap. En studie av hur barn urskiljer grafiska symboler i hem och förskola.* Göteborg 2013

345. ANNA-LENA LILLIESTAM *Aktör och struktur i historieundervisning. Om utveckling av elevers historiska resonerande.* Göteborg 2013

346. KRISTOFFER LARSSON *Kritiskt tänkande i grundskolans samhällskunskap. En fenomenografisk studie om manifesterat kritiskt tänkande i samhällskunskap hos elever i årskurs 9.* Göteborg 2013

347. INGA WERNERSSON och INGEMAR GERRBO (red) *Differentieringens janusansikte. En antologi från Institutionen för pedagogik och specialpedagogik vid Göteborgs universitet.* Göteborg 2013

348. LILL LANGELOTZ *Vad gör en skicklig lärare? En studie om kollegial handledning som utvecklingspraktik.* Göteborg 2014

349. STEINGERDUR OLAFSDOTTIR *Television and food in the lives of young children.* Göteborg 2014

350. ANNA-CARIN RAMSTEN *Kunskaper som byggde folkhemmet. En fallstudie av förutsättningar för lärande vid teknikskiften inom processindustrin.* Göteborg 2014

351. ANNA-CARIN BREDMAR *Lärares arbetsglädje. Betydelsen av emotionell närvaro i det pedagogiska arbetet.* Göteborg 2014

352. ZAHRA BAYATI *"den Andre" i lärarutbildningen. En studie om den rasifierade svenska studentens villkor i globaliseringens tid.* Göteborg 2014

353 ANDERS EKLÖF *Project work, independence and critical thinking.* Göteborg 2014

354 EVA WENNÅS BRANTE *Möte med multimodalt material. Vilken roll spelar dyslexi för uppfattandet av text och bild?* Göteborg 2014

355 MAGNUS FERRY *Idrottsprofilerad utbildning – i spåren av en avreglerad skola.* Göteborg 2014

356 CECILIA THORSEN *Dimensionality and Predictive validity of school grades: The relative influence of cognitive and socialbehavioral aspects.* Göteborg 2014

357 ANN-MARIE ERIKSSON *Formulating knowledge. Engaging with issues of sustainable development through academic writing in engineering education.* Göteborg 2014

358 PÄR RYLANDER *Tränares makt över spelare i lagidrotter: Sett ur French och Ravens maktbasteori.* Göteborg 2014

359 PERNILLA ANDERSSON VARGA *Skrivundervisning i gymnasieskolan. Svenskämnets roll i den sociala reproduktionen.* Göteborg 2014

360 GUNNAR HYLTEGREN *Vaghet och vanmakt - 20 år med kunskapskrav i den svenska skolan.* Göteborg 2014

361 MARIE HEDBERG *Idrotten sätter agendan. En studie av Riksidrottsgymnasietränares handlande utifrån sitt dubbla uppdrag.* Göteborg 2014

362 KARI-ANNE JØRGENSEN *What is going on out there? - What does it mean for children's experiences when the kindergarten is moving their everyday activities into the nature - landscapes and its places?* Göteborg 2014

363 ELISABET ÖHRN och ANN-SOFIE HOLM (red) *Att lyckas i skolan. Om skolprestationer och kön i olika undervisningspraktiker.* Göteborg 2014

364 ILONA RINNE *Pedagogisk takt i betygssamtal. En fenomenologisk hermeneutisk studie av gymnasielärares och elevers förståelse av betyg.* Göteborg 2014

365 MIRANDA ROCKSÉN *Reasoning in a Science Classroom.* Göteborg 2015

366 ANN-CHARLOTTE BIVALL *Helpdesking: Knowing and learning in IT support practices.* Göteborg 2015

367 BIRGITTA BERNE *Naturvetenskap möter etik. En klassrumsstudie av elevers diskussioner om samhällsfrågor relaterade till bioteknik.* Göteborg 2015

368 AIRI BIGSTEN *Fostran i förskolan.* Göteborg 2015

369 MARITA CRONQVIST *Yrkesetik i lärarutbildning - en balanskonst.* Göteborg 2015

370 MARITA LUNDSTRÖM *Förskolebarns strävanden att kommunicera matematik.* Göteborg 2015

371 KRISTINA LANÅ *Makt, kön och diskurser. En etnografisk studie om elevers aktörsskap och positioneringar i undervisningen.* Göteborg 2015

372 MONICA NYVALLER *Pedagogisk utveckling genom kollegial granskning: Fallet Lärande Besök utifrån aktör-nätverksteori.* Göteborg 2015

373 GLENN ØVREVIK KJERLAND *Å lære å undervise i kroppsøving. Design for utvikling av teoribasert undervisning og kritisk refleksjon i kroppsøvingslærerutdanningen.* Göteborg 2015

374 CATARINA ECONOMOU *”I svenska två vågar jag prata mer och så”. En didaktisk studie om skolämnet svenska som andraspråk.* Göteborg 2015

375 ANDREAS OTTEMO *Kön, kropp, begär och teknik: Passion och instrumentalitet på två tekniska högskoleprogram.* Göteborg 2015

376 SHRUTI TANEJA JOHANSSON *Autism-in-context. An investigation of schooling of children with a diagnosis of autism in urban India.* Göteborg 2015

377 JAANA NEHEZ *Rektorers praktiker i möte med utvecklingsarbete. Möjligheter och hinder för planerad förändring.* Göteborg 2015

378 OSA LUNDBERG *Mind the Gap – Ethnography about cultural reproduction of difference and disadvantage in urban education.* Göteborg 2015

379 KARIN LAGER *I spänningsfältet mellan kontroll och utveckling. En policystudie av systematiskt kvalitetsarbete i kommunen, förskolan och fritidshemmet.* Göteborg 2015

380 MIKAELA ÅBERG *Doing Project Work. The Interactional Organization of Tasks, Resources, and Instructions.* Göteborg 2015

381 ANN-LOUISE LJUNGBLAD *Takt och hållning - en relationell studie om det oberäkneliga i matematik-undervisningen.* Göteborg 2016

382 LINN HÅMAN *Extrem jakt på hälsa. En explorativ studie om ortorexia nervosa.* Göteborg 2016

383 EVA OLSSON *On the impact of extramural English and CLIL on productive vocabulary.* Göteborg 2016

384 JENNIE SIVENBRING *I den betraktades ögon. Ungdomar om bedömning i skolan.* Göteborg 2016

385 PERNILLA LAGERLÖF *Musical play. Children interacting with and around music technology.* Göteborg 2016

386 SUSANNE MECKBACH *Mästarcoacherna. Att bli, vara och utvecklas som tränare inom svensk elitfotboll.* Göteborg 2016

387 LISBETH GYLLANDER TORKILDSEN *Bedömning som gemensam angelägenhet – enkelt i retoriken, svårare i praktiken. Elevers och lärares förståelse och erfarenheter.* Göteborg 2016

388 cancelled

389 PERNILLA HEDSTRÖM *Hälsocoach i skolan. En utvärderande fallstudie av en hälsofrämjande intervention.* Göteborg 2016

Editors: Åke Ingerman, Pia Williams and
Elisabet Öhrn

390 JONNA LARSSON *När fysik blir lärområde i förskolan.* Göteborg 2016

391 EVA M JOHANSSON *Det motsägelsefulla bedömningsuppdraget. En etnografisk studie om bedömning i förskolekontext.* Göteborg 2016

392 MADELEINE LÖWING *Diamant – diagnoser i matematik. Ett kartläggningsmaterial baserat på didaktisk ämnesanalys.* Göteborg 2016

393 JAN BLOMGREN *Den svårfångade motivationen: elever i en digitaliserad lärmiljö.* Göteborg 2016

394 DAVID CARLSSON *Vad är religionslärar-kunskap? En diskursanalys av trepartssamtal i lärarutbildningen.* Göteborg 2017

395 EMMA EDSTRAND *Learning to reason in environmental education: Digital tools, access points to knowledge and science literacy.* Göteborg 2017

396 KATHARINA DAHLBÄCK *Svenskämnets estetiska dimensioner - - i klassrum, kursplaner och lärares uppfattningar.* Göteborg 2017

397 K GABRIELLA THORELL *Framåt marsch! – Ridlärarrollen från dåtid till samtid med perspektiv på framtid.* Göteborg 2017

398 RIMMA NYMAN *Interest and Engagement: Perspectives on Mathematics in the Classroom.* Göteborg 2017

399 ANNIKA HELLMAN *Visuella möjlighetsrum. Gymnasieelevers subjektsskapande i bild och medieundervisning.* Göteborg 2017

400 OLA STRANDLER *Performativa lärarpraktiker.* Göteborg 2017

401 AIMEE HALEY *Geographical Mobility of the Tertiary Educated – Perspectives from Education and Social Space.* Göteborg 2017

402 MALIN SVENSSON *Hoppet om en framtidsplats. Asylsökande barn i den svenska skolan.* Göteborg 2017

403 CATARINA ANDISHMAND *Fritidshem eller servicehem? En etnografisk studie av fritidshem i tre socioekonomiskt skilda områden.* Göteborg 2017

404 MONICA VIKNER STAFBERG *Om lärarblivande. En livsvärldsfenomenologisk studie av bildningsgångar in i läraryrket.* Göteborg 2017

405 ANGELICA SIMONSSON *Sexualitet i klassrummet. Språkundervisning, elevsubjektivitet och heteronormativitet.* Göteborg 2017

406 ELIAS JOHANNESSON *The Dynamic Development of Cognitive and Socioemotional Traits and Their Effects on School Grades and Risk of Unemployment.* Göteborg 2017

407 EVA BORGFELDT *"Det kan vara svårt att förklara på rader". Perspektiv på analys och bedömning av multimodal textproduktion i årskurs 3.* Göteborg 2017

408 GÉRALDINE FAUVILLE *Digital technologies as support for learning about the marine environment. Steps toward ocean literacy.* Göteborg 2018

409 CHARLOTT SELLBERG *Training to become a master mariner in a simulator-based environment: The instructors' contributions to professional learning.* Göteborg 2018

410 TUULA MAUNULA *Students' and Teachers' Jointly Constituted Learning Opportunities. The Case of Linear Equations.* Göteborg 2018

411 EMMALEE GISSLEVIK *Education for Sustainable Food Consumption in Home and Consumer Studies.* Göteborg 2018

412 FREDRIK ZIMMERMAN *Det tillåtande och det begränsande. En studie om pojkars syn på studier och ungdomars normer kring maskulinitet.* Göteborg 2018

413 CHRISTER MATTSSON *Extremisten i klassrummet. Perspektiv på skolans förväntade ansvar att förhindra framtida terrorism.* Göteborg 2018

414 HELENA WALLSTRÖM *Gymnasielärares mentorshandlingar. En verksamhetsteoretisk studie om lärararbete i förändring.* Göteborg 2018

415 LENA ECKERHOLM *Lärarperspektiv på läsförståelse. En intervjustudie om undervisning i årskurs 4-6.* Göteborg 2018

416 CHRISTOPHER HOLMBERG *Food, body weight, and health among adolescents in the digital age: An explorative study from a health promotion perspective.* Göteborg 2018

417 MAGNUS KARLSSON *Moraliskt arbete i förskolan. Regler och moralisk ordning i barn-barn och vuxen-barn interaktion.* Göteborg 2018

418 ANDREAS FRÖBERG *Physical Activity among Adolescents in a Swedish Multicultural Area. An Empowerment-Based Health Promotion School Intervention.* Göteborg 2018

419 EWA SKANTZ ÅBERG *Children´s collaborative technology-mediated story making. Instructional challenges in early childhood education.* Göteborg 2018

420 PER NORDÉN *Regnbågsungar: Familj, utbildning, fritid.* Göteborg 2018

421 JENNY RENDAHL *Vem och vad kan man lita på? Ungdomars förhållningssätt till budskap om mat och ätande utifrån ett forskarinitierat rollspel.* Göteborg 2018

422 MARTINA WYSZYNSKA JOHANSSON *Student experience of vocational becoming in upper secondary vocational education and training. Navigating by feedback.* Göteborg 2018

423 MALIN NILSEN *Barns och lärares aktiviteter med datorplattor och appar i förskolan.* Göteborg 2018

424 LINDA BORGER *Investigating and Validating Spoken Interactional Competence – Rater Perspectives on a Swedish National Test of English.* Göteborg 2018

425 ANNA-MARIA FJELLMAN *School choice, space and the geography of marketization – Analyses of educational restructuring in upper secondary education in Sweden.* Göteborg 2019

426 ANNELI BERGNELL *Med kroppen som illustration: Hur förskolebarn prat-skapar naturvetenskap med hjälp av multimodala och kroppsförankrade förklaringar.* Göteborg 2019

427 ANNE SOLLI *Handling socio-scientific controversy: Students' reasoning through digital inquiry.* Göteborg 2019

428 MARTIN GÖTHBERG *Interacting - coordinating text understanding in a student theatre production.* Göteborg 2019

429 SUSANNE STRÖMBERG JÄMSVI *Unpacking dominant discourses in higher education language policy.* Göteborg 2019

430 KURT WICKE *Läroböcker, demokrati och medborgaskap. Konstruktioner i läroböcker i samhällskunskap för gymnasiet.* Göteborg 2019

431 KATARINA SAMUELSSON *Teachers' Work in Times of Restructuring. On Contextual Influences for Collegiality and Professionality.* Göteborg 2019

432 HELÉNE BERGENTOFT *Lärande av rörelseförmåga i idrott och hälsa ur ett praktikutvecklande perspektiv.* Göteborg 2019

433 JANNA MEYER-BEINING *Assessing writers, assessing writing: a dialogical study of grade delivery in Swedish higher education.* Göteborg 2019

434 DAN FRANSSON *Game demands and fatigue profiles in elite football – an individual approach -Implications of training and recovery strategies.* Göteborg 2019

435 ELIN ARVIDSON *Physiological responses to acute physical and psychosocial stress – relation to aerobic capacity and exercise training.* Göteborg 2019

436 SUSANNE STAF *Skriva historia – literacyförväntningar och elevtexter i historieämnet på mellan- och högstadiet.* Göteborg 2019

437 VERONICA SÜLAU *Vad händer i lärares kollegiala samtalspraktik? En studie av mötet mellan en nationell kompetensutvecklingsinsats och en lokal fortbildningspraktik.* Göteborg 2019

438 MARIA OHLIN *How to Make Bicycling Safer – Identification and Prevention of Serious Injuries among Bicyclists.* Göteborg 2019

439 LINUS JONSSON *An empowerment-based school physical activity intervention with adolescents in a disadvantaged community: A transformative mixed methods investigation.* Göteborg 2019

440 ELIN NORDENSTRÖM *Feedback and instructional guidance in healthcare simulation debriefings.* Göteborg 2019

441 KATEŘINA ČERNÁ *Nurses' work practice in chronic care: knowing and learning in the context of patients' self-monitoring data.* Göteborg 2019

442 MARGARETHA HÄGGSTRÖM *Estetiska erfarenheter i naturmöten. En fenomenologisk studie av upplevelser av skog, växtlighet och undervisning.* Göteborg 2020

443 PANAGIOTA NASIOPOULOU *The professional preschool teacher under conditions of change – competence and intentions in pedagogical practises.* Göteborg 2020

444 ANNA TOROPOVA *Teachers meeting the challenges of the Swedish school system. Agents within boundaries.* Göteborg 2020

446 ULF RYBERG *Att urskilja grafiska aspekter av derivata – hur elevernas möjligheter påverkas av innehållets behandling i undervisningen.* Göteborg 2020

447 KASSAHUN WELDEMARIAM *Reconfiguring Environmental Sustainability in Early Childhood Education: a Postanthropocentric Approach.* Göteborg 2020

448 ANNE KJELLSDOTTER *Didactical Considerations in the Digitalized Classroom.* Göteborg 2020

449 CARINA PETERSON *Val, omröstning, styrning. En etnografisk studie om intentioner med, villkor för och utfall av barns inflytande i förskolan.* Göteborg 2020

450 LOTTA WEDMAN *The concept* concept *in mathematics education: A concept analysis.* Göteborg 2020

451 MARLENE SJÖBERG *Samtal om undervisning i naturvetenskap. Ämnesdidaktisk kollegial utveckling i lärarutbildning och lärarprofession.* Göteborg 2020

452 LENNART SVENSSON *Kontextuell analys – En forskningsmetodologi och forskningsansats.* Göteborg 2020

453 JOHN DOHLSTEN *Vad möjliggör och begränsar en hållbar elitfriidrott? Aktionsforskning i elitidrottspraktiker inom Göteborgs friidrottsförbund.* Göteborg 2020

454 LENA SOTEVIK *Barbiebröllop och homohundar. Barn och barndomar i relation till queerhet och (hetero)normativa livslinjer.* Göteborg 2020.

455 FRIDA SIEKKINEN *Att vara och inte vara. Elevpositioner(ingar) i spänningsfältet mellan svenska och svenska som andraspråk.* Göteborg 2021.

456 ANN-CHARLOTT WANK *Meningsskapande samtal. En studie om barns meningsskapande med fokus på processer och innehåll relaterat till förskolans praktik.* Göteborg 2021.

457 ANDREAS LUNDBERG ZACHRISSON *Overuse injuries in Swedish elite athletics. Incidence, occurrence, athlete availability, and risk factors.* Göteborg 2021.

458 ANNA NORRSTRÖM *Samtal under lärarlagsmöten. Diskursorienteringar i den professionella praktiken.* Göteborg 2021.

459 JOHANNA MELLÉN *Stability and Change. Policy, options, and choice in Swedish upper secondary education.* Göteborg 2021.

460 JONATAN JUNGMALM *Running-related injuries among recreational runners. How many, who, and why?* Göteborg 2021.

461 ELISABETH OHLSSON *Den synliggjorda vokabulären och praktiken. Gymnasieelevers akademiska skrivande på svenska.* Göteborg 2021.

462 VICTORIA ROLFE *Exploring socioeconomic inequality in educational opportunity and outcomes in Sweden and beyond.* Göteborg 2021.

463 JONAS LINDBÄCK *Värsta bästa skolan. Om unga i förorten och segregationen i skolan.* Göteborg 2021.

Editors: Christel Larsson, Elisabeth Öhrn, Pia Williams

464 ANNE-MARIE CEDERQVIST *Seeing the parts, understanding the whole. A technology education perspective on teaching and learning in processes of analysing and designing programmed technological solutions.* Göteborg 2021.

465 MARIE GRICE *Epistemic beliefs and conceptions of competence in education for sustainable development.* Göteborg 2021.

466 KRISTINA HUNEHÄLL BERNDTSSON *Digitala sexuella trakasserier i skolan: Elevperspektiv på sexting, utsatthet och jämställdhet.* Göteborg 2022.

Editors: Christel Larsson, Elisabeth Öhrn, Pia Williams och Olof Franck

467 STINA JERDBORG *Learning Principalship: Becoming a Principal in a Swedish Context. A study of Principals in Education and Practice.* Göteborg 2022.

468 ALEXANDRA SÖDERMAN *Digital studentkultur – om slutna grupper på Facebook som icke-formell arena i högre utbildning.* Göteborg 2022.

469. OLA HENRICSSON *"Som att hålla tiden i sin hand" – Didaktiskt perspektiv på muntligt berättande.* Göteborg 2022.

470 ERIKA MAJOROS *Linking recent and older IEA studies on mathematics and science.* Göteborg 2022.

471 JENNY SVANTESON WESTER *Teaching and learning mathematics with integrated small-group discussions. A learning study about scaling geometric figures.* Göteborg 2022.

472 JASMINE BYLUND *Everyday Language Practices and the Interplay of Ideologies, Investment and Identities. Language Use and Dispositions among Young Adolescents in Multilingual Urban Settings in Sweden.* Göteborg 2022.

473 AGNETA PIHL *Children retelling stories. Responding, reshaping, and remembering in early childhood education and care.* Göteborg 2022.

474 KATARINA NILFYR *Interaktionsmönster, social anpassning och emotioner i förskolan – En mikrosociologisk studie av interaktion mellan förskollärare och barn i målorienterade aktiviteter.* Göteborg 2022

475 LEAH NATASHA GLASSOW *Teacher sorting and the opportunity gap. A cross-national investigation of institutional differentiation and educational equity.* Göteborg 2022

476 ÅSA ANDERSSON *Sustainable inclusion without sustainability. Working with equal participation and unforeseen movement in physical education, sports, and research.* Göteborg 2023

477 INGELA FINNDAHL *Young students' Language Choice in Swedish compulsory school – expectations, learning and assessment.* Göteborg 2023

478 MIKAEL R KARLSSON *Skolförändring, reformer och professionella villkor – en etnografisk studie.* Göteborg 2023

479 LINUS BYLUND *Differentiation, didactics and inequality. How rich and poor populations are educated for sustainability.* Göteborg 2023

480 JONATAN FRIDOLFSSON *Statistical advancements in analyzing accelerometer-measured physical activity intensity* Göteborg 2023

481 OLA FLENNEGÅRD *Uppdrag: Historia och demokrati Perspektiv på studieresor till Förintelsens minnesplatser.* Göteborg 2023

482 MALIN BRÄNNSTRÖM *Mellan osynlighet och avvikelse – nyanlända elever med kort skolbakgrund i grundskolans senare årskurser.* Göteborg 2023

483 EMELIE STAVHOLM *Teacher professional learning in response to contemporary challenges in early childhood education and care.* Göteborg 2024

484 TANYA K OSBORNE*: Unicorns in Moderation Gender and Epistemology on Stack Overflow.* Göteborg 2024

485 DIMITRIOS PAPADOPOULOS: *Individualising processes in adult education: The case of Swedish for immigrants (SFI).* Göteborg 2024

486 ANNA-LENA BORG: *Trygga och otrygga platser En etnografisk studie om våld och utsatthet bland barn i fritidshem.* Göteborg 2024

487 PETER JOHANNESSON: *Lärares lärande i gränslandet mellan skolans och vetenskapens praktiker. Aktionsforskning som socialt lärande.* Göteborg 2024

Editors: Christel Larsson, Pia Williams and Olof Franck

488 LEA ELDSTÅHL-AHRENS: *Learning to argue in primary school: A sociocultural study of group discussions. with argumentative tasks.* Göteborg 2024

489 YI DING: *The Self in the School Context: Mathematics Self-concept and Self-efficacy in PISA.* Göteborg 2024

490 CLARA PALM: *Skrivundervisning i svenska som andraspråk inom vuxenutbildning.* Göteborg 2024

491 PANTEA RINNEMAA: *Reading and learning from civics textbooks: exploring challenges and opportunities from students' and teachers' perspective.* Göteborg 2024

492 ALENA SEREDKO: *Doing knowledge@scale: Sociomaterial Practices and Professional Learning of Software Developers on Stack Overflow.* Göteborg 2024

493 SOLVEIG E.S. HAUSKEN-SUTTER: *Interdisciplinary research and youth sport injury. Developing methodological insights.* Göteborg 2024494

494 MARINA KARLSSON: *The Quality Dialogue. An activity theoretical study on systematic quality work in a municipal preschool administration.* Göteborg 2024

495 ANNA L.V. LUNDBERG: *Att lära om statisk och dynamisk proportionalitet – En studie av den didaktiska transpositionen av svenska matematikuppgifter med proportionalitet.* Göteborg 2024

496 HADIL ELSAYED*: Health Promotion in Swedish schools: Navigating Institutional, Social and Professional Landscapes.* Göteborg 2024

497 NATALIE DAVET*: Generationsmötets ambivalenser. Kritiska perspektiv på ålder, tid och rum.* Göteborg 2024

498 SOFIJE SHENGJERGJI: *Play-Responsive Teaching. Navigating Semiotic Repertoires and Digital Technologies in Early Childhood Education and Care.* Göteborg 2025

499 JULIA WANSELIUS: *Estimating added and free sugars intake in Swedish adolescents - methods, food sources, nutritional implications, and potential food label impact.* Göteborg 2025

500 MARI LINDSTRÖM*: Teachers' Professional Competence and Working Conditions in Swedish Schools. Relationships with student achievement.* Göteborg 2025

Editors: Stefan Johansson, Olof Franck, Christel Larsson and Pia Williams

501 MARI WOLLMAR: *Opportunities and challenges with the shift to climate-adapted food consumption.Balancing nutrition, climate impact, and acceptance in public and private meals.* Göteborg 2025

502 LENA ASP: *Does teaching quality matter for student learning outcomes? A student perspective of a mathematics classroom.* Göteborg 2025

GOTHENBURG STUDIES IN EDUCATIONAL SCIENCES 502

High-quality teaching is assumed to provide students with learning opportunities that may mitigate educational inequities and narrow achievement gaps. However, empirical findings are mixed, while some studies report significant, positive relations between teaching quality and student learning outcomes, others do not. Measuring the multidimensional construct of teaching quality presents conceptual and methodological challenges. In the thesis, teaching quality is operationalised through aggregated student perceptions of mathematics teachers' practices and instruction, yielding a valid and reliable measure.

This thesis comprises three empirical studies using secondary data from the Trends in International Mathematics and Science Study (TIMSS) 2019 for Grade 4. Study I examines the construct validity of the mixed-worded mathematics confidence scale and the linguistic equivalence across translated questionnaire items. Study II investigates the relationships between aspects of teaching quality and the two outcomes of mathematics confidence and mathematics achievement in Sweden, both at the student and classroom levels. Study III extends the analysis to include a cross-national comparison of four Nordic countries, investigating classroom composition effects while accounting for student background factors.

The findings showed that classroom management related positively to mathematics achievement, while instructional clarity was significantly related to mathematics confidence. There are indications that teaching quality may mitigate the negative influence of low SES on academic achievement. Classroom-level composition effects were observed across the Nordic countries, with SES and home language as key factors for classroom-level mathematics achievement.

**Lena Asp** holds an M.SC in Education with specialisation in Learning Communication and Information Technology. She has worked as a teacher and school leader. Her research interests include teaching quality, educational measurement, and equity. She is involved in teacher education at the Department of Education and Special Education, University of Gothenburg.

ISBN 978-91-7963-219-9

9789179632199

ACTA UNIVERSITATIS GOTHOBURGENSIS