Tove Stjern Frønes
Andreas Pettersen
Jelena Radišic
Nils Buchholtz *Editors*

Equity, Equality and Diversity in the Nordic Model of Education

Springer

Equity, Equality and Diversity in the Nordic Model of Education

Tove Stjern Frønes • Andreas Pettersen
Jelena Radišić • Nils Buchholtz
Editors

Equity, Equality and Diversity in the Nordic Model of Education

Springer

Editors
Tove Stjern Frønes
Department of Teacher Education and
School Research
University of Oslo
Oslo, Norway

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

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

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

Preface

Due to increasing diversity within the school landscape, educational equity and equality have received increased emphasis in educational science discourse. This diversity not only touches on important aspects of equal opportunities in education, but also responds to the demand for ethically more responsible research to improve and reduce disadvantage and discrimination.

The Nordic countries are generally perceived as countries with the most pronounced equality of educational opportunity, which has been repeatedly confirmed by national and international large-scale studies. This equality is generally explained by the common orientation of these countries within a common Nordic model of education, which was developed in these countries after the Second World War and realises an egalitarian view of society within the framework of national education policy. In recent years, however, this common model has been called into question due to the increasing influences of globalisation and growing migration movements. Nordic countries are taking new paths in educational policy that, politically and socially motivated, prefer a more performance-oriented and economically efficient educational system and that, when implemented through educational policy reforms, call into question the image of an egalitarian society. Surprisingly, however, there has been little quantitative research on educational justice and equality of opportunity that focuses specifically on the Nordic countries and could provide a stocktake of the current situation.

The members of the Large-scale Educational Assessment (LEA) research group at the University of Oslo therefore wanted to use research data from the accessible international large-scale studies (e.g. the Trends in Mathematics and Science Study [TIMSS] and the Programme for International Student Assessment [PISA]) and national tests (e.g. mapping tests in Norway) in a constructive and exploratory manner to compile corresponding findings about the various Nordic countries. The LEA group combines scientific expertise from different areas (e.g. mathematics, reading and ICT) and has been working for years in the field of analysis of student outcome data and investigation of educational systems at different levels. The editors are all

part of the working group, and since 2018, they have brought together scientists from different Nordic (and other) countries to realise this book as an international endeavour.

Oslo, Norway	Tove Stjern Frønes
Oslo, Norway	Andreas Pettersen
Oslo, Norway	Jelena Radišić
Oslo, Norway	Nils Buchholtz

Acknowledgement

Contents

Chapter 1
Equity, Equality and Diversity in the Nordic Model of Education—Contributions from Large-Scale Studies

Tove Stjern Frønes, Andreas Pettersen, Jelena Radišić, and Nils Buchholtz

1.1 Introduction

In education, the 'Nordic model' refers to the similarities and shared aims of the education systems developed in the five Nordic countries—Denmark, Finland, Iceland, Sweden and Norway—after World War II. Traditionally, there have always been many similarities and links between the Nordic countries through their historical connections and geographical proximity. The common experience of solidarity and political oppression during World War II also created the basis for a common political orientation in the postwar period, which was also reflected in the education systems during the development of the countries' economies and their establishment of welfare states. At the same time, this process has been strongly supported by social-democratic governance in these countries in the 1960s and 1970s (Blossing, Imsen, & Moos, 2014). The model is based on a concept of *Education for All*, where equity, equal opportunities and inclusion are consistently cited as the goal of schooling and orientation (Blossing et al., 2014; Telhaug, Mediås, & Aasen, 2006). This corresponds to the egalitarian idea of a classless society, which is characterised by individual democratic participation, solidarity and mutual respect and appreciation for all. This idea was manifested in, for example, major reallocations of economic resources through the tax systems and free schooling for all, which arose out of the principle that parents' lack of economic resources should not prevent children from obtaining a good quality education. The equalisation of structural inequalities and creation of equity was—and still is—the task of the education system in the Nordic countries. Worldwide, especially within the Nordic countries, the view is being shared that the education system should be fair and provide access and opportunities for further education, regardless of where someone lives, the status of the parental home, where someone comes from, what ethnic background someone has, what age or gender someone is, what skills one has or whether someone has physical disabilities (Blossing et al., 2014; Quaiser-Pohl, 2013). Some

T. S. Frønes (✉) · A. Pettersen · J. Radišić · N. Buchholtz
Department of Teacher Education and School Research, University of Oslo, Oslo, Norway
e-mail: t.s.frones@ils.uio.no

special features of the Nordic system are therefore deeply embedded in the school culture in the countries, for example, through the fact that access to free and public local schools and adapted education is statutory, which is in contrast to many other countries, even other European ones (further developed and discussed in Chap. 2). The Nordic model is widely considered a good example of educational systems that provide equal learning opportunities for all students. Achieving equity, here meaning the creation of fairness, is expressed concretely in political measures to distribute resources equally and strengthen the equality of marginalised groups by removing the barriers to seize educational opportunities, for example, when mixed-ability comprehensive schools are created or the educational system is made inclusive regarding students with special needs (UNESCO, 1994; Wiborg, 2009). Equality is roughly connoted with 'sameness in treatment' (Espinoza, 2007), while equity takes further in consideration also the question of how well the requirements of individual needs are met. Thus, the goal of equity is always linked to the concept of justice, provided that an equality of opportunities is created. If, however, one looks at individual educational policy decisions on the creation of educational justice in isolation, one must weigh which concept of equity or equality is present in each case. For example, it is not enough to formally grant equal rights in the education system to disadvantaged groups, but something must also be done actively to ensure that marginalised groups can use and realise this equality. The complexity of the terms becomes even greater when one considers that to achieve equality, measures can be taken that presuppose an unequal distribution of resources or unequal treatment and, therefore, are not fair e.g., when resources are bundled especially for disadvantaged groups and these are given preferential treatment (will be further developed and discussed in Chap. 2). Thus, equality and equity rely on each other and are in a field of tension comprising multiple ideas (Espinoza, 2007).

1.2   Challenges Put to the Nordic Model

Because of migration movement beginning in the late 1970s, economic growth and differentiated welfare distribution, social inequality has increased, especially in the last decade. Therefore, teachers in the Nordic educational systems are faced with increasing student diversity. Beyond gender and the students' physical or mental abilities, this diversity very much includes heterogeneity in students' social, cultural and economic background, hence not automatically warranting support and equal learning opportunities for all. More students today than when the common Nordic model was developed have the difficulty of following lessons in the national language, one that is not their mother tongue or even easily spoken at home.

  A much more complex diversity has also emerged involving several factors, such as multi-cultural or transnational affiliation, access to (digital) educational resources and mobility in a globalized world, establishing a more complex group structure than in previous decades. Conversely, in policy, general support for the equalising idea behind the Nordic model is decreasing, with claims that globalisation has

forced the Nordic countries to compete on an international scale and that in the recent neoliberal educational policy, 'the concept of a *School for All* is no longer part of the rhetoric' (Blossing et al., 2014, p. 2). As Lundahl (2016) describes, there have always been differences in how the individual Nordic countries shape their education policies. This is partly because of different traditions, the rural character of the countries and different public management mechanisms. Under today's changing conditions, however, differences can now be identified (such as the opening of the educational sector for private schools in Sweden, hegemonic measures for dealing with cultural diversity in Finland and Iceland or the introduction of soft streaming models and ability grouping in Denmark) that call into question the guidelines of the Nordic model. Nevertheless, education policy measures are usually justified by the strengthening of equity. How can this be understood? With these measures, however, it is somewhat unclear how inequalities within the national systems are dealt with and whether educational policy addresses issues of educational inequality or inequity (Espinoza, 2007). Some even question whether a unified approach in the Nordic countries truly exists anymore (Antikainen, 2006; Lundahl, 2016).

1.3   The Outline of This Volume

Although previous analyses of how the Nordic model is enacted in practice take into account the ideological and economic aspects of the educational policy to reduce inequality and strengthen equity in the school systems, their findings have only been backed up occasionally with empirical evidence. Of course, the question of a common Nordic model and how the different countries achieve these aims can be traced in a historical review of the model's origins and development and an analysis of syllabi, curricula guidelines and policy documents in education from the individual Nordic countries. However, here, international comparative studies with a large number of participating students and/or teachers can make a distinctive contribution allowing entire education systems to be observed and compared.

Large-scale studies make use of standardised measuring instruments that meet high-quality assurance standards. In quantitative analyses, the underlying structures in education systems can be traced from a comparative perspective, and references can be made to the similarities and differences in the individual Nordic countries. Empirical evidence from both national and international large-scale assessment studies on the relationship between socio-economic status (SES), different cultural background, different learning opportunities and student achievement has become increasingly important for policy makers when making decisions about educational means to close the achievement gap between different student groups and reduce educational inequality. Based on 20 years of data from the Trends in International Mathematics and Science Study (TIMSS), a recent volume by Broer, Bai and Fonseca (2019) reports, for example, on the changes in the relationship between SES and student achievement and educational inequality. However, cross-country

comparisons on equity and equality and studies that take into account different stakeholders in the Nordic educational system are scarce (e.g., OECD, 2019; Rühle, 2015; Volante, Klinger, & Bilgili, 2019; Volckmar, 2019).

In this volume, we acknowledge and underline the importance of considering the context of education when investigating equity, equality and diversity. We attempt to provide a better understanding of both the functions and the foundations of the Nordic model in education through our theoretical and methodological discussions and our examinations of studies conducted in the Nordic countries. The book consists partly of chapters discussing conceptual, philosophical and methodological issues and partly of chapters presenting key findings from secondary analyses of data from studies of educational outcomes. In the theoretical and methodological chapters, we give systematic presentations of how the results of various large-scale national and international assessment studies can be used as indicators of equity, equality and diversity. The empirical part of the book provides relevant empirical analyses of the different factors related to equity, equality and diversity by considering the impact of factors operating at different levels. There are contributions both related to the school or class levels of equity and on equity at the student level, here inspecting groups of students systematically.

The data pulled together in this book stem from various large-scale assessment studies and are analysed by authors from different countries across the Nordic area. Thereby, we have aimed for a carefully crafted collection of chapters using international and national large-scale assessment studies, including TIMSS, Programme for International Student Assessment (PISA), Progress in International Reading Literacy Study (PIRLS), Teaching and Learning International Study (TALIS), International Computer and Information Literacy Study (ICILS) and national tests (e.g., Norwegian mathematical literacy mapping test). Usually, only internal comparisons within one of these studies are conducted. In this volume, we attempt to contrast the findings from comparable studies to bring in different perspectives. In the different chapters of the book, the investigations address various subject domains (i.e., mathematics, science, reading), different age cohorts and various grades. However, each investigation addresses the aspects pertinent to the topics of equity, equality and diversity across the education systems in the Nordic countries.

Although the theme of this book is the Nordic model, it could however not be realised that all chapters analyse data from all Nordic countries. In most chapters, it is not the model per se that is examined, but the chapters do consist of studies of how features of the model appear at different times in some of these countries and how equity unfolds in the region. In the same way, it is natural that a book based on a Norwegian research group uses Norway as a case. In some large-scale assessments or cycles, data are also available for only some countries (e.g., PIRLS, ICILS, TALIS). Several of the chapters have also been limited to countries that have common features, either educational or cultural. It is nevertheless the case that all considered countries are Nordic.

The book is grounded in the collaboration of a large group of contributing authors of the LEA (large-scale educational assessment) research group from the Faculty of Educational Sciences at the University of Oslo in Norway. The common strength of

the international comparative studies that are conducted by this group lies in the fact that they all deal with the topics of equity and equality on different levels and assess national profiles as a way to inform policy makers. The idea that the group's research results should be made available to other researchers and cooperation with scientists from other universities in the Nordic region, but also from other European countries, was promoted. The studies were also conducted by researchers situated in both the different theoretic and didactic fields—science, mathematics, reading and digital competency—but also in the assessments themselves: the researchers behind this book are 'insiders' in these large-scale assessments, and this contributes to situated, rich analyses of data from international large-scale assessments. Many of the authors are closely involved in the reporting of large-scale studies within their respective countries as well, which is why the creation of this book is all the more valuable for us. The authors have dared each other to be curious, open and transparent by exploring both the basic concepts and methods in our traditional line of research. For this reason, however, it is not to be expected that a fundamental collective and interdisciplinary reappraisal of the topic of equity and equality can take place in the short time of about 2 years it took to write this book. In this respect, the book does not represent basic research or framework development in the field of equity and equality theory, nor is it able to present the results of research programmes, some of which have been running for several years, such as Broer et al.'s (2019) research. Nevertheless, in structuring the book, we have not only taken empirical aspects into account, but also the overarching philosophical-theoretical considerations and a systematic synopsis. Furthermore, we asked a colleague from qualitative instructional research to take a critical stance to the book. With this structure, the findings of the book will be relevant and interesting for researchers, policy makers and practitioners.

1.4   Content and Structure of the Book

Overall, the book comprises four principal sections. The first section contains two chapters on the theoretical-philosophical and methodological considerations of equity, equality and diversity in the Nordic model of education. Chapter 2 starts with the philosophical contribution of Buchholtz, Stuart and Frønes (2020), who discuss the concepts of equity, equality and diversity and their relevance to the idea of equality in the Nordic model of education. The three concepts are interrelated and set as critical keystones in the international comparative debate on educational justice. With the notion that the discourse of the concepts and educational policies based on them reflect the cultural traditions and orientations of the Nordic countries and that evidence on achieved equity has to be interpreted with caution when looking at the findings from large-scale studies, the chapter concretises the ideas in describing and discussing different educational policies of the Nordic countries and challenges the scientific research on equity, equality and diversity.

The section ends with a chapter by Mittal, Nilsen and Björnsson (2020) that addresses diverse methodological and analytical approaches in connection to equity and equality. In particular, they focus on the comparability of the equity measures used and the manner in which countries' level of equity is viewed regarding the different standards and analytical approaches employed. Taking a methodological stance, the authors contribute to the overall discussion on the impact that diverse approaches may have and the implications these hold for educational research as the discussions on equity and educational policy in the Nordic countries evolve.

The second part of the book presents a collection of studies related to the teacher and the school variables in connection to understanding equity, equality and diversity in the Nordic countries. Björnsson (2020) starts by investigating teachers' attitudes and experiences of teaching in a multicultural setting (Chap. 4). Using the TALIS data, the author focuses on the variations in self-efficacy in multicultural classrooms across the Nordic countries, and whether this variation can be explained by different aspects of teacher background.

Continuing with the TALIS data (Chap. 5), Yang Hansen, Radišić, Liu and Glassow (2020) focus on the diversity in the relationship between different aspects of teacher quality and job satisfaction across the Nordic countries. The authors comparatively examine these mechanisms by taking into account both the system characteristics and ongoing changes in each, discussing how these are enacted in an everyday school environment that serves students of different backgrounds and educational needs.

Rohatgi, Bundsgaard and Hatlevik (2020) continue with a comparative perspective (Chap. 6) focusing on Norwegian and Danish schools, here on the topic of digital inclusion and how collaboration between teachers, their professional development, attitude and ICT use affect students' ICT literacy. Taking data from the ICILS study, the authors examine the variation in computer and information literacy in the two countries, where these policies are warranted at the national level.

Nilsen, Scherer, Gustafsson, Teig and Kaarstein (2020) contribute further with their investigation of teachers' role in enhancing equity with the aid of TIMSS data, focusing primarily on the different aspects of teacher qualifications and how these possibly moderate the relationship between students' outcomes and their social background taking into account teachers' instructional quality.

In Chap. 8, using PISA data, Scherer (2020) critically assesses reported evidence for positive and significant SES–achievement relations and the substantial variation of this relation, both in strength and proposed underlying mechanisms, across educational contexts, such as classrooms, schools and educational systems. Using a Nordic lens, he tests three hypotheses, that is, the *compensation, mediation* and the *moderation hypothesis* on the interplay between students' SES, the disciplinary climate and achievement in science.

Finally, in Chap. 9, Nortvedt, Bratting, Kovpanets, Pettersen and Rohatgi (2020) report on how a national-level assessment initiative can contribute to equity in school by improving the opportunities to learn for students identified as at risk for lagging behind in mathematics. Using student data from implementations of the Norwegian mapping test, in addition to data from teacher interviews, the author

team addresses both what happens to an assessment that is exposed over time and how it serves the purpose of supporting teachers and their ongoing practice. The chapter also addresses what happens to the students identified through the mapping tests.

Section three of the book focuses on the empirical studies related to the student-level variables in the context of equity, equality and diversity, exploring the learning opportunities of different student groups. Bergem, Nilsen, Mittal and Ræder (2020) investigate the importance of teachers' instructional quality for student motivation in the view of their diverse socio-economic backgrounds (Chap. 10). Using TIMSS data for Norway, the authors seek to identify how different dimensions of instructional quality are related to motivation for students with different socio-economic background in both grade 5 and 9.

Exploring diverse student profiles is the focus of Chap. 11, which is authored by Radišić and Pettersen (2020). The authors start by investigating the motivational profiles of resilient and non-resilient student groups in Sweden and Norway by using TIMSS data with a person-centred approach. Furthermore, the authors investigate the characteristics of the classroom and school environment pertinent to the identified profiles and compare the results of the two countries.

Frønes, Rasmusson and Bremholm (2020) contribute to the discussion by investigating equity and diversity in reading comprehension through the lenses of the PISA reading assessment (Chap. 12). This chapter studies the reading performance of diverse student groups in the period from 2000 to 2018, including comparisons between Norway, Sweden and Denmark and reading policy development in the countries. The authors in particular, address the introduction of new text formats as multiple and dynamic, which is made possible through the change of assessment delivery mode.

The importance of the delivery mode in reading tests is the focus of Engdal Jensen (2020) in Chap. 13. The author explores the Norwegian case as she examines to what extent delivery mode influences student's outcomes. The chapter strongly focuses on the gender perspective, raising the question whether the change in delivery mode affects boys' and girls' results on reading comprehension tests in the same way.

With the aid of PIRLS data, in Chap. 14, Støle, Wagner and Schwippert (2020) focus on the learning environment at home and investigate whether the children of parents who read in their spare time and have positive attitudes towards reading activities do better on reading assessments, even if these parents have a low level of education. The analyses compare the five Nordic countries, discussing implications in the context of immediate school surroundings and the compensating effects schools may provide for particularly vulnerable groups of students.

The fourth and final section of the book has two parts. It comprises both a critical overview of the book provided in the commentary by the Finnish educational researcher Fritjof Sahlström and a concluding chapter where the editors of the volume provide a brief summary of the book's findings and respond to the commentary.

In his commentary, Sahlström (2020) has been invited to comment on this book for several reasons. First, we wanted the perspectives on this volume from his

qualitative point of view—because over the course of his research career, he has been inspecting the inner workings of the educational system through in-depth case studies, not data from large-scale assessments. Because the goal of this volume is intentionally and deliberately restricted to large-scale international assessment resources, we wanted an outsider perspective to comment on the findings about the Nordic model and the benefits and limitations of the methods and data used.

In the concluding chapter of the book, 'Equity, Equality and Diversity in the Nordic Countries—Final Thoughts and Looking Ahead', by Frønes, Pettersen, Radišić and Buchholtz (2020), we synthesize the findings and possible implications from the empirical chapters and also comment on the benefits and limitations and indicate areas for future research, showing where our findings can be a point of departure for diverse methodological approaches.

References

Antikainen, A. (2006). In search of the Nordic model in education. *Scandinavian Journal of Educational Research, 50*(3), 229–243.

Bergem, O. K., Nilsen, T., Mittal, O., & Ræder, H. G. (2020). Can teachers' instruction increase low-SES students' motivation to learn mathematics? In T. S. Frønes, A. Pettersen, J. Radišić, & N. Buchholtz (Eds.), *Equity, equality and diversity in the Nordic model of education* (pp. 251–272). Cham, Switzerland: Springer.

Björnsson, J. K. (2020). Teaching culturally diverse student groups in the Nordic countries—What can the TALIS 2018 data tell us? In T. S. Frønes, A. Pettersen, J. Radišić, & N. Buchholtz (Eds.), *Equity, equality and diversity in the Nordic model of education* (pp. 75–97). Cham, Switzerland: Springer.

Blossing, U., Imsen, G., & Moos, L. (Eds.). (2014). *The Nordic education model. A 'school for all' encounters neo-liberal policy*. Dordrecht, The Netherlands: Springer.

Broer, M., Bai, Y., & Fonseca, F. (2019). *Socioeconomic inequality and educational outcomes. Evidence from twenty years of TIMSS*. Cham, Switzerland: Springer.

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* (pp. 13–41). Cham, Switzerland: Springer.

Engdal Jensen, R. (2020). Implications of changing the delivery mode on reading tests in Norway—A gender perspective. In T. S. Frønes, A. Pettersen, J. Radišić, & N. Buchholtz (Eds.), *Equity, equality and diversity in the Nordic model of education* (pp. 337–362). Cham, Switzerland: Springer.

Espinoza, O. (2007). Solving the equity—Equality conceptual dilemma: A new model for analysis of the educational process. *Educational Research, 49*(4), 343–363.

Frønes, T. S., Pettersen, A., Radišić, J., & Buchholtz, N. (2020). Equity, equality and diversity in the Nordic countries—Final thoughts and looking ahead. In T. S. Frønes, A. Pettersen, J. Radišić, & N. Buchholtz (Eds.), *Equity, equality and diversity in the Nordic model of education* (pp. 397–412). Cham, Switzerland: Springer.

Frønes, T. S., Rasmusson, M., & Bremholm, J. (2020). Equity and diversity in reading comprehension—A case study of PISA 2000–2018. In T. S. Frønes, A. Pettersen, J. Radišić, & N. Buchholtz (Eds.), *Equity, equality and diversity in the Nordic model of education* (pp. 305–335). Cham, Switzerland: Springer.

Lundahl, L. (2016). Equality, inclusion and marketization of Nordic education: Introductory notes. *Research in Comparative & International Education, 11*(1), 3–12.

Mittal, O., Nilsen, T., & Björnsson, J. K. (2020). Measuring equity across the Nordic education systems—Conceptual and methodological choices as implications for educational policies. In T. S. Frønes, A. Pettersen, J. Radišić, & N. Buchholtz (Eds.), *Equity, equality and diversity in the Nordic model of education* (pp. 43–71). Cham, Switzerland: Springer.

Nilsen, T., Scherer, R., Gustafsson, J.-E., Teig, N., & Kaarstein, H. (2020). Teachers' role in enhancing equity—A multilevel structural equation modeling with mediated moderation. In T. S. Frønes, A. Pettersen, J. Radišić, & N. Buchholtz (Eds.), *Equity, equality and diversity in the Nordic model of education* (pp. 173–196). Cham, Switzerland: Springer.

Nortvedt, G. A., Bratting, K. B., Kovpanets, O., Pettersen, A., & Rohatgi, A. (2020). Improving equity through national-level assessment initiatives. In T. S. Frønes, A. Pettersen, J. Radišić, & N. Buchholtz (Eds.), *Equity, equality and diversity in the Nordic model of education* (pp. 225–248). Cham, Switzerland: Springer.

OECD. (2019). *PISA 2018 results (Volume II): Where all students can succeed.* Paris: OECD Publishing. https://doi.org/10.1787/b5fd1b8f-en

Quaiser-Pohl, C. (2013). Diversity in education and the concept of 'diversity' as a topic for educational science studies. In C. Quaiser-Pohl, V. Ruthsatz, & M. Endepohls-Ulpe (Eds.), *Diversity and diversity management in education—A European perspective* (pp. 9–22). New York: Waxmann.

Radišić, J., & Pettersen, A. (2020). Resilient and nonresilient students in Sweden and Norway—Investigating the interplay between their self-beliefs and the school environment. In T. S. Frønes, A. Pettersen, J. Radišić, & N. Buchholtz (Eds.), *Equity, equality and diversity in the Nordic model of education* (pp. 273–304). Cham, Switzerland: Springer.

Rohatgi, A., Bundsgaard, J., & Hatlevik, O. E. (2020). Digital inclusion in Norwegian and Danish schools—Analysing variation in teachers' collaboration, attitudes, ICT use and students' ICT literacy. In T. S. Frønes, A. Pettersen, J. Radišić, & N. Buchholtz (Eds.), *Equity, equality and diversity in the Nordic model of education* (pp. 139–172). Cham, Switzerland: Springer.

Rühle, S. (2015). *Diversität, curriculum und bildungsstrukturen. Eine vergleichende Untersuchung in Deutschland und Finland* [Diversity, curriculum and educational structures. A comparative study in Germany and Finland]. Münster, Germany: Waxmann.

Sahlström, F. (2020). The black box of Nordic education held against the light of large-scale international assessment resources—A critical commentary. In T. S. Frønes, A. Pettersen, J. Radišić, & N. Buchholtz (Eds.), *Equity, equality and diversity in the Nordic model of education*)pp. 387–396. Cham, Switzerland: Springer.

Scherer, R. (2020). The case for good discipline? Evidence on the interplay between disciplinary climate, socioeconomic status, and science achievement from PISA 2015. In T. S. Frønes, A. Pettersen, J. Radišić, & N. Buchholtz (Eds.), *Equity, equality and diversity in the Nordic model of education* (pp. 197–224). Cham, Switzerland: Springer.

Støle, H., Wagner, A. K. H., & Schwippert, K. (2020). The importance of parents' own reading for 10-year old students' reading achievement in the Nordic countries. In T. S. Frønes, A. Pettersen, J. Radišić, & N. Buchholtz (Eds.), *Equity, equality and diversity in the Nordic model of education* (pp. 363–384). Cham, Switzerland: Springer.

Telhaug, A. O., Mediås, O. A., & Aasen, P. (2006). The Nordic model of education: Education as part of the political system in the last 50 years. *Scandinavian Journal of Educational Research, 50*(3), 245–283.

UNESCO. (1994). *The Salamanca statement and framework for action on special needs education.* UNESCO. http://www.unesco.de/fileadmin/medien/Dokumente/Bildung/Salamanca_Declaration.pdf

Volante, L., Klinger, D., & Bilgili, O. (2019). *Immigrant student achievement and education policy. Cross-cultural approaches.* Dordrecht, The Netherlands: Springer.

Volckmar, N. (2019). The enduring quest for equity in education: Comparing Norway and Australia. *Scandinavian Journal of Educational Research, 63*(4), 617–631.

Wiborg, S. (2009). *Education and social integration. The development of comprehensive schooling in Europe*. New York: Palgrave Macmillan.

Yang Hansen, K., Radišić, J., Liu, X., & Glassow, L. N. (2020). Exploring diversity in the relationships between teacher quality and job satisfaction in the Nordic countries—Insights from TALIS 2013 and 2018. In T. S. Frønes, A. Pettersen, J. Radišić, & N. Buchholtz (Eds.), *Equity, equality and diversity in the Nordic model of education* (pp. 99–137). Cham, Switzerland: Springer.

Part I
Theoretical and Methodological Considerations

Chapter 2
Equity, Equality and Diversity—Putting Educational Justice in the Nordic Model to a Test

Nils Buchholtz (iD)**, Amelie Stuart, and Tove Stjern Frønes**

Abstract Equity, equality and diversity are often linked to educational policy within the Nordic countries in the form of goals and principles. This can be traced back to the common educational tradition of these countries within the Nordic model of education. Because the terms are often used interchangeably, it seems appropriate to first grasp the theoretical and philosophical understanding of the terms before concrete educational policy measures can be assessed regarding to these goals. The chapter provides an overview of the terms and concretises educational policy measures to achieve equity, equality and diversity in the context of the Nordic countries. Today, societal developments and political changes call into question the common ground of the Nordic countries when it comes to matters of educational equity. Among other things, it will be discussed what contribution large scale international comparative studies can make to understanding equity, equality and diversity.

Keywords Equity · Equality · Diversity · School for all · Nordic model of education · Educational justice · ILSA

The Nordic model of education and its idea of a "School for All" is recognised by various parties as a realisation of greater equity (Blossing, Imsen, & Moos, 2014; Telhaug, Mediås, & Aasen, 2006). But equity can refer to various aspects of the educational policy discourse, and it is not always synonymous with equality,

N. Buchholtz (✉) · T. S. Frønes
Department of Teacher Education and School Research, University of Oslo, Oslo, Norway
e-mail: N.f.Buchholtz@ils.uio.no

A. Stuart
Max Weber Centre for Advanced Cultural and Social Studies, University of Erfurt, Erfurt, Germany

13

especially when it comes to the question of the realisation of life chances and educational justice. Furthermore, the classroom today is no longer as homogeneous as it was when the model was developed in light of experiences of common solidarity among the Nordic countries. So, what do equity and equality mean in the "School for All" today? Increased diversity among students suggests that the "All" has changed in the "School for All", and that the idea of equity is today confronted with changed and more differentiated individual needs. Thus, if the Nordic model is to maintain its idea of a "School for All", then the "School" must also change. When equity is used interchangeably with equality and thus means equal treatment by educational administrations, increased inequality in terms of needs is difficult to address, as some special needs might be overlooked or insufficiently resolved. In Norway, for example, education policy planning over the last decade has shifted from an understanding of "equity through equality", and thus standardisation and uniformity, to a new policy of "equity through diversity" with less dependency on central authorities (Solstad, 1997). Policy document analyses by Haugen (2010), however, reveal that how equity is understood and can be achieved is not a given and a matter of educational policies that are based on certain ideological groundings. This requires a theoretical and philosophical reflection on the concepts of equity, equality and diversity in education and how they are interpreted and implemented in educational policies in the Nordic countries. In this chapter, we will move from central philosophical theories on equity, equality and diversity in the international debate to how they mirror central features of the Nordic context.

A country's educational system plays a key role when addressing questions of fairness and equality. It lies at the centre of important normative questions concerning, for instance, equal opportunities for all members of society or respecting individual diversity. The prospect of equality of opportunities in education is a hope shared but also doubted by almost all education systems in the world. As political scientist Iris Young (2011, p. 21) notes with regard to the US:

> While there are vast disagreements about why, almost no one in American society today thinks that educational opportunity is equal. There are vast and growing disparities in the quality of education to which Americans have access, and these shamefully track race and class […]. The turn-of-the-twentieth-century hope that public education can equalize the relationship among children of very unequal parents, giving each child an equal chance to compete with others from more privileged backgrounds, seems like a strange dream.

When policy makers intend to counterbalance or eradicate social and economic inequalities in order to achieve social justice, justice considerations with respect to the educational system must be addressed. This is because inequality is "manifested in the family environment, in occupational status and level of income; [and] it is also evident in educational opportunities, aspirations, attainment and cognitive skills" (Espinoza, 2007, p. 344). The main reason for looking at the impact of the educational system in this respect is, as Espinoza notes, that "educational systems […] are involved in the reproduction and change of class relationships" (2007, p. 344). Thus, putting the question to an empirical study would mean first analysing which aspects and procedures within the educational system (implicitly) maintain

or reproduce inequalities and then – starting from this analysis – working to develop policy measures to address the inequality and to change procedures.

In the first part of the chapter, we will focus on the theoretical underpinnings of the concepts of equity, equality and diversity in education, whereby we will reflect on the international origins of the respective discourses. As Espinoza points out, it is important to clarify the terminology when discussing matters of just and unjust equality (cf. Espinoza, 2007, p. 344). Drawing from his work on the concepts of "equality" and "equity", we will briefly introduce and discuss the main differences between these concepts and their relevance for our discussion of justice and equality of educational opportunity. For the discussion of educational justice, the concept of diversity also plays a role, which we will present in a complementary manner and place in the context of equity and equality. Vertovec (2007, 2010) describes diversity in contemporary societies as "super-diversity", pointing to a new and emerging complexity. According to Vertovec, this super-diversity involves several factors and comprises groups based not only on religion, country of birth and language, but also on the social rights and status achieved by different immigrant groups who have arrived at varying times and with different social statuses (2007, 2010). As we will see in the second part of the chapter, in the Nordic countries, this super-diversity is today matched with the legal counterpart of "inclusion". As an example, we will examine different approaches to establishing educational equity in the Nordic countries with regard to dealing with minority language students in order to concretize the previously elaborated concepts of equity, equality and diversity in different educational systems. The question of how united the Nordic countries are today with regard to a common Nordic model of education will be addressed in our final discussion.

2.1 Equity and Equality in Educational Contexts

In general, the concept of "equity" means being equal in quantity and quality and can be associated with justice in the sense of fairness, according to Espinoza (2007, pp. 344, 346). In this sense, individual circumstances and differences related to individual needs and requirements in the educational context are taken into consideration. The concept of "equality", on the other hand, can be associated with the idea of sameness in treatment, which is based on the normative ideal of the equality of all persons. It might seem that the implementation of "equality" would lead to more justice and equal opportunities within the educational context. However, we will show in which ways the concept of equity can provide further, important criteria for enhancing justice. We will therefore start our investigation of equality and equity with what is often seen as the most basic of the two concepts: equality.

2.1.1 Equality

Von der Pfordten (2010) summarises the consequences of equality as a normative ideal as follows: "Every individual which has to be considered ethically must also be considered equally concerning her interests" (p. 201). According to von der Pfordten, there are essentially two options for assessing and evaluating the interests of those parties to be considered equally: equal treatment (e.g. concerning the allocation of resources or in taxation) and equality in society (e.g. when relating individuals to each other or to the society as a whole). However, apart from this formal consideration of equality, the material aspect of having the same opportunities to realise these claims is significant. As the economist and philosopher Amartya Sen points out, under conditions of extreme poverty, whether a child owns, for example, a bicycle could be the decisive factor in getting an education, because having a bicycle would ensure that the child can go to school even if the school is far away (cf. Sen, 1983).

In relation to this, the philosopher G.A. Cohen has emphasised that the focus on possessing rights is not sufficient. It is at least as important to ask whether a person actually has the opportunity to exercise these rights. Cohen therefore distinguishes between a "lack of freedom" and "unfreedom" (cf. Cohen, 1983). An example of this would be the right to freedom of movement, which is enjoyed by all citizens and which stands in contrast to the impossibility of exercising this right due to economic or other constraints. Here, the responsibility of the state is extended from merely guaranteeing citizens' rights to also working against the causes of this "unfreedom" of citizens. Another example to illustrate the concept, coming from education, would be the idea of national school curricula for equal educational attainment, no matter where students live or what kind of school they attend with respect to their abilities. The legal philosopher Martha Nussbaum similarly points out that guaranteeing rights is not enough for an autonomous, fulfilled life. What is crucial is a set of capabilities that every person must develop in order to be able to live out his or her rights (cf. Nussbaum, 2003, p. 37). Over the years, Nussbaum has compiled and supplemented a list of these capabilities. Among other things, this list is intended as a normative test for state action and governmental duties, since it is primarily the responsibility of a government to ensure conditions under which all citizens can develop their respective abilities. If these conditions are not properly in place or only partly fulfilled, then citizens will be unable to develop all their capabilities and will thus be prevented from leading happy and autonomous lives.

2.1.2 Equity

While the concept of equality in educational contexts can be discussed along the dimensions of access to education, educational provision and organisation, survival, output and outcome (cf. Antikainen, 2006; Espinoza, 2007; Farell, 1999), the

concept of equity "demands fair competition but tolerates and, indeed, can require unequal results" (Espinoza, 2007, p. 346). This means that equality can be assessed quantitively by, for example, asking how many people in any given society have how much access to highly demanded goods. Equity, however, is assessed both quantitively and qualitatively, which means that it includes a moral judgment of a certain distribution of opportunities or goods. In this respect, it is much more difficult to assess equity because of, for example, subjective differences in how the quality and extent of inequalities are assessed (see Chap. 3). Despite this ambiguity, Espinoza attempts to provide an orientation: "The fundamental idea underlying the 'equity' theory is that fairness in social relationships occurs when rewards, punishments and resources are allocated in proportion to one's input or contributions" (2007, p. 348). Here, "input" refers to what an individual contributes to his or her success or the outcome of a process – for example, ambition or talents. But it is important to define what this means: a person's contribution. This needs to be addressed before questions of justice can be considered. We might ask: What counts as a person's natural gift and what does not? What (kind of) influence do virtues and vices, such as diligence or ambition, have on our assessment of these factors? For these issues, it is helpful to compare one individual's contribution in relation to the benefits he or she enjoys to the contributions and benefits of other individuals (Espinoza, 2007, p. 349), since this comparison might enable us to evaluate the fairness of the outcome (e.g. a person's allocated resources).

This outcome becomes manifest in the socio-economic background of a person, which is often described and measured as socio-economic status (SES). We could also describe socio-economic background as a person's position within a social structure. This position can be either advantageous and associated with benefits and opportunities or disadvantageous and associated with obstacles and discriminations. The assumption here is that each position within the social structure lends to each person specific possibilities and limitations. One proponent of this model is Iris Young. She assumes that societies can be depicted on the basis of a structural model. The way in which a social structure is built has very profound consequences for the opportunities and liberties of each member of society. Young borrows this structural model from sociology, referring to, for example, Pierre Bourdieu's field theory: "He conceives structures as 'fields' on which individuals stand in varying positions in relation to one another, offering possibilities for interpretation and action" (Young, 2006, p. 112) These so-called structures surround every human being and influence their freedom of action and their decision-making abilities. These structures also coordinate and shape collective action, and they are confirmed by the actual adherence to social norms, which are themselves based on this social structure.

If the effects of the structure within which a person finds himself or herself are disadvantageous, then according to Young, one could call this "structural injustice". Young defines this kind of injustice as follows:

> The wrong is structural injustice, which is distinct from at least two other forms of harm or wrong, namely, that which comes about through individual interaction, and that which is attributable to the specific actions and policies of states or other powerful institutions. (2011, p. 45)

This means that in the case of structural injustice, a person's limitations or disadvantages are caused neither by his or her own actions and decisions nor by governmental decisions or laws that would limit personal freedoms. Rather, structural injustice is influenced by the relative position a person has within the hierarchical social structure, a position which is based on access to, or control over, wealth, prestige and power (Mueller & Parcel, 1981) and the possibilities and limitations resulting from this position. Initially, a person has no control over this position, as it is ascribed at birth. And yet, one's opportunities in life depend on this original position within the social structure. Correspondingly, the socio-economic status of a person is used in many studies on educational equity and equality as an important indicator for determining if a person belongs to a marginalised group. Willms and Tramonte (2019), for example, propose for educational research to investigate equity by examining differences among sub-populations with different SES in terms of their access to key measures of educational provision, such as quality instruction, taking up the perspective of distributive justice. On the other hand, equality can be studied if differences in student outcomes can be attributed to differences in SES, thus raising the perspective of equality. SES can be operationalised in studies in different ways (each with different explanatory power). Chapter 3 is therefore devoted to the different ways of measuring this construct statistically and how different statistical methods can be used to estimate equity (see Chap. 3).

The concept of equity, as introduced above, can be illustrated in this respect when we want to compare the status of one group or individual with the status of another group or individual. One example for this could be by looking at their access to higher education and by asking whether each member has the same access to higher education despite their different positions within the structure. It would therefore be insufficient to look only at the formal establishment of equal access.

The importance of this for justice considerations is addressed by John Rawls in the following way, describing the ideal of equality of opportunity:

> The thought here is that positions are to be not only open in a formal sense, but that all should have a fair chance to attain them. Offhand it is not clear what is meant, but we might say that those with similar abilities and skills should have similar life chances. More specifically, assuming that there is a distribution of natural assets, those who are at the same level of talent and ability, and have the same willingness to use them, should have the same prospects of success regardless of their initial place in the social system. In all sectors of society there should be roughly equal prospects of culture and achievement for everyone similarly motivated and endowed. The expectations of those with the same abilities and aspirations should not be affected by their social class. (1999, p. 63)

This quote draws attention to one of the core aspects of educational justice: life chances. The right to education, which is guaranteed (for example in the Universal Declaration of Human Rights) to each child out of justice concerns, can be viewed as an instrument for enhancing a person's life chances. But there is a crucial difference between having the formal right to education and equal treatment and actually being able to make use of this right by, for example, being able to regularly attend school, as we have seen earlier.

Thus, when looking for an answer to the initial normative question about overcoming inequalities, according to Rawls, a society would need the ideal of equality of opportunity for at least two reasons:

(1) The "outcomes of natural chance or the contingency of social circumstances" (Rawls, 1999, p. 11) need to be counterbalanced if they have a negative influence on the realisation of life plans. Equality of opportunity as an ideal of political decision making is important in cases when members of a community have an unequal share of benefits and burdens because they were born into unequal social positions, i.e. for contingent reasons.
(2) Equality of opportunity is needed to compensate for possible negative consequences and disadvantages if they are the result of certain individual decisions, e.g. starting a family.

Especially, this second reason shows that even though Rawls never uses the term "equity", he clearly has this concept in mind when he discusses the ideal of equality of opportunity, since he is very much concerned with individual circumstances and their relation to fairness.

In the Nordic countries, the idea of "education for all" has been particularly strong, as a basis for the Nordic welfare model, but also in the legal sense. Not only do students have the formal right to education – the right to access schooling – but all students, even those with special needs, have statutory rights to attend their local school and receive compulsory schooling up to 16 years of age (Imsen & Volckmar, 2014). More remarkable in an international context is students' statutory right to adapted education and student-centred learning for equalisation and inclusion – which can thus be linked to the concept of equity. The purpose clauses of the Nordic education systems are explicitly linked to equalisation – introduced to increase mobility in society and reduce differences among various groups, primarily social disparities (Imsen & Volckmar, 2014, p. 46). Telhaug et al. (2006) show that this links back to the main goals of compulsory schooling in Nordic societies after the Second World War, namely to establish social virtues such as equal opportunity, cooperation, adaptation and solidarity.

2.1.3   The Tension Between Equity and Equality

What becomes clear from this consideration of equity and equality is the fact that these concepts are not necessarily complementary to one another – instead, there is a tension between them. Enhancing equity might not necessarily also enhance equality between the members of a society. As we have seen earlier, the concept of equity focuses on a fair distribution of highly demanded goods, whereas the concept of equality "is associated with the democratic ideal of social justice [and] demands equality of results" (Espinoza, 2007, p. 346). In other words: While the concept of equity focuses on distributive justice, the concept of equality mainly looks at procedural justice (e.g. when people are treated equally; Espinoza, 2007, p. 349; for a

detailed discussion on distributive justice, see Lamont & Favor, 2017). Consequently, the problem might arise that "if we wish to produce equal results, it is likely that we will need to generate an unequal distribution of resources" (Espinoza, 2007, p. 348). In other words: A policy aiming at greater equity within the educational system might entail a reduction of equality at the same time (cf. Espinoza, 2007, p. 346). Blossing et al. (2014, p. 7) give a concrete example for the possible dilemma in the education sector:

Should more resources be allocated to the most able pupils in order to maximise the national economic benefit of the school system, or is it more appropriate to channel more resources to those that are in need of the most help and support? If the distribution of resources is equal for all pupils, the result will probably be increasing social differences in educational outcomes, so this is an odd issue in the question of equity.

It is important to recall here that the concept of equality means, broadly speaking, sameness in treatment. This sameness does not necessarily also have to be just. For example, when we consider a group of different individuals with different needs or abilities, an equal treatment of all group members might not as a matter of fact be also regarded as "just", since some of them have certain needs that are not shared by all and, as a result, these members might need a treatment different from that of the rest of the group. In relation to this, we need to emphasise one aspect: Difference in treatment (and thus inequality) can be regarded as just only if it does not harm the other group members or puts them in a disadvantageous position compared to the position of the others (Rawls, 1999).

This example of a justified difference in treatment of certain members of a group is yet very much in accordance with the concept of equity, since it takes individual circumstances into account. In the context of education, we can see that in the concept of equity there lies a concern that students are different along several dimensions that have an impact on their need for learning and follow-up in the educational system. Opheim (2004) describes the need for fair learning environments, taking into account that most students are not alike. If all were alike, equity in education would simply be a question of providing an equal distribution of educational resources to all students – and thus it would turn out to be the concept of equality. But because students are different both individually and in the type and amount of resources they have obtained from their family and environment and which they bring with them into the classroom, their individual need for training will vary (Opheim, 2004), and therefore the concept of equity is needed.

2.1.4   Diversity in Educational Contexts

In the context of education, the tension between equity and equality must also be expanded by current educational policy challenges. Educational research is primarily concerned with the emergence and effects of inequality and selection in relation to educational pathways and in the social and economic sense. So far, we have

outlined that, in relation to equity, distributive justice plays a role in the allocation of highly demanded goods (which sometimes leads to the acceptance of unequal distributions), and that the concept of equality refers to procedural aspects of justice in, for example, the dismantling of barriers to access to higher education. However, there is something missing. As Blossing et al. (2014, p. 7) point out: "Both equity and equality are terms that seem to be connected to the adjective equal, which is defined as being the same in quality, size, degree or value. These definitions miss the notion of being different, but of equal worth".

When looking at education, due to forms of migration, transnationalisation and hybridisation, other aspects of individuality and equality in educational contexts also come into focus (Robak, Sievers, & Hauenschild, 2013). In addition to age, gender or socio-economic differences as a cause of structural inequality, there are other factors that play a role in the attribution of life chances, such as different cultural backgrounds, national-ethno-cultural (multiple) affiliations, cultural values, religions, languages, physical conditions and individual abilities (Robak et al., 2013, p. 15). Such forms of social diversification have increasingly been subsumed under the term "diversity" over the past 10 years (see Nestvogel, 2008; Prengel, 2013; Robak et al., 2013). While the concepts of equity and equality seem to be conceptually differentiated and refer to philosophical and sociological theories, the use of the term "diversity" has so far referred not so much to a unified concept as it has to a discourse that is concerned with the question of the appropriate political, legal, economic and educational handling of social diversity as influenced by particular theories (Hofmann, 2012; Robak et al., 2013).

With regard to one of the origins of the discourse, we will draw attention to the well-known context of diversity in the US. The political debate on diversity began in the US as early as the 1960s, when the so-called Grassroots Movement, the civil rights movement and the women's movement fought for equality at the workplace and in society (see Quaiser-Pohl, 2013). A central concern was the abolition of racial segregation in public schools. Even though African Americans were formally allowed to attend the same type of school as white students at that time, due to inequalities in housing and patterns of racial segregation in neighbourhoods, there were racially segregated "Black schools" that were worse equipped and harder to reach than schools for white people. Consequently, equality in education was one of the main themes of the movement, while the focus was placed on the categories of gender, race and class.

The attempt to assimilate the Sámi people in Norway can be seen as a complementary example from the Nordic countries (Gaski, 2008). Assimilation was an official Norwegian policy up until the Second World War, one which sought to compel Sámi people to discard their indigenous identity in favour of an ethno-national Norwegian identity and state citizenship (Gaski, 2008, p. 220). Gaski points to the results of this intensive policy of assimilation from the seventeenth century onwards, highlighting the radical decline in people who identified themselves as Sámi, extensive impoverishment, political powerlessness and a lack of knowledge about Sámi history and culture. A turning point for the political organisation of Sámi interests

came in the 1950s, when a revitalization movement focused on the Sámi identity led to a new Sámi self-image (Jakobsen, 2011).

Even today, the concept of diversity in education is inextricably linked with the concept of equality (cf. Quaiser-Pohl, 2013; Volckmar, 2019). In addition to the education sector, in the economic sciences in the early 2000s, a so-called diversity management branch developed in reaction to increasing globalisation and internationalisation. In contemporary human resource management, the diversity of employees is used constructively as a resource to increase the efficiency and competitiveness of enterprises (cf. Robak et al., 2013). Concerning the US, contemporary researchers refer to the so-called "Big 8" as central categories for addressing diversity: race/ethnicity, gender, nationality, class/socio-economic status, age, sexual orientation, mental/physical ability and religion (Plummer, 2003; Quaiser-Pohl, 2013). This categorisation is also commonly used, albeit in varying forms, throughout the international research society.

In the educational sciences, Annedore Prengel brought together the political discourse, which stems from the anti-discrimination debate, and the utilitarian discourse on dealing with diversity in organisations. In her theoretical description of a resource-oriented diversity education, diversity is roughly understood as synonymous with difference and heterogeneity (Prengel, 2007, 2013). Using different categories of differentiation (such as ethnicity, gender and disability), Prengel describes how marginalised groups are discriminated against and socially marginalised, as well as how these groups fight for recognition as different and for overcoming difference. Thus, the term "diversity" comprises two levels: an analytical level and a normative level. Diversity is directed against discrimination based on attribution. The constitutional processes of the lines of difference vary in each case and must be reconstructed empirically in order to make them understandable and amenable to analysis (see Robak et al., 2013). In normative terms, this approach provides a power-critical analysis of exclusion based on attributions and, with the appreciation of the individuality of each person, also includes a bridge to the topic of inclusion (Prengel, 2013). Contemporary international large-scale assessment studies, such as PISA, address the issue of social diversity and operationalise the concept by looking separately at the performance of marginalised groups of students (e.g. with immigrant status or in terms of resilience) or by examining the variation in student performance both between and within schools (OECD, 2018). While more extensive aspects of heterogeneity and diversity are addressed here when measuring equity (cf. Chap. 3), students are still considered to belong to a certain risk group.

Although the concept of heterogeneity, according to its Greek etymology, describes the non-uniformity of the elements of a set and thus does not prescribe a hierarchy, the concepts of heterogeneity or difference can contain negative connotations, due to their duality, because they can be understood as a disturbance or deviation from assumed or expected homogeneity (Nestvogel, 2008, p. 21). The concept of difference is used by Prengel, however, to emphasise the uniqueness of individuals on the basis of different social criteria of difference (Robak et al., 2013). Diversity has a positive connotation because the term includes an appreciative attitude and openness to the differences of people. Moreover, according to UNESCO

(2009), cultural diversity, also referred to as sociodiversity, is a concept that is regarded as a resource for innovation:

> In a globalizing world, such changes are pervasive and make for the increased complexity of individual and group identities. Indeed, the recognition — and even affirmation — of multiple identities is a characteristic feature of our time. One of the paradoxical effects of globalization is thus to provoke forms of diversification conducive to innovation of all kinds and at all levels. (UNESCO, 2009, p. 28)

The term "diversity" has therefore increasingly replaced the concept of heterogeneity in educational debates. In the discourse on diversity, the terms "equal opportunities", "equal justice" and "educational justice" are frequently used, but their meanings are not always clear. The term "inequality", for example, is gradually being replaced by diversity, which leads to differentiation practices being increasingly discussed separate from political questions of distribution and justice, a trend which has been criticised (Hofmann, 2012, p. 30). In its normative orientation, the concept of diversity can be linked to Rawls' theory of justice. According to him, justice is to be understood as fairness, which refers to the freedom of the individual and equal opportunities based on performance:

> (a) Each person has the same indefeasible claim to a fully adequate scheme of equal basic liberties, which scheme is compatible with the same scheme of liberties for all; and (b) Social and economic inequalities are to satisfy two conditions: first, they are to be attached to offices and positions open to all under conditions of fair equality of opportunity; and second, they are to be to the greatest benefit of the least-advantaged members of society (the difference principle). (Rawls, 2001, p. 42)

If all members of society are to be given the opportunity to freely choose and pursue their goals in life, then disadvantages in education must be eliminated by a compensatory redistribution of resources. However, overcoming these disadvantages, which are caused by contingencies such as a child's birthplace or cultural background, should not only address the segregation of groups in terms of performance characteristics, which can be countered, for example, by the provision of special educational measures, as Stojanov (2011) notes:

> In negative terms, this means that the central manifestations of injustice in education are emotional neglect, disregard for subjectivity, and ignoring and disregarding the potential abilities of individuals. In this context, the isolated focus on achievement as an alleged criterion for the "fair" distribution of life chances in and through educational institutions appears to prevent insight into the actual target norms of educational justice. (p. 24, translated by the authors).

> In concrete terms, this means that the unequal treatment of people with a migration background, for example, takes place because these people "are regarded as *determined* by their origin or as *products* of a family enculturation that is postulated to be deviant" (Stojanov, 2011, p. 42), and in the process experience a disregard. It therefore requires a shift in the educational policy discourse towards an appreciative recognition of diversity as a resource, one which can be seen and empirically analysed, for example, in the attitudes of teachers, changes in curricula, national school

policy developments and the general strengthening of the autonomy of marginalised groups in education.

2.2 Equality, Equity and Diversity in the Educational Systems of the Nordic Countries

In the Nordic countries, the influence of multicultural and diverse groups on social and educational contexts has long been discussed. The experiences of the political reorganisations after the French Revolution and the Napoleonic Wars, but especially those of the Second World War, which were characterised by strong solidarity but also by political oppression, led to a socially, broadly supported understanding of democracy in the Nordic countries with high political participation and beliefs of equality. Equitable education was seen as one of the keys to achieving the goals of the Nordic welfare model. Accordingly, education for democracy, solidarity and social commitment was the core objective of educational policy in the following decades. Whereas the rural character of the Nordic countries was originally characterised by regional differences in education, each of these countries introduced state-controlled public comprehensive schools to varying degrees in the second half of the twentieth century (Antikainen, 2006; Blossing et al., 2014; Telhaug et al., 2006). From this, the Nordic countries developed the ideal model of a "School for All", which is also discussed in educational policy discourse under the term "Nordic Model of Education" (Lundahl, 2016; Telhaug et al., 2006). The model follows an egalitarian philosophy of the education of a classless society based on solidarity, which sees the task of nation-state action in the equalisation of social differences and recognises the extension of this task to the scholastic education of future generations. Correspondingly, the comprehensive schools gradually replaced forms of schooling based on organisational differentiation or ability grouping and consisted essentially of unstreamed, mixed-ability classes. Until the 1970s and 1980s, the implementation of the model was strongly influenced in the individual countries by long-lasting social democratic governments, which were established in parallel to the economic construction of welfare states based on the general principle of equality without large income disparities. At the same time, the cultural homogeneity of the population in the individual Nordic countries remained relatively stable during this period, with the exception of the historically developed treatment of cultural minorities, such that equity in the education system tended to relate mainly to the compensation of regional differences, gender differences or skills disadvantages (Blossing et al., 2014). However, the experience with work-related immigration movements since the 1970s has posed challenges for this orientation. As Blossing et al. (2014) note:

> Since the mid 1980s new forms of governance and discourses have been introduced. Triggered by the entrance into and the competition on the global market place, all Nordic countries have brought political neoliberal thinking and governance, including new public

management systems and social technologies, into their education systems, although in different ways and with different consequences for school practice. (p. 5)

The new neo-liberal and conservative policies of the 1980s and 1990s, which emphasised competition and individualism, have been discussed as being incompatible with the traditional egalitarianism of the Nordic countries (Blossing et al., 2014; Tjeldvoll, 1998). State intervention in the school system, decentralisation measures that allowed municipalities more freedom in allocating school resources, the handling of an increasing number of students with an immigration background, and stronger selection and segregation processes in the education system have all been discussed as effects of this change in educational policy. As a consequence, under the impression of growing social inequality in some Nordic countries, the model of the "School for All" and the extent to which it still corresponds to educational policy realities is currently a controversial topic (Antikainen, 2006; Blossing et al., 2014; Lundahl, 2016).

Coming back to our initial observation, despite the general orientation of the Nordic model of education with a "School for All", we indeed find differences between the countries in terms of how they understand equity and with regard to their strategies for coping with the increasing heterogeneity of their students. Some of these differences have partly arisen historically, such as the extent to which state action should guide the education sector, or how strongly the expansion of the comprehensive school system is linked to the formation of state identity and regional politics (Antikainen, 2006). In the following, we will present important educational policy measures and historical developments in the individual Nordic countries in connection with equity, equality and diversity when dealing with marginalised groups. The focus of our overview will be on dealing with national minority language students and students with an immigrant background. This issue is a benchmark across all Nordic countries facing similar challenges, as large percentages of immigrants in these countries tend to be concentrated in socioeconomically disadvantaged neighbourhoods and are overrepresented in "disadvantaged schools" (defined as schools with the highest proportion of students whose mothers have low levels of education) (see Quaiser-Pohl, 2013, p.17). Particularly, Norway, Sweden, Finland and Denmark have developed strong integration policies following increased immigration after 2000 and in light of the fact that 2 out of 5 immigrant students to these nations are socioeconomically disadvantaged (OECD, 2019a). Of course, with respect to our considerations of diversity thus far, we are aware that this is an inadequate reduction. We unfortunately cannot go into detail about all the educational policy backgrounds and measures that could be discussed in connection with diversity, such as the status of inclusion in schools or how to deal with students with special needs (Arnesen & Lundahl, 2006; Egelund, Haug, & Persson, 2006; Lundahl, 2016), the handling of religious plurality (Skeie, 2009), the processing of regional educational differences and inequality between schools (OECD, 2019a), or responses to gender differences, including the consistent female dominance in performance and academic attainment rates in the Nordic countries (OECD, 2012, 2019a; Pekkarinen, 2012).

2.2.1   The Case for Norway

The educational system and its handling of cultural differences contributed greatly to Norway's development from a poor country to one of the richest nations in the world. The issue of dealing with linguistic minorities, such as Sámi people, in schools was raised as early as the eighteenth century in Norway (Engen, Kulbrandstad, Kulbrandstad, & Lied, 2018). Increased immigration movements by workers from Pakistan, India and Turkey in the 1970s and the admission of Vietnamese, Chilean and Iranian refugees in the 1980s led to political discussions about cultural and linguistic homogeneity within the country. These discussions were initiated by the Norwegian Sámi Association and by strikes led by the Immigrant Children's Parents Union, which drew attention to the poor performance of their children in Norwegian schools. These political disputes in the early 1980s led to educational policy reforms regarding equal treatment and to the establishment of formal equality in education through changes in school curricula in 1987 (Engen et al., 2018). These changes guaranteed functional bilingualism for minority language students, but this was changed in the 1990s due to massive political pressure from the anti-immigration and pro-assimilation movement. From then on, minority language students lost the right to be taught in their mother tongue as soon as they had sufficient knowledge of Norwegian to be able to follow regular lessons (so-called transitional bilingualism). Municipalities were more or less free to provide such education, and mother tongue education is therefore clearly marginalised in Norway, with only 2–6% of all contemporary minority language students taking part in such education, although there are regional differences (Loona & Wennerholm, 2017; Statistisk sentralbyrå, 2017). Today, immigration to Norway is comparatively moderate, although multilingual diversity is present at the classroom level in all urban areas. In early 2020, 18.2% of Norway's population were either immigrants or had immigrant parents (Statistisk sentralbyrå, 2020). Norway has a public school system that is divided into three levels: primary and lower secondary school, which is compulsory for students from 9–16 years of age, and three-year upper secondary school, including vocational schooling, which ensures the possibility of obtaining equivalent educational qualifications according to performance and ability. The comprehensive school system thus follows the social democratic and multicultural model of a "School for All", but despite the objective of levelling social inequality through education, studies have repeatedly confirmed that social differences in learning outcomes have been greater in Norway than in other, comparable countries due to the large gender gap found in the PISA assessment of reading literacy (OECD, 2019a; Opheim, 2004). Nevertheless, Norway is still among the countries with the lowest impact of socio-economic factors on student performance (OECD, 2012, 2019b), and there is no significant difference in the performance of disadvantaged students in either advantaged or disadvantaged schools (OECD, 2018). In Imsen and Volckmar's (2014) analysis of the Norwegian school system, they list a number of studies that indicate regional differences between schools due to the decentralisation policy in education in the 1990s, as well as studies that

identify performance differences between social groups. In response, recent educational policies, such as the 2006 Knowledge Promotion Reform (Kunnskapsløftet), placed emphasis on adapted education and individualisation in teaching, which compelled teachers to devote much time to individual student support (Imsen & Volckmar, 2014). Generally, globalisation and international comparison have put the Norwegian educational system continuously to the test. Over the last 20 years, therefore, the influences of neo-liberal education policies have been noticeable in the education sector, including consistent monitoring of student performance and educational outcomes through standardised achievement tests and early intervention in performance (Imsen & Volckmar, 2014). Current surveys show that the performance of immigrant students across all school types is still significantly lower than that of other students (OECD, 2018), although there are differences in their performance in reading or English, and there is a tendency for second-generation immigrant students to partly overcome their disadvantages (Statistisk sentralbyrå, 2017). Boys from immigrant families are, however, identified as a particularly disadvantaged group, as they have comparatively lower rates of completion of regular secondary schooling and are also less likely to take up university studies (Statistisk sentralbyrå, 2017).

2.2.2   The Case for Sweden

After the Second World War, equality and diversity have been explicit goals in the Swedish education system in terms of core values to be taught (e.g. Husén, 1989; Rosén & Wedin, 2018; SOU, 2014). From being one of the world's most centralised school systems, the Swedish school system has been transformed since the early 1990s into one of the world's most decentralised (Gustafsson, et al. 2014). The organisation and governance of Swedish schools changed radically when the responsibility for carrying out education was decentralised to municipalities and independent principals and a new state school administration was created. Gustafsson et al. (2014) especially emphasise how the "School for All" was challenged by the deregulated distribution of resources, freedom of choice between municipal and independent schools, free establishment in the school market with tuition fees as financial incentives, and a new grading system. The independent school reform in 1992 allowed private profit-making school providers to enter the education sector (Lundahl, 2016). These publicly funded, privately run independent schools have become a substantial part of contemporary schooling: 15.2% of compulsory school students and 27.6% of upper secondary school students attended such schools in 2018–2019. In the early 1990s, Sweden implemented the free school choice policy, allowing students to choose the school of their preference. Such a policy breaks with the former proximity principle of recruiting students with the intention of promoting equity and reducing residential segregation. However, empirical evidence has demonstrated the negative consequences of this policy on educational equity and justice, e.g. intensifying school segregation (e.g. Fjellman,

Yang Hansen, & Beach, 2018; Gustafsson et al., 2014; Söderström & Uusitalo, 2010; Yang Hansen and Gustafsson, 2016).

During the 1960s and 1970s, Sweden was the first country in Europe to adopt the idea of multiculturalism in educational policy, and the social democratic policy of Olof Palme strengthened the cultural autonomy and mother tongue education for immigrant students in Sweden (cf. Loona & Wennerholm, 2017). With increasing work-related immigration movements in the 1980s and the rise of asylum seekers in the 1990s (immigrants mostly coming from the Middle East, Latin America and former Yugoslavia), similar developments as in Norway applied to Sweden, and the society became more stratified (see e.g. Svanberg & Tydén, 1999). In 2019, the proportion of the Swedish population with a foreign background (either immigrant or immigrant parents) was 25.5% (SCB, 2020), which is currently the highest among the Nordic countries. As early as 1983, an educational policy decree stipulated that schools adopt intercultural learning methods (SOU, 1983), which is in line with the Swedish notion of a "School for All", one based upon values of equality, community and integration (Egelund et al., 2006; Rosén & Wedin, 2018, p. 58). Rosén and Wedin note this as a shift in discourse from a previous focus on multicultural education in terms of specific activities for children with migration backgrounds towards an intercultural education that includes all students. The policy changed slightly after economic crises in the 1990s. For example, state financial support for the municipalities was initially suspended, but since 2002 minority language students have again been supported in the acquisition of both languages by corresponding guidelines. However, the municipalities are relatively free to decide how and whether to provide appropriate services for minority language students (Loona & Wennerholm, 2017). The formal right to mother tongue instruction for minority language students in Swedish schools is at present marginalised, as Loona and Wennerholm note, even if the proportion of minority language students taking part in this kind of instruction is still comparatively high, at 54% (2017, p. 316). Possible reasons for this are the underfunding of the courses, a lack of teachers of Swedish as a second language, and the fact that such school courses are offered peripherally – for example, at off-peak times after school hours. The Swedish school system is a public comprehensive compulsory system and consists of both primary and lower secondary education. Few students also attend a special equivalent Sámi School for the first six years. A non-compulsory three-year strand of upper secondary education follows, attended by 99% of the age cohort (Båvner, Barklund, Hellewell, & Svensson, 2011).

School curricula are set centrally in Sweden, schools and student performance are monitored centrally by school inspections and national tests, and classes are made up of mixed-ability classrooms in accordance with the political approach (e.g. Blossing & Söderström, 2014). The analyses of Eklund (2003) on how diversity has been handled in education in Sweden since 1960 show, however, that there is a mismatch between the political aims of a "School for All", the curriculum and the views of students when, for example, looking at findings on the school-related segregation of minority language student groups. The influence of the socio-economic background of students on their performance has increased over the past two decades

(e.g. Gustafsson & Yang Hansen, 2017) and is at the same level as other OECD countries and the highest among the Nordic countries (OECD, 2019b). PISA has consistently found that immigrant students constantly perform worse than their native peers, even when controlled for socio-economic background (OECD, 2019b). In addition, a large-scale admission reform in Stockholm, which introduced freedom of choice of schools based on grades alone, led to a significant increase in segregation by family background in 2000, and especially segregation between immigrants and natives (Söderström & Uusitalo, 2010). The approach of heterogeneous classes is repeatedly undermined by homogenisation within the schools through ability grouping, which is often used as an organisational solution to deal with students' learning differences (Båvner et al., 2011; Blossing & Söderström, 2014), although according to a Swedish Skoleverket report, this has shown no effect on student performance (Skolverket, 2009). Blossing and Söderström (2014) conclude their analysis by stating that the Swedish school system with its political approach is today exposed to a neo-liberal educational policy that focuses strongly on educational output and therefore may lose sight of the establishment of equity. In Sweden in particular, this calls into question the idea of the Nordic model of education, which is also addressed in the discussion of the results of cross-national analyses in Chap. 3.

2.2.3   The Case for Iceland

Iceland also has a growing proportion of immigrants in its population, reaching 14.1% in 2019 (Statistics Iceland, 2019). Large groups of immigrants come from Poland, Lithuania, the Philippines and Thailand. However, immigration began somewhat later in Iceland than in the other Nordic countries, starting in the 1990s (Ragnarsdóttir & Lefever, 2018). Consequently, there are still few students with foreign backgrounds in the Icelandic school system, but increased immigration is expected to change this in the coming years (Garðarsdóttir & Hauksson, 2011; OECD, 2019b). Iceland has a compulsory public school system that spans preschool to higher education, with widespread enrolment in upper secondary level. In Iceland, equal access to education irrespective of gender, economic status, geographic location, religion, disability, and cultural or social background has been anchored in the Icelandic constitution since 1944. The school system changed by law to a comprehensive system with mixed-ability groups in 1974, no longer disadvantaging students in rural areas who had to take part in ambulatory schooling or students who were grouped according to their reading ability regardless of age. In their analysis of the Icelandic school system, Sigurðardóttir, Guðjonsdóttir, and Karlsdóttir (2014) describe the development of the understanding of the Icelandic concept of a "School for All", moving from creating equality for students in rural areas in the beginning to an inclusive school system at present. They also list aspects of equity achievements, like broad-based inclusion in school (less than 1% of students attend special schools), a national curriculum based on adapted teaching and the recent emphasis

on individualised learning. Since English is widely spoken in Iceland, the Icelandic Language Council together with the Icelandic Ministry of Education, Science and Culture changed the official language policy in 2008 in an effort to increase exposure to the Icelandic language in the educational sector. The "Icelandic for Everything Language Policy" emphasised that all students who have a heritage language other than Icelandic have the right to receive instruction in Icelandic as a second language, and that all schools must have reception plans in place for minority language students (Jónsdóttir, Ólafsdóttir, & Einarsdóttir, 2018; Ragnarsdóttir & Lefever, 2018). As a result of these measures, Iceland is repeatedly recognised as having attained a high level of equity in education (OECD, 2012, 2019b; Sigurðardóttir et al., 2014). The success of these efforts must, however, be viewed in the national context. Since Iceland is in the group of countries where immigrants are either highly skilled or come from high-income countries (OECD, 2019a), socio-economic background is less a factor in immigrant children's school performance, and its impact has even been decreasing in recent years (OECD, 2018). However, the gap in reading performance between immigrant students and non-immigrant students in the PISA is large (OECD, 2019a, 2019b). On the other hand, efforts to integrate immigrant students do not seem to have contributed to the envisioned equality in education, as shown by a study from 2011 on the educational success of migrants in Iceland (Garðarsdóttir & Hauksson, 2011). The study showed that only about 60% of male and 40% of female immigrants pass a secondary school examination, far less than in other European countries (cited in Ragnarsdóttir & Lefever, 2018). In general, the dropout rates from secondary school are comparatively high in Iceland (around 30%), which Sigurðardóttir et al. (2014) see as a major challenge for Icelandic educational policy.

2.2.4   The Case for Finland

In Finland, compared to other Nordic countries, the proportion of people with a migrant background is comparatively low: 7.2% in 2018 (Statistics Finland, 2020). Similar to Iceland, immigration to Finland only started in the 1990s, with most immigrants coming from the former Soviet Union, e.g. Estonia. But it is not only immigration that creates a need for multiculturalism in the Finnish school system: both the Evangelical Lutheran Church and the Orthodox Church are established by law and enjoy special privileges. In this regard, students have the right to instruction based on their own religious affiliation. A minority of the Finnish population, 5.4%, politically strengthened in their cultural autonomy already by the constitution of 1919, speaks Swedish, and 0.03% of the population speaks Sámi (Graeffe & Lestinen, 2012). Culturally, the impression of a relatively homogeneous population still exists, even though the Swedish minority attends its own schools and exists relatively parallel to the Finnish majority society (Holm & Londen, 2010). A constitutional reform of 1999 guarantees minorities equality based on the principle of a multicultural state, which in the educational sector also embraces functional

bilingualism and multiculturality for immigrant populations. Immigrant students are provided with special individual support measures to establish their schooling and learn Finnish and Swedish (Graeffe & Lestinen, 2012). The government also recommends and enables the teaching of Finnish as a second language or teaching in the mother tongue, and it is estimated that about 75% of minority language students participate in such programmes (2012). After a multi-sectional school system was successively replaced by a fundamental school reform in the 1970s, Finland implemented a nine-year, single-structured comprehensive school system. Since 2004, but even more so with the new curriculum that came into effect in 2016, the national curriculum is based on the model of Finland as a multi-ethnic state and take into account multicultural, intercultural and international education (Räsänen, 2007; Rühle, 2015). However, the excessively narrow definition of cultural diversity, the formulation of only particular educational goals for individual minority groups instead of universal goals for all students, and the failure to take other aspects of diversity into account are the object of criticism that the political orientation towards multiculturalism is intended as a "one-way process" and is related primarily to the hegemonic integration of immigrants into the majority society (Holm & Londen, 2010; Zilliacus, Holm, & Sahlström, 2017; see also Rühle, 2015). At 5.8%, Finland currently has only a small proportion of immigrant students in education, like Iceland. Since immigrants from the former Soviet Union tend to be better educated than the average population, Finland has, in recent years, been able to demonstrate how well minority language students are integrated in the education system. International large-scale studies, such as PISA, have shown that immigrant students in Finnish schools perform significantly better than immigrant students in other countries (Graeffe & Lestinen, 2012). However, the performance differences between immigrant students and their native peers are the largest among the Nordic countries, not least because of the good performance of Finnish students (OECD, 2019b). Gender differences in student performance in Finland are also the largest among the Nordic countries, preferring girls (OECD, 2019b). Ahonen (2014) suspects the consequences of deregulation of school financing as the root of these findings, as many schools have cut financial resources for remedial teaching. In his analysis of the Finnish school system, Ahonen further shows the influences of neo-liberal education policy in Finland since the 1990s. For example, the introduction of marketisation and parental choice of primary schools has led to increasing segregation and polarisation between schools with respect to socio-economic background, which is also reflected in a widening gap between schools in PISA, at least between 2000 and 2009 (Ahonen, 2014).

2.2.5   The Case for Denmark

Denmark, too, had experience with guest workers from Southern and Eastern Europe, the Middle East and Asia as early as the 1960s and 1970s, whose families are now part of the Danish population. At the beginning of 2020, 13.8% of the

Danish population had a migrant background (Statistics Denmark, 2020). However, there is a clear disadvantage in the academic performance of students with a migrant background. Immigrant students from a non-western background perform less well than their native peers in standardised tests, even when controlled for socio-economic status (Houlberg, Andersen, Bjørnholt, Krassel, & Pedersen, 2016; OECD, 2019b; Rangvid, 2010). Against this background, however, the liberal-conservative Danish education policy of the last 20 years – strongly influenced by the right wing Dansk folkeparti (The Danish People's Party) – has been pursuing a strictly hegemonic course in the sense of a Danish unified culture since the mid-2000s (Horst & Gitz-Johansen, 2010). On the one hand, this can be seen, for example, in the fact that learning Danish as a second language is only offered individually and is autonomously initiated by school principals (Andersen et al., 2012; Houlberg et al., 2016). On the other hand, political influence can be seen in the distribution of minority language students to different school districts with a higher proportion of non-immigrant students, which has been controlled since 2006 by the largely autonomous municipalities. This was done because it was estimated that schools with a student population composed of 50% immigrant students would experience a deterioration in academic performance (Calmar Andersen & Thomsen, 2011). This system has been supported by changing governments, but as the scheme is still optional, only some municipalities have chosen to implement it. Overall, a relatively constant average percentage of 10–11% immigrant students at Danish schools has been observed over the years (Houlberg et al., 2016; OECD, 2019b). Although re-distribution is seen as a measure to establish educational equality, as is made clear in Horst and Gitz-Johansen's (2010) analysis of education policy documents from 2003–2005, the strategy conveys a reading of equality in the sense of a deprivation paradigm "where the interpretation of underachievement is closely related to the child's ethnicity, family and locality, including lower socio-economic status. This is mirrored in an absence of recognition of ethnic diversity as linguistic and cultural resources" (Horst & Gitz-Johansen, 2010, p. 143). Denmark, with its Folkeskole, has a 10-year, non-streamed comprehensive public school system. Denmark also has a long tradition of students attending private schools (Lundahl, 2016), and a substantial proportion of students, 15%, attend these schools (Rasmussen & Moos, 2014). One reason for the increase in Danish students attending private schools is the many regional school closures in the first decade of the twenty-first century, following the 2007 municipalities reform.

Danish schools follow the model of the "School for All", which, as the analysis of the Danish school system by Rasmussen and Moos (2014) shows, has changed under the influence of Denmark's transformation from a welfare state after the Second World War to a globally competitive economy from the 1990s onwards. In 2004, for example, similar to Sweden, Danish education policy decrees focused more on the evaluation of student performance and established stronger governance in the education system by national agencies for quality assurance (Houlberg et al., 2016). Correspondingly, the Folkeskole Act from 2006 represented a reordering of the purposes of schooling, and the purpose of preparing Danish students for further education and work has accordingly been strengthened. Approximately at the same

time, the schools' (and students') performances were made public, increasing competition between them. In addition, some schools have established special classes for talented students, which seems to reflect a soft form of ability grouping in the school system (Rasmussen & Moos, 2014). The overall performance of Danish students varies across different programmes. For example, Danish students show mediocre to good performance in PISA, TIMSS and PIRLS (Houlberg et al., 2016), but their performance in information-related subjects surveyed in ICILS is clearly superior due to the broadly established technical infrastructure in Danish schools and the widespread integration of digital media in teaching (Bundsgaard, Pettersson, & Puck, 2014). Similar to Sweden, the socio-economic background of Danish students has a much greater influence on school performance than in the other Nordic countries, although in recent years – as in Iceland – it has become much less significant. Thus, in international comparisons, Denmark is seen to have established equity in education (OECD, 2018).

2.3   Discussion

So, the question remains: How can diversity be maintained and respected while at the same time guaranteeing educational justice and equality of opportunity to all students?

If we look at how different educational systems address this question, we can observe that focusing solely on achieving homogeneity and assimilation through education seems to be problematic from the standpoint of diversity theories as well as from justice-based considerations. Achieving equality in this way becomes unjust if it comes at the expense of certain groups (Rawls, 1999). A pure homogenisation of differences fails to recognise the different individual needs that prevail in a diverse educational landscape and which can lead to segregation effects. What constitutes the shift of the idea of homogenisation in the education system can be seen, for example, in the No Child Left Behind orientation in the US debate on education. In the US, with the reform intended to attain equality of achievement between students, a large part of the resources was used to compensate for disadvantages such that disadvantaged students actually scored better in mathematical performance tests (Dee & Jacob, 2010). However, the simultaneous threat of sanctions at the school level for failure to achieve the set goals led to so-called "teaching-to-the-test" effects and to the disadvantaging and blaming of schools with a high proportion of low-scoring students (Darling-Hammond, 2007). Generally, within individual school classes, an unequal distribution of resources can also lead to injustice, because high achievers are disadvantaged and are no longer adequately supported under the premise of promoting the disadvantaged. This, then, does not provide a fair learning environment for all groups of students. Some researchers argue that this focus on homogenisation is also present in some Nordic countries and educational policies. Lundahl (2016), for example, describes how the introduction of free choice of schools by parents has led to segregation processes, which in turn increases

the differences between schools, and how special needs education is increasingly treated as a problem of management and accountability (and is therefore only seen as a deviation from the norm). Especially when it comes to examining the Nordic model of education in relation to its implementation of educational justice within the individual Nordic countries, we believe that several aspects are crucial, which we will discuss in the context of our previous consideration:

First, the terms "equity" and "equality" are not always used consistently in educational policy documents and, as Espinoza (2007) notes, are often confused or used interchangeably. The fact that the terms can be used in different ways, depending on the specific situation, means that the achievement of equity and equality can be interpreted and misinterpreted in many ways in educational policy documents. This makes it difficult to assess whether educational justice in the sense of the Nordic model has actually been achieved. Correspondingly, how equity and equality are operationalised and examined in empirical studies differs, with far-reaching consequences. As Blossing et al. (2014) critically note, OECD reports already speak of the achievement of equity when, for example, information about the realisation of educational opportunities is the outcome of a quantitative analysis of the socio-economic background and its relation to achievements, especially when looking at different groups like immigrant students or disadvantaged students. In the context of scientific research into educational inequalities from the perspective of educational effectiveness, such a reduction is certainly justified (see Chap. 3), but it is clearly debatable whether this mirrors current diversity discourses and their proponents (who see themselves as being exposed to the danger of being led ideologically).

Second, the formal establishment of equity alone is not sufficient to meet the requirements of a moral evaluation of fairness. For example, it must be investigated how the provision of resources in the education system can enable disadvantaged social groups to claim the right to equal educational opportunities. Furthermore, which measures can ensure the acceptance of diversity as a resource for the education system, apart from focusing on student outcomes, must be analysed. A focus on outcomes does not provide findings on how disadvantaged social groups perceive themselves and their achievements in the education system, nor whether they are valued and given the necessary attention. Such indicators are, however, particularly relevant when – as envisaged in the Nordic model – it comes to the inclusion and participation of individuals in democratic nation states, since this is one of the factors that determines how someone will behave as part of society in the future.

Third, the disadvantageous and completely contingent background conditions of some children pose a responsibility for political decision makers, teachers and society in general. These agents must strike a balance between the demands of equality (sameness in treatment) and the demands of equity (fairness of access, procedures, output and outcome). For considerations of equal opportunity, it is a government's responsibility to guarantee to each and every child an equal right to education. But in order to achieve and foster equity and diversity, their responsibility is not only to guarantee formal rights. They also need to guarantee that all individuals have the capability to realise their rights and have the material resources to do so. As we have pointed out, especially for children, the availability of basic (learning) materials is

crucial, and without them their education is impossible. This raises the question of what and how resources in education are to be used if equality of opportunities is to be achieved. An equal distribution of resources is not fair because needs are not equal. Since educational policy in the Nordic countries is increasingly based on the control and distribution of financial resources by the state, and since the educational systems of Nordic countries require considerable state resources more so than other European countries (Telhaug et al., 2006), the question of effectiveness is all the more important. However, this is not at all due to neo-liberal political considerations or the question of profitability, but is instead due to the interests of marginalised groups: How do political measures best reach those who need them? Do these measures really meet the needs of the disadvantaged?

Fourth, the orientation of educational policy in the Nordic countries towards key figures and its comparability in terms of international competition is criticised from various quarters, as this policy does not follow the original intention of equality but instead results in a stronger orientation towards school performance and stronger state governance (Blossing et al., 2014; Telhaug et al., 2006). However, the criticism fails to recognise that from the perspective of educational effectiveness, only the use of standardised and internationally comparable instruments makes it possible as objectively as possible to assess the performance of school systems and thereby give a non-biased indication of the level of achievement of equity and equality in the educational system, at least in part. The SES of students as a psychometric construct is defined in various large-scale studies using a conglomerate of different variables and is linked to students' performance in order to obtain scientifically justified statements. The index takes into account a wide range of information on parents' education, occupations, possessions, such as access to the Internet, the existence of a workplace or the number of books at home (OECD, 2019a; see also Chap. 3). From a justice perspective, this broad anchoring is to be welcomed: Despite all justified criticism of the oversimplification, the index takes into account – across all countries – whether the conditions for fulfilling the criterion of being able to pursue the formal right to education are met. At the same time, however, there is a danger that policy makers will rely too much on these indicators and will subsequently only work on changing them instead of changing the conditions that foster them.

Finally, there is still the potential to improve scientific research on equality and equity in order to provide a better basis for policy makers. For example, in light of current diversity concepts, it no longer seems appropriate to focus research on equity and equality on the attribution of immigrant student status, SES or gender differences. It is true that studies now differentiate more broadly between various marginalised groups, such as disadvantaged students, immigrant students, second-generation immigrant students or students-at-risk (OECD, 2019a). What reporting on equity and equality has in common across studies, however, is that ascriptions of being deviate from the norm are used and different groups are compared against each other. Group membership is without a doubt important for identifying the causes of inequalities but should be secondary in the description of equity. Thomsen (2013, p. 175) describes the consequence of this orientation towards attribution:

> The narrativisation of equity-as-equal outcomes and equal-opportunity-as-the-removal-of-barriers has become in national policy the arithmetic equation of the distribution of goods/benefits among population groupings in roughly the same proportion as they are in the wider society […] It is a distributive notion of equity and social justice […].

Thomsen bemoans that this notion suggests that "all those below the median/average are just 'behind' […]. When students are homogenized in this way, difference becomes a problem rather than a potential resource and strength" (2013, p. 176). Individual efforts, talents, diligence as well as lack of ambition also need to be considered in order to reach a conclusive idea of educational justice. It is, as we have shown, a matter of considering these normative aspects together, over a long period of time, since the dynamics of the social structure need to be addressed when evaluating the benefits, opportunities, obstacles and discriminations experienced by children. Here, large-scale studies being conducted to support educational policy making in different educational systems are continuously striving to develop more inclusive constructs and standards for equity. These studies should not settle for mere descriptions of differences between different groups of students; rather, against the background of the diversity discourse, they should also find measures useful for accessing the justice and fairness aspect of educational equity (including the qualitative aspects). Furthermore, these studies also has the opportunity to emphasise the positive aspects of group attributions (such as resilience) and be used to identify potential avenues by which to use information about marginalised groups as a resource for achieving equity.

References

Ahonen, S. (2014). A school for all in Finland. In U. Blossing, G. Imsen, & L. Moos (Eds.), *The Nordic Education Model. A "School for All" encounters neo-liberal policy* (pp. 77–93). Dordrecht, The Netherland: Springer.

Andersen, D., Jensen, V. M., Nielsen, S. S. Pedersen, K. C. Z., Petersen, D. S., & Thorsen, K. M. (2012). *Indsatser for tosprogede elever* [Efforts for bilingual students]. Kortlægning og analyse, SFI – Det Nationale Forskningscenter for Velfærd, København.

Antikainen, A. (2006). In search of the Nordic model in education. *Scandinavian Journal of Educational Research, 50*(3), 229–243.

Arnesen, A.-L., & Lundahl, L. (2006). Still social and democratic? Inclusive education policies in the Nordic welfare states. *Scandinavian Journal of Educational Research, 50*(3), 285–300.

Båvner, P., Barklund, A., Hellewell, A., & Svensson, M. (2011). *OECD – Overcoming school failur* (Country background report Sweden). Regerninskansliet.

Blossing, U., Imsen, G., & Moos, L. (2014). Nordic schools in a time of change. In U. Blossing, G. Imsen, & L. Moos (Eds.), *The Nordic Education Model. A "School for All" encounters neo-liberal policy* (pp. 1–14). Dordrecht, The Netherland: Springer.

Blossing, U., & Söderström, M. (2014). A school for every child in Sweden. In U. Blossing, G. Imsen, & L. Moos (Eds.), *The Nordic Education Model. A "School for All" encounters neo-liberal policy* (pp. 17–34). Dordrecht, The Netherland: Springer.

Bundsgaard, J., Pettersson, M., & Puck, M. R. (2014). *Digitale kompetencer – IT i danske skoler i et internationalt perspektiv* [Digital competencies –IT in Danish schools in an international perspective]. Aarhus Universitetsforlag.

Calmar Andersen, S., & Thomsen, M. K. (2011). Policy implications of limiting immigrant concentration in Danish public schools. *Scandinavian Political Studies, 34*(1), 27–52.

Cohen, G. A. (1983). The structure of proletarian unfreedom. *Philosophy & Public Affairs, 12*(1), 3–33.

Darling-Hammond, L. (2007). Race, inequality and educational accountability: The irony of 'No Child Left Behind'. *Race, Ethnicity and Education, 10*(3), 245–260.

Dee, T. S., & Jacob, B. S. (2010). *The impact of No Child Left Behind on students, teachers and schools* (Brooking papers of economic activity, pp. 149–207).

Egelund, N., Haug, P., & Persson, B. (2006). *Inkluderande Pedagogik i Skandinaviskt Perspektiv* [Inclusive pedagogy in a scandinavian perspective]. Liber.

Eklund, M. (2003). *Interkulturellt lärande. Intentioner och realiteter till svensk grundskola sedan 1960.* Doctoral dissertation; Intercultural learning. Intentions and realities to Swedish primary school since 1960. Luleå University of Technology.

Engen, T. O., Kulbrandstad, L. I., Kulbrandstad, L. A., & Lied, S. (2018). Education and diversity: Introduction of a multidisciplinary research group. In *Norwegian perspectives on education and cultural diversity* (pp. 2–37). Newcastle upon Tyne, UK: Cambridge Scholars Publishing.

Espinoza, O. (2007). Solving the equity – Equality conceptual dilemma: A new model for analysis of the educational process. *Educational Research, 49*(4), 343–363.

Farell, J. (1999). Changing conceptions of equality of education: Forty years of comparative evidence. In R. Arnover & C. Torres (Eds.), *Comparative education: The dislectic of the global and the local* (pp. 149–178). New York: Rowan & Littlefield.

Fjellman, A. M., Yang Hansen, K., & Beach, D. (2018). School choice and implications for equity: The new political geography of the Swedish upper secondary school market. *Educational Review, 71*(4), 518–539.

Garðarsdóttir, Ó., & Hauksson, G. (2011). *Young immigrants and persons with immigrant background in islandic society and islandic schools 1996–2011.* Netla*:* Online Journal on Pedagogy and Education. Special Edition – Menntakvika Conference 2011. http://netla.hi.is/menntakvika2011/020.pdf

Gaski, L. (2008). Sami identity as a Discursive Formation: Essentialism and ambivalence. In H. Minde, L. Gaski, S. Jentoft & G. Midré (Eds.), *Indigenous Peoples: Self-determination, Knowledge, Indigeneity* (pp. 219–236). Eburon Academic Publishers.

Graeffe, L., & Lestinen, L. (2012). Towards multicultural education in Finland. In J. A. Spinthourakis, J. Lalor, & W. Berg (Eds.), *Cultural diversity in the classroom. A European comparison* (pp. 109–128). Wiesbaden, Germany: Springer.

Gustafsson, J.-E., Cliffordson, C. & Erickson, G. (2014). *Likvärdig kunskapsbedömning i och av den svenska skolan – problem och möjligheter* [Equivalent assessment of knowledge in and by the Swedish school – problems and opportunities]. SNS Förlag.

Gustafsson, J.-E., & Yang Hansen, K. (2017). Changes in the impact of family education on student educational achievement in Sweden 1988–2014. *Scandinavian Journal of Educational Research, 62*(5), 719–736.

Haugen, C. R. (2010). Equity and life-long learning: An analysis of White Paper no. 16 (2006/2007) of Norway. *Scandinavian Journal of Educational Research, 54*(4), 357–376.

Hofmann, R. (2012). *Gesellschaftstheoretische Grundlagen für einen reflexiven und inklusiven Umgang mit Diversitäten in Organisationen* [Social theoretical foundations for a reflexive and inclusive approach to diversity in organizations]. In R. Bendl, E. Hanappi-Egger, & R. Hofmann, Diversität und Diversitätsmanagement (pp. 23–60). University of Vienna.

Holm, G., & Londen, M. (2010). The discourse on multicultural education in Finland: Education for whom? *Intercultural Education, 21*(2), 107–120.

Horst, C., & Gitz-Johansen, T. (2010). Education of ethnic minority children in Denmark: Monocultural hegemony and counter positions. *Intercultural Education, 21*(2), 137–151.

Houlberg, K., Andersen, V. N., Bjørnholt, B., Krassel, K. F., & Pedersen, L. H. (2016). *Country background report – Denmark*. OECD Review of Policies to Improve the Effectiveness of

Resource Use in Schools. Det Nationale Institut for Kommuners og Regioners Analyse og Forskning.

Husén, T. (1989). The Swedish school reform—Exemplary both ways. *Comparative Education, 25*(3), 345–355.

Imsen, G., & Volckmar, N. (2014). The Norwegian school for all: Historical emergence and neoliberal confrontation. In U. Blossing, G. Imsen, & L. Moos (Eds.), *The Nordic Education Model. A "School for All" encounters neo-liberal policy* (pp. 35–56). Dordrecht, The Netherlands: Springer.

Jakobsen, J. (2011). Education, recognition and the Sami people of Norway. In H. Niedrig & C. Ydesen (Eds.), *Writing postcolonial histories of intercultural education*. Frankfurt, Germany: Peter Lang Publishing Group.

Jónsdóttir, K. L. S., Ólafsdóttir, S., & Einarsdóttir, J. T. (2018). Linguistically diverse children in Iceland: Their family language policy and Icelandic phonological awareness. In H. Ragnarsdóttir & S. Lefever (Eds.), *Icelandic studies on diversity and social justice in education* (pp. 7–38). Newcastle upon Tyne, UK: Cambridge Scholars Publishing.

Lamont, J., & Favor, C. (2017). Distributive justice. In E. Zalta (Ed.), *Stanford encyclopedia of philosophy* (Winter 2017).. https://plato.stanford.edu/entries/justice-distributive/

Loona, S., & Wennerholm, M. (2017). Heritage language education in Norway and Sweden. In O. E. Kagan, M. M. Carreira, & C. Hitchens Chik (Eds.), *The Routledge handbook of heritage language education: From innovation to program building* (pp. 313–326). New York: Routledge.

Lundahl, L. (2016). Equality, inclusion and marketization of Nordic education: Introductory notes. *Research in Comparative & International Education, 11*(1), 3–12.

Mueller, C. W., & Parcel, T. L. (1981). Measures of socioeconomic status: Alternatives and recommendations. *Child Development, 52*(1), 13–20.

Nestvogel, R. (2008). Diversity studies und Erziehungswissenschaften [Diversity studies and educational sciences]. In GPJE (Ed.), *Diversity studies und politische Bildung* [Diversity studies and political education] (pp. 21–33). Schwalbach/Ts, Germany: Wochenschau Verlag.

Nussbaum, M. C. (2003). Capabilities as fundamental entitlements: Sen and social justice. *Feminist Economics, 9*(2–3), 33–59.

OECD. (2012). *Equity and quality in education. Supporting disadvantaged students and schools*. OECD Publishing. https://doi.org/10.1787/9789264130852-en.

OECD. (2018). *Equity in education. Breaking down barriers to social mobility*. OECD Publishing. doi:https://doi.org/10.1787/9789264073234-en.

OECD. (2019a). *PISA 2018 results (Volume II): Where all students can succeed*. OECD Publishing https://doi.org/10.1787/b5fd1b8f-en.

OECD. (2019b). *PISA 2018 results. Combined executive summaries* (Vols. I, II & III). OECD Publishing. https://www.oecd.org/pisa/Combined_Executive_Summaries_PISA_2018.pdf

Opheim, V. (2004). Equity in education. Country Analytical Report Norway. Oslo: NIFU STEP rapport 7/2004.

Pekkarinen, T. (2012). Gender differences in education. *Nordic Economic Policy Review, 1*, 165–195.

Plummer, D. L. (2003). Overview of the field of diversity management. In D. L. Plummer (Ed.), *Handbook of diversity management* (pp. 1–49). Lanham, MD: University Press of America.

Prengel, A. (2007). Diversity education – Grundlagen und Probleme der Pädagogik der Vielfalt. [Diversity education – Basics and problems of the pedagogy of diversity]. In G. Krell et al. (Eds), *Diversity Studies. Grundlagen und disziplinäre Ansätze* [Diversity studies. Foundations and disciplinary approaches] (pp. 45–67). Frankfurt am Main, Germany: Campus-Verlag.

Prengel, A. (2013). Geleitwort: Diversität und Bildung [Preface: Diversity and education]. In K. Hauenschild, S. Robak, & I. Sievers (Eds.), *Diversity education. Zugänge – Perspektiven – Beispiele* [Diversity education. Approaches – Perspectives – Examples] (pp. 11–15). Frankfurt am Main, Germany: Brandes & Apsel.

Quaiser-Pohl, C. (2013). Diversity in education and the concept of "diversity" as a topic for educational science studies. In C. Quaiser-Pohl, V. Ruthsatz, & M. Endepohls-Ulpe (Eds.), *Diversity and diversity management in education – A European perspective* (pp. 9–22). New York: Waxmann.

Ragnarsdóttir, H., & Lefever, S. (2018). Introduction. In H. Ragnarsdóttir & S. Lefever (Eds.), *Icelandic studies on diversity and social justice in education* (pp. 1–6). Newcastle upon Tyne, UK: Cambridge Scholars Publishing.

Rangvid, B. S. (2010). Sourche country differences in test score gaps: Evidence from Denmark. *Education Economics, 18*(3), 269–295.

Räsänen, R. (2007). Education for intercultural, multi-levelled citizenship in Europe: The case of Finland. In H.-G. Kotthoff & S. Moutsios (Eds.), *Education policies in Europe. Economy, citizenship, diversity* (pp. 221–237). München, Germany: Waxmann.

Rasmussen, A., & Moos, L. (2014). A School for Less than all in Denmark. In U. Blossing, G. Imsen, & L. Moos (Eds.), *The Nordic Education Model. A "School for All" encounters neo-liberal policy* (pp. 57–75). Dordrecht, The Netherlands: Springer.

Rawls, J. (1999). *A theory of justice*. Cambridge, MA: Harvard University Press.

Rawls, J. (2001). In E. Kelly (Ed.), *Justice as fairness. A restatement*. Cambridge, MA: Harvard University Press.

Robak, S., Sievers, I., & Hauenschild, K. (2013). Einleitung diversity education: Zugänge und Spannungsfelder [Introduction diversity education: Approaches and areas of tension]. In K. Hausenschild, I. Sievers, & S. Robak (Eds.), *Diversity education. Zugänge – Perspektiven – Beispiele* [Diversity education. Approaches – Perspectives – Examples] (pp. 15–35). Frankfurt am Main, Germany: Brandes & Apsel.

Rosén, J., & Wedin, A. (2018). Same but different. Negotiating diversity in Swedish pre-school teacher training. *Journal of Multicultural Discourses, 13*(1), 52–68.

Rühle, S. (2015). *Diversität, Curriculum und Bildungsstrukturen. Eine vergleichende Untersuchung in Deutschland und Finland* [Diversity, curriculum and educational structures. A comparative study in Germany and Finland]. Waxmann.

SCB. (Statistiska Centralbyrån). (2020). Befolkningsstatistik i sammandrag 1960–2019. https://www.scb.se/hitta-statistik/statistik-efter-amne/befolkning/befolkningens-sammansattning/befolkningsstatistik/pong/tabell-och-diagram/helarsstatistik%2D%2Driket/befolkningsstatistik-i-sammandrag/#Fotnoter

Sen, A. (1983). Poor, relatively speaking. *Oxford Economic Papers, 35*, 153–169.

Sigurðardóttir, A. K., Guðjonsdóttir, H., & Karlsdóttir, J. (2014). The development of a School for all in Iceland: Equality, threats and political conditions. In U. Blossing, G. Imsen, & L. Moos (Eds.), *The Nordic Education Model. A "School for All" encounters neo-liberal policy* (pp. 95–113). Dordrecht, The Netherlands: Springer.

Skeie, G. (2009). *Religious diversity and education. Nordic perspectives*. Münster, Germany: Waxmann.

Skolverket., S. (2009). *What influences educational achievement in Swedish schools? A systematic review and summary analysis*. Stockholm, Sweden: Swedish National Agency for Education.

Söderström, M., & Uusitalo, R. (2010). School choice and segregation: Evidence from an admission reform. *The Scandinavian Journal of Economics, 112*(1), 55–76.

Solstad, K. J. (1997). *Equity at risk.Planned educational change in Norway: Pitfalls and progress*. Oslo, Norway: Scandinavian University Press.

SOU. (1983). Olika ursprung – gemenskap in Sverige: Education for the vulnerable and culturally diverse. *Main focus of the Språk- und kulturarvsutredningen*. Liber/Allmänna forel.

SOU. (2014). *Staten får inte abdikera–om kommunaliseringen av den svenska skolan* [The government must not abdicate – About the municipalization of Swedish school]. Utbildningsdepartementet.

Statistics Denmark. (2020). Immigrants and their descendants. https://www.dst.dk/en/Statistik/emner/befolkning-og-valg/indvandrere-og-efterkommere/indvandrere-og-efterkommere

Statistics Finland. (2020). Immigrants in the population. https://www.stat.fi/tup/maahanmuutto/maahanmuuttajat-vaestossa_en.html

Statistics Iceland. (2019). *Immigrants and persons with foreign background 2019*. https://statice.is/publications/news-archive/inhabitants/immigrants-and-persons-with-foreign-background-8903

Statistisk sentralbyrå. (2017). *Hvordan går det med innvandrere og deres barn i skolen?* [How are immigrants and their children doing at school?] https://www.ssb.no/utdanning/artikler-og-publikasjoner/hvordan-gar-det-med-innvandrere-og-deres-barn-i-skolen

Statistisk sentralbyrå. (2020). *Immigrants and Norwegian-born to immigrant parents*. https://www.ssb.no/en/befolkning/statistikker/innvbef

Stojanov, K. (2011). *Bildungsgerechtigkeit. Rekonstruktionen eines umkämpften Begriffs* [Educational justice. Reconstructions of a contested concept]. Springer.

Svanberg, I., & Tydén, M. (1999). Ett mångkulturellt Sverige? [A multicultural Sweden?] In I. Svanberg & M. Tydén (Eds.), *I nationalismens bakvatten. Om minoritet, etnicitet och rasism* [In the backwaters of nationalism. About minority, ethnicity and racism] (pp. 78–96). Lund, Sweden: Studentlitteratur.

Telhaug, A. O., Mediås, O. A., & Aasen, P. (2006). The Nordic model of education: Education as part of the political system in the last 50 years. *Scandinavian Journal of Educational Research, 50*(3), 245–283.

Thomson, P. (2013). Romancing the market: Narrativising equity in globalising times. *Discourse: Studies in the Cultural Politics of Education, 34*(2), 170–184.

Tjeldvoll, A. (1998). Quality or equality? Scandinavian education towards the year 2000. In A. Tjeldvoll (Ed.), *Education in the Scandinavian welfare state in the year 2000. Equality, policy and reform* (pp. 3–23). New York: Garland.

UNESCO. (2009). UNESCO world report. Investing in Cultural Diversity and Intercultural Dialogue. UNESCO.

Vertovec, S. (2007). Super-diversity and its implications. *Ethnic and Racial Studies, 30*(6), 1024–1054. https://doi.org/10.1080/01419870701599465.

Vertovec, S. (2010). Towards post-multiculturalism? Changing communities, conditions and contexts of diversity. *International Social Science Journal, 61*(199), 83–95. https://doi.org/10.1111/j.1468-2451.2010.01749.

Volckmar, N. (2019). The enduring quest for equity in education: Comparing Norway and Australia. *Scandinavian Journal of Educational Research, 63*(4), 617–631.

Von der Pfordten, D. (2010). *Normative ethik*. De Gruyter.

Willms, J. D. & Tramonte, L. (2019). The measurement and use of socioeconomic status in educational research. In L. E. Suter, E. Smith & B.D. Denman (Eds.), *The SAGE handbook of comparative studies in educational research* (pp 289–304). SAGE.

Yang Hansen, K., & Gustafsson, J.-E. (2016). Causes of educational segregation in Sweden–school choice or residential segregation. *Educational Research and Evaluation, 22*(1–2), 23–44.

Young, I. (2006). Responsibility and global justice. *Social Philosophy and Policy, 23*(1), 102–130.

Young, I. (2011). *Responsibility for justice*. New York: Oxford University Press.

Zilliacus, H., Holm, G., & Sahlström, F. (2017). Taking steps towards institutionalizing multicultural education – The national curriculum of Finland. *Multicultural Education Review, 9*(4), 231–248.

Chapter 3
Measuring Equity Across the Nordic Education Systems—Conceptual and Methodological Choices as Implications for Educational Policies

Oleksandra Mittal, Trude Nilsen, and Julius K. Björnsson

> *Producing sound analyses should not only be done out of methodological considerations; the quality of analysis may also have strong political consequences.*
>
> *(Duru-Bellat & Mingat, 2011)*

Abstract Ever since international large-scale student assessments made it possible to rank countries according to their equitability, Nordic countries have topped these rankings. Nevertheless, a decline in equity has been reported lately. However, the process of empirical enquiry that leads to specific inferences on equity partly stays obscure to education decision-makers. This unawareness of the boundaries of specific methodological and analytical approaches may lead to wrong interpretations and policy implications. Therefore, our aim is to discuss and empirically illustrate how the array of choices taken throughout the research process, from equity conceptualization and operationalization to its measurement, may affect the inferences on educational equity for Nordic countries. Our sample includes fourth- and eighth-grade students from Norway, Sweden, Denmark and Finland who participated in TIMSS 2015. We applied two-level multigroup regression models within the structural equation modelling framework to investigate the sensitivity of the countries' level of equity to: (a) operationalization of the socioeconomic status measure; (b) operationalization of equity or, in other words, the method of analysis employed (e.g., bivariate analysis versus univariate); (c) single-level against multilevel analytical approaches; (d) the grade/age of students; and (e) the choice of the learning outcome across subject domains. Prior to the analyses, we estimated the comparability of SES as a latent construct between Nordic countries. Our results confirmed that some of the most common choices to measure educational equity do matter.

O. Mittal (✉) · T. Nilsen · J. K. Björnsson
Department of Teacher Education and School Research, University of Oslo, Oslo, Norway
e-mail: oleksandra.mittal@ils.uio.no

Thus, we would encourage a researcher to report elaborately on the research process and inform on its limitations because if interpreted wrongly, it may have unfavourable consequences for a particular group of individuals.

Keywords Equity · Nordic countries · TIMSS · Methodological choices

With every cycle of international large-scale assessments (ILSAs), there has been a "horse-race" with regard to not only academic outcomes (De Lange, 2006), but extending further to the creation of league tables for which country has the most equitable education system (Egelund, 2008; Heyneman & Lee, 2014; Mullis, Martin, & Foy, 2008; Mullis, Martin, Foy, & Hooper, 2017; Organisation for Economic Co-operation and Development [OECD], 2018, 2019; Schleicher, 2019). This strive for equity has been significantly shaped by the OECD and Nordic countries. Specifically, the OECD with Programme for International Student Assessment (PISA) has influenced the discourse on how equity is conceptualized and measured, and the Nordic education model has stood out as exemplary in ensuring social cohesion, justice, and security with equal access and learning opportunities for all (Telhaug, Aasen, & Mediås, 2004; Telhaug, Mediås, & Aasen, 2006; Witoszek & Midttun, 2018).

Nordic countries have topped the educational equity rankings over most of the ILSA cycles; nevertheless, a few recent studies have reported a decline in equity (e.g., Bakken & Elstad, 2012; Gustafsson, Nilsen, & Hansen, 2018; Gustafsson & Yang Hansen, 2018; OECD, 2013, 2016; Yang Hansen, 2015). This finding expands one's horizons to seek new underlying factors and examine closer the decisions that researchers take when doing inferences on the equity. Thus, in our chapter, we will illustrate empirically how a high ranking on the "equity league table" represents more of a "broad-brush picture" (Leung, 2014), as this ranking is very sensitive to the choices made by researchers throughout the process of empirical inquiry. Such rankings may hence not necessarily be a goal to strive for.

The overarching aim of this chapter is twofold: to broaden the discussion of Chap. 2 on equity and equality by adding an educational measurement perspective, and to investigate some of the challenges that are common, but not restricted, to the analysis of educational equity within the framework of ILSAs. Therefore, we intend for the theoretical part of the chapter first to give a brief explanation of what equity stands for. Next, we will describe how the current understanding of equity in education is based on UNESCO's perspective on equity as a fourth sustainable development goal. Third, we will outline the approaches to measure equity from UNESCO, the OECD, and broader perspectives. In the fourth section, we will highlight how different ways of conceptualizing and measuring equity may affect different groups of individuals. In the concluding part of the overview, we will outline the scope of research on equity in schools and discuss the operationalization of a socioeconomic status (SES) measure.

The discussion will be followed by empirical illustrations of how an equity league table of Nordic countries may change with each methodological and analytical decision taken when doing a cross-country comparative analysis with the ILSA data. To the best of our knowledge, this study represents the first attempt to address the gaps in existing research on equitability through the joint study of four Nordic education systems. Moreover, the issues investigated reflect some of the most common conceptual and methodological choices made. Thus, they will encompass: (a) the choices of a SES measure for studying equity and the comparability of SES as a latent construct between the Nordic countries; (b) the sensitivity of countries' level of equity to the method of analysis employed (e.g., bivariate analysis versus univariate); (c) single-level against multi-level analytical approaches; (d) effects of the grade/age of students on inferences about equity; and (e) changes in equity rankings related to the choice of the learning outcome across subject domains.

As a result, the second empirical part of our chapter may be regarded both as complementary to our theoretical discussion and as a stand-alone investigation. It does not address all of the problems discussed in the first part, but it serves as an example of the common thread of choices made when investigating educational equity within and across countries. In particular, these choices are to be made when academic performance is used as the criterion against which developed countries' education systems[1] are tested for fairness and inclusion (OECD, 2019). The illustrations will emphasize how fragile the conclusions on equity can be and raise a concern for how a seemingly straightforward process of investigating equity may have policy implications. Our findings would encourage researchers to report informatively on the research process (Leamer, 1983) in order to enlighten different political and educational actors about the boundaries and limitations of conceptualizing, measuring, and analysing equity within and across schools. Further, our research may contribute to disentangle the complicated question of educational equity in the Nordic countries.

3.1   Overview

In this section, we focus on the interpretation of the OECD's and UNESCO's perspectives on equity and equality. To dive deeper into the philosophical perspectives on equity in education and to see its multidimensionality, one may want to refer to Chap. 2. We further describe a number of methods to measure equity and emphasize the role our empirical inferences may have for different sub-groups of individuals.

---

[1] We specify only developed countries for a reason. If academic performance is the criterion against which equity is studied in developed countries, many developing countries still struggle to ensure equal access to education and high educational attainment. Thus, the latter remains the criterion against which the education systems of developing countries are tested for equity (Kim, Cho, & Kim, 2019).

The overview section is concluded by a summary of SES and its operationalization.

3.1.1   What Is Equity?

Equity is one of the most widely discussed topics since the end of 1990s due to economic, social, and cultural globalization, as well as a shift in the understanding of twenty-first-century values. Both the result and accelerator of these processes – namely ILSAs – further contribute to putting equity on the agenda. The concept itself, however, is not new; in fact, Coleman's (1966) report on *Equality of Educational Opportunity* stirred decades of sociological research in education revolving around the concepts of equality, equity, and equality of educational opportunity. Since then the definition of equity has undergone many transformations (see Chap. 2). From being purely theoretical, the concept of equity has become more practical and measurable in the field of education, standing alongside the concepts of educational excellence (Van den Branden, Van Avermaet, & Van Houtte, 2011) and quality (Kyriakides & Creemers, 2011). Furthermore, equity is at the heart of the post-2015 Education for All (EFA) goals set by UNESCO (Rose, 2015).

To measure educational equity, researchers commonly refer to the OECD and its broad formulation of equity as variances in learning outcomes not attributable to variances in the socioeconomic background of students (OECD, 2018). This latter definition by the OECD encompasses many ways to measure equity, which are discussed in our further sections. According to the OECD Report *"No More Failures: Ten Steps to Equity in Education"* (Field, Kuczera, & Pont, 2007), equity is divided into fairness and inclusion aspects (OECD, 2012). Inclusion implies that all acquire the minimum set of skills necessary to be a functional member of society. Fairness at the same time ensures that personal and social circumstances do not hamper educational success.

Equity in education, can also be interpreted as the concept of a "fair learning environment" (Opheim, 2004). According to this concept, each student should have access to all levels of schooling and a fair chance to succeed based on his or her abilities and needs, irrespective of background characteristics, biased expectations, and stereotypes. As a result, this interpretation of equity may lead to specific educational policies aimed at compensating for the effects of students' different socioeconomic backgrounds. Such policies may contribute to unequal treatment of students or unequal distribution of school resources, which however should not lead to discrimination of any group of students. Educational effectiveness research (EER) then investigates the extent to which schools and teachers can compensate for unjustifiable differences in both cognitive and non-cognitive outcomes (Creemers & Kyriakides, 2008; Kyriakides & Creemers, 2011). Hence, equity implies that schools have to reduce the impact of students' socioeconomic background, gender and ethnicity on their learning outcomes.

The Nordic education model is based on a drive for fairness and inclusion, as well as Rawls' principles of distributive justice and "fair equality of opportunity" (1999; Chap. 2). These egalitarian principles are foundational to the Nordic society. Consequently, small achievement gaps between students or sameness in their learning outcomes irrespective of their wealth, social status, ethnicity, cultural resources, and gender are considered to be the ideal of the equitable education system (Blossing, Imsen, & Moos, 2014; Strietholt, 2014).

It is necessary to mention that the concept of educational equity is often used interchangeably with equality. Although specific boundaries are set between the two in theory (see, e.g., Espinoza, 2007; Farrell, 1999; Holsinger & Jacob, 2009; Chap. 2), it is still challenging to address them in attempts to measure the concepts and conduct cross-country comparisons with the data and instruments at hand. In addition, cultural and political contexts within each country heavily influence the way equity is perceived and measured. For example, for the Nordic region, equality for all is essential and fair; however, some other countries believe in excellence and meritocracy[2] as the cornerstone of an equitable education system. Therefore, it is important to remember that both equality and equity in education are two sides of the same coin, and maintaining the balance between these concepts is imperative. For example, it is indeed impossible to equalize students' academic outcomes for a number of reasons. First, we all are different in so many ways[3] (Tomlinson, 1999), and distributing educational resources equally may only increase the achievement gap. Second, while two students are not likely to get the same job in their adulthood, each one must have an equally fair chance to become a productive, well-paid, and happy member of society. Thus, when inequalities in access to education and academic performance in schools arise, researchers should investigate whether and to what extent those inequalities are justified. Moreover, researchers should be aware that their decisions, including the choices of theory, definition, sample, method, analytical tools, and indicators, might have an irreversible impact on educational policies that can imbalance the scales of justice for a particular group of individuals.

3.1.2 Equity in Education as a Sustainable Development Goal

Equity has always been both a philosophical and a political concept underpinned by a variety of theoretical approaches. However, the way it is defined and measured in education currently is closely connected to the EFA goals set in 1990 at the World Conference on EFA organized by the United Nations Development Programme (UNDP), UNESCO, the United Nations International Children's Emergency Fund

---

[2] In meritocratic approach to educational equity, the emphasis is on students' effort, persistence, and initiative (Van den Branden et al., 2011). Thus, the main determinant of (non)fairness is the extent to which the students' academic performance correlates with their individual abilities and characteristics, irrespective of SES, cultural belonging, or gender (Espinoza, 2007; UIS, 2018).

[3] See Chap. 2 for the in-depth overview of diversity theories.

(UNICEF), and World Bank with Denmark, Finland, Norway, and Sweden among its co-sponsors (World Conference on Education for All [WCEFA], 1990). At the time, broad statements were made on developing human values and lifelong learning as the main goals of equity in education. Nevertheless, the focus was mainly narrowed to ensuring universal access to primary education as well as decentralization and devolution of authority and responsibility for the administration of basic education to the community. All the Nordic countries aligned with these goals, with Sweden eventually having a higher decentralized and ability-stratified educational system. In Sweden, a free school choice was implemented in the early 1990s, and researchers have claimed that this is the reason for the increased differences between schools (Gustafsson & Yang Hansen, 2018). In Norway, government officials placed a new emphasis on "equity through diversity" somewhere between 1980 and 1990 to replace the idea of "equity through equality", which had driven education reforms in Norway for a century (Solstad, 1997).

When leaders at the World Education Forum in 2000 established the Dakar Framework for Action with six education goals for the years 2000–2015, the emphasis shifted from universal primary education for all and the elimination of gender disparities to a focus on quality education, excellence for all, and equitable access to appropriate learning and life-skills programmes for young people and adults (World Education Forum, 2000). In 2015, the *Global Monitoring Report* was published by UNESCO, which had monitored progress towards the EFA goals and the two education-related Millennium Development Goals: "Achieve Universal Primary Education" and "Promote Gender Equality and Empower Women" (UNESCO, 2015). The report made it clear that educational goals and targets set back in 1990 and by the Dakar framework in 2000 were not realized to the full extent because they were vague and hardly measurable. With the new post-2015 education targets included in the fourth sustainable development goal, the focus remained on educational quality but this time centred on equity, which should be clearly articulated, realistic, and measurable (Rose, 2015). This goal mirrors the new dynamic model of educational effectiveness (Creemers & Kyriakides, 2008) that incorporates equity and quality in the studies of school effectiveness.

With equity at heart, the overarching post-2015 target from the EFA Steering Committee proposal to the UN states: "*By 2030, all girls and boys complete free and compulsory quality basic education of at least nine years and achieve relevant learning outcomes, with particular attention to gender equality and the most marginalized*" (EFA Steering Committee Technical Advisory Group, 2014). This declaration, of course, brings many equity problems to the discussion, including improving mean scores, setting minimum learning standards as introduced in some policies across the nations, estimating performance variation; and investigating gaps in learning outcomes between different groups of students, such as between top-achieving students and low-achieving students or the top 10% affluent students and the 10% most disadvantaged students (Schleicher, 2019). Other equity issues include analysing to what extent the variation in performance is attributable to students' SES, gender, or ethnicity; the equity of the distribution of secondary education; the quantity, quality, and distribution of the teaching force and educational

resources; equity and inclusiveness in education expenditures; and targeting marginalized groups of students. All of these challenges are part of the broad educational equity context and may be investigated using different types of analyses depending on the set of research questions.

3.1.3   How Can We Measure Equity?

After the unsatisfactory results presented in the *Global Monitoring Report* in 2015, and in order to make the targets on inclusive and equitable quality education clearly defined and adequately measured, the Education 2030 Framework for Action mandated the development of new indicators, statistical approaches, and monitoring tools for the assessment of progress towards the fourth sustainable development goal (UNESCO, 2015). In response, the UNESCO Institute for Statistics (UIS) published *The Handbook on Measuring Equity* in 2018. This handbook offered a set of guidelines for researchers on how equity can be defined and measured including examples of various types of analyses that can be undertaken. *The Handbook on Measuring Equity* outlined five possible methods for equity conceptualization and measurement: *minimum standards* (minimum achievement definition; Gordon, 1972), *equality of condition* (distribution of an educational variable or achievement gaps), *impartiality* (close to the concepts of horizontal equity[4] and equality of opportunity; Berne & Stiefel, 1984; Stewart, 2005), *meritocracy* (academic outcomes depend only on the child's abilities, persistence, and effort, but not on background characteristics; Gewirtz & Cribb, 2009; Van den Branden et al., 2011), and *redistribution* (re-distributing resources in favour of disadvantaged sub-groups of students, also known as vertical equity; Berne & Stiefel, 1984)*.

Like UNESCO's publications, the OECD (2004, 2018) reports on *Equity in Education* have been setting standards on equity against which countries' education systems are compared. The earlier report (OECD, 2004) touched upon equality of opportunity and "vertical equity", and took up the egalitarian stand (Rawls, 1999). The recent report formulated a broader approach to defining equity which states that, regardless of differences between students' learning outcomes, the aim is for those differences to be "unrelated to their background or to economic and social circumstances over which students have no control" (OECD, 2018). This quite open-ended definition highlights the breadth of opportunities for empirical investigation within a school effectiveness paradigm, some of which are outlined in the present book.

School performance is one of the main criteria against which developed countries' education systems are tested for fairness and inclusion. When measuring

---

[4] Horizontal equity can be interpreted as equality between different groups of individuals within a society. These groups are constructed based on individuals' cultural, social, ethnical, and geographical characteristics. Another definition is the "equal treatment of equals", which means that everyone deserves equal treatment and, therefore, an equal amount of resources (UNESCO, 2018).

equity within the framework of ILSAs at the stage of educational achievement, researchers commonly study the following (OECD, 2019; Strietholt, 2014):

1. variation in students' academic performance between and within schools which can be analysed through estimating standard deviations (SDs) (Burroughs et al., 2019), the proportion of students with minimum competency level, or achievement gaps between low- and high-achievers;
2. the width of inequality between groups estimated through bivariate multigroup analysis, such as estimating disparities in learning outcomes, for example, between socioeconomically disadvantaged and advantaged groups of students, boys and girls, or ethnic minorities and majority;
3. the extent to which educational outcomes correlate with students' social, economic, and/or cultural capital through bivariate or multivariate analysis (Sirin, 2005; White, 1982); or
4. different mediating and moderating mechanisms, represented by individual and school-level factors underlying or affecting the SES–achievement relationship (Guo et al., 2018; Gustafsson et al., 2018; Johnson, McGue, & Iacono, 2007; Kriegbaum & Spinath, 2016; Liu, Van Damme, Gielen, & Van Den Noortgate, 2015; Mood, Jonsson, & Bihagen, 2012; Rjosk et al., 2014; Steinmayr, Dinger, & Spinath, 2010).

Despite limitations when measuring and making inferences on equity within and across countries based on ILSAs' analyses (for an extended discussion see, e.g., Rutkowski & Rutkowski, 2010, 2013; Schuelka, 2013), the impact ILSAs have had on education systems worldwide within the past 20–25 years is undeniable (Grek, 2009; Schwippert & Lenkeit, 2012). Nevertheless, their potential to aid educational policies has not been fully tapped (Strietholt & Scherer, 2017). Thus, it is more important than ever to use large-scale survey data while exercising wisdom in the research (Hopfenbeck et al., 2018), as researchers bear responsibility for the policy implications their studies may have for a sub-group of individuals. Specific groups, such as high-performing students, may be left behind if educational policy focuses on one group only.

3.1.4 Who Gets Left Behind?

According to the OECD reports, equity comprises two dimensions: fairness and inclusion (Field et al., 2007; OECD, 2012). However, as the previous review revealed, the methodological approaches to study equity are mainly tailored for children from disadvantaged backgrounds or low-achieving students. While this focus is crucial, it is essential to remember that whenever researchers focus on, for instance, one specific sample of students or are driven by their own value judgements, they inevitably imbalance the scales of justice. The body of students is always heterogeneous, everyone with their own needs and abilities. There is no single solution for all, which implies educational policies should be as heterogenous

as possible. Thus, it is imperative for researchers to describe their thread of decisions starting from the theory and ending with the choice of analytical tools. Further, researchers should present implications that the obtained inferences have in a global perspective for the whole school, district, country, or internationally.

To give an example, measuring achievement disparities in the Nordic countries (and other countries) illustrates how reducing gaps between weak and strong students may increase the proportion of academically capable students (Gustafsson et al., 2018; Kyriakides & Creemers, 2011; Mullis, Martin, & Loveless, 2016; OECD, 2016). Norway is, however, an exception because, despite having small achievement gaps, Norwegian students still exhibit average or below-average academic performance with few top-performing students (Mullis et al., 2016). A possible explanation for this finding is the so-called "zero-sum game" (Rutkowski, Rutkowski, & Plucker, 2012), meaning that focusing on the low-achieving students may be at the expense of highly capable students not getting a fair opportunity to succeed. Just like there is a need for varied teaching and differentiated instruction for disadvantaged students (OECD, 2004, 2018), there is an equal need for students with higher learning potential to get appropriate support in order to realize their potential. To this end, this issue becomes one of equity, excellence, and improving knowledge economy on a global scale.

Bringing balance to education is thus important, and the More to Gain policy of the Norwegian Ministry of Education and Research reflects such an attempt, as it aims is to provide differentiated instruction not only to students who need extra support but also to those who "have special talents or potential to achieve on the highest level" (Official Norwegian Reports NOU, 2016). Therefore, when it comes to reporting on a specific type of equity, it is advisable to discuss what the results mean for different groups of individuals and what consequences they might have on educational policies in general.

3.1.5 SES, Equity, and Operationalization

Decades of educational research has shown that student family SES remains one of the most influential factors in predicting academic achievement (Sirin, 2005; White, 1982). In a meta-analysis of 499 quantitative studies, Hattie (2009) discovered that this relationship has the biggest effect size (d = .57), meaning that SES explained 57% of the variance in academic achievement. Consequently, the overarching aim for increasing equity is to prevent differences in student outcome from being attributable to SES indicators such as parents' wealth and income, power, or possessions. Several studies have investigated the relation between such background factors and student achievement (e.g., Bellens, Van Damme, Van Den Noortgate, Wendt, & Nilsen, 2019; Burkam & Lee, 2002; OECD, 2012, 2018). On the global scale, however, extensive educational reforms introduced across countries have not minimized the positive relationship between SES and educational outcomes, leading to a conclusion that educational equity has not improved (Marks, 2013, p. 172).

The linear relationship between SES and academic achievement is, of course, considerably more complex, and students' learning outcomes[5] are the result of interplay between different educational actors (Caro, Sandoval-Hernández, & Lüdtke, 2014). To understand the mechanisms behind the SES and learning outcomes association, a number of studies in the last decade have explored mediating and moderating factors, which can better explain this relationship. For example, Liu et al. (2015) investigated the mediating effects of school processes influencing the relationship between school SES and mathematic literacy. School climate and instructional quantity and quality are the most common factors explored as mediating the effect of school and classroom SES on achievement (Rjosk et al., 2014). Gustafsson et al. (2018) explored the moderating power of these predictors within schools across 50 countries participating in TIMSS 2011. In PISA 2018, a new conceptual framework for measuring equity included mediating mechanisms focusing on access to educational resources, concentration of disadvantage, and stratification policies between schools (OECD, 2019). These factors were presented in the *PISA 2018 Results* report as mediators between learning outcomes and background characteristics such as SES, immigrant status, and gender.

There exist a number of ways, both unidimensional and multidimensional, to operationalize socioeconomic background, and researchers have extensively argued that a multidimensional SES construct including social, cultural, and economic factors is more valid than a unidimensional construct (e.g., Yang, 2003; Yang & Gustafsson, 2004). This three-dimensional view of SES, which was inspired to a great extent by Bourdieu's (1986) theory, has been used as a proxy for ILSAs' SES construct. Nevertheless, in a meta-analysis of peer-SES effects, Van Ewijk and Sleegers (2010) concluded that an extensive amount of research has neglected a generally accepted three-component view of SES and operationalized it through even dichotomous variables, like reduced price lunch status, which had low effect size. Conversely, Van Ewijk and Sleegers (2010) found that the use of a thoroughly constructed composite SES led to the higher effect estimate. In our study, a number of SES indicators will be used to see the extent to which the operationalization of SES may affect inferences on educational equity.

As a composite or multidimensional indicator, SES represents a combination of different types of capital or resources that influence children's development (Bourdieu, 1986; Coleman, 1988). Researchers have investigated PISA 2000 data to determine how much of the educational outcome variance can be explained by different types of resources, namely cultural, economic, and social capital (Marks, Cresswell, & Ainley, 2006; Turmo, 2004). These studies have concluded that family cultural resources within SES constructs, most often represented by number of books at home, parental education, and/or home study supports, explain more of the variance in students' educational outcomes than economic resources for most of the countries. The same conclusion applied to five Nordic countries (Turmo, 2004),

---

[5] Students' learning outcomes represent here a broader concept including cognitive and non-cognitive domains, namely, academic performance, motivation, well-being, self-beliefs, and expectations for the future (Kyriakides & Creemers, 2011; OECD, 2019).

where only cultural capital explained a significant percentage of socioeconomic inequality (inequity), which was up to 21% in Denmark and 18% in Norway. On the contrary, in a few cases, economic and social capital explained very little variation in academic achievement, between 0% and 2% for social capital and 10% maximum for economic capital in Denmark only.

This finding about cultural capital being the most important for students' attainment and achievement can be explained by more varied cultural experiences that highly educated parents may provide for their children (Steinmayr, Dinger, & Spinath, 2012), as well as more complex and demanding communication styles or linguistic codes (Bernstein, 1971) the parents of higher education may use. This is one of the reasons for us in this study to choose specific indicators representing both unidimensional and multidimensional constructs for SES.

The overview of SES concludes our aim to discuss and review a number of issues related to the conceptualization and operationalization of equity in education. Our further aim is to present empirical evidence on the way methodological and analytical choices may alter inferences on the equitability of the Nordic education systems.

3.2 Methodology

In the empirical section of our study, we used data from Trends in International Mathematics and Science Study (TIMSS) 2015 to investigate how an equity league table of Nordic countries changed with different types of analysis.

3.2.1 Data and Sample

Our sample included all Nordic countries whose students participated in TIMSS 2015 Grades 4 and 8. TIMSS 2015 was the sixth cycle of the large-scale comparative study of fourth- and eighth-grade students' knowledge in the curriculum areas of mathematics and science, administered every four years by the International Association for the Evaluation of Educational Achievement (IEA) since 1995 (Mullis, Martin, Foy, & Hooper, 2016a, 2016b). In the fourth grade cohort, Denmark ($N = 3710$), Finland ($N = 5015$), Norway ($N = 4164$), and Sweden ($N = 4142$) participated in the survey; however, the eighth grade cohort included only two Nordic participants, Sweden ($N = 4090$) and Norway ($N = 4795$).

A two-stage stratified cluster sample design with a systematic random sampling approach[6] applied in TIMSS, with students nested in classrooms and classrooms nested in schools, results in substantial intraclass correlation (ICC) within groups,

---

[6] For more on the sampling approach, please see: https://timssandpirls.bc.edu/publications/timss/2015-methods/chapter-3.html

which violates standard statistical tests' assumption of the independency of observations (Hox, Moerbeek, & van de Schoot, 2010). For example, ICC varied from 0.06 to 0.21 for the mathematics domain in fourth grade, with the lowest ICC in Finland and the highest ICC in Denmark. These results indicate that 6% to 21% of variance in student mathematics performance in TIMSS 2015 is explained by school variability. The ICCs for science were larger and varied from 0.07 to 0.27 for Finland and Sweden, respectively. It is imperative that coefficients should be at least below 0.1 in order to avoid biased standard error estimates and type I error. In the case of ICC coefficients larger than 0.1, a multilevel analysis is usually required (Hox, Maas, & Brinkhuis, 2010).

3.2.2   Measures

We used a number of different indicators for SES in our study. We measured the first construct of SES as a latent variable that included the number of books at home and father's and mother's highest level of education. The second construct was a composite indicator of SES represented in TIMSS 2015 as a continuous variable named Home Resources for Learning that included five indicators in Grade 4: number of books at home, number of children's books at home, home study supports (i.e., own room and/or internet connection), highest level of parental education, and highest level of parental occupation. In Grade 8, the composite SES indicator was named Home Educational Resources that comprised three indicators: number of books at home, number of home study supports, and highest level of parental education. Both composite variables were index variables estimated through item response theory (IRT) internationally.[7] In addition, we included the following unidimensional indicators for SES: number of books at home, highest level of mother's education, and highest level of father's education. The number of books at home was measured through students' ratings on a five-point scale in both grades, while parents' level of education was measured by parents' ratings in fourth grade and students' ratings in eighth grade on a seven-point scale.

3.2.3   Analyses

We conducted all analyses in Mplus Version 8.4 (Muthén & Muthén, 1998–2010) and used SPSS Version 26 for preparing the data. Based on the estimates of the ICCs above (see *Data and Sample*), it was appropriate to apply two-level models when implementing a regression of mathematics and science achievement scores on SES and checking for variance both within and between schools. In addition, we were interested in explaining between-school variation in achievement.

---

[7] See http://timssandpirls.bc.edu/timss2015/international-results/timss-2015/mathematics/home-environment-support/

Hence, we applied two-level (students and schools) multi-group (across countries) regression models to data within the structural equation modeling (SEM) framework. The latent SES variable at Level 1 (within level) was aggregated to Level 2 (between level) within the multilevel SEM framework. SEM is a multivariate statistical analysis technique which takes on a confirmatory (hypothesis-testing) approach in examining the relationships between multiple observed and unobserved variables while providing explicit estimates of error variance parameters. SEM generates factor loadings of indicators on the underlying latent factor, as well as model fit indices, thereby providing measures of reliability and construct validity (Byrne, 2012; Khine, 2013). It has been widely and effectively used in studying relationships between predictors and outcomes within the framework of ILSAs of students' competencies such as TIMSS, PISA, and PIRLS (Muijs, 2012). In addition, we performed measurement invariance (MI) analyses. The test for MI allows researchers to obtain information about whether the latent construct has the same meaning for participants belonging to different groups or, in our case, to different countries. In the Mplus software, we utilized the convenience option MODEL = Configural Metric Scalar to specify, estimate, and compare different invariance models. This option resulted in common goodness-of-fit indices (Comparative Fit Index (CFI), Root Mean Square Error Of Approximation (RMSEA), and Standardized Root Mean Square Residual (SRMR)). Three levels of invariance from the basic and less restricted to the most restricted are commonly used (Rutkowski & Rutkowski, 2013). A test for *configural invariance* estimates whether the same number of indicators is loaded per latent variable across groups, while *metric invariance* tests whether the factor loadings are the same across groups, and *scalar invariance* reflects whether the scale's item thresholds are the same across groups. Metric invariance is the minimum requirement for the relations between two constructs to be compared across two countries (Vandenberg & Lance, 2000).

3.3 Findings

In the following sections, we present our findings according to the following structure: (a) estimation of measurement invariance of the SES latent construct; (b) operationalization of the SES measure; (c) levels of analysis (single- versus two-level regression): correlation between SES and performance in fourth and eighth grades, mathematics versus science domains; (d) dispersion of achievement scores among fourth- and eighth-grade students in mathematics and science domains (standard deviation); and (e) achievement gaps between the highest-SES and lowest-SES groups of fourth- and eighth-grade students in mathematics and science domains (multigroup analysis).

3.3.1   SES Latent Construct: Measurement Invariance

Before proceeding with comparing results across four Nordic countries that participated in the TIMSS 2015 cycle, it is imperative to test whether the main latent construct of SES is invariant and thus comparable across these countries. Table 3.1 shows the corresponding suggested cut-offs for the goodness-of-fit indices and their incremental changes to evaluate metric invariance (Chen, 2007; Rutkowski & Svetina, 2017).

Table 3.1 shows that the latent construct SES created out of the measures of number of books at home and father's and mother's education was invariant at the configural and metric levels. The incremental differences in the CFI and RMSEA between the models assuming metric and scalar invariance exceed the suggested cut-offs. Hence, while there is evidence for the presence of metric invariance, scalar invariance may not be met. As such, we may compare relationships between SES and other constructs or variables, but we may not compare the means of SES across countries.

3.3.2   Operationalization of SES

According to the central definition of equity within the main ILSAs' framework (e.g., TIMSS and PISA; see Mullis et al., 2016a; OECD, 2018), researchers should investigate to what extent students' learning outcomes are correlated with their background characteristics like SES, ethnicity, and gender, over which they do not have control. The operationalization of an SES measure is a complex issue and we aim to illustrate how it may affect the countries' equity ranking.

For this analysis, we used data on fourth-grade students from Denmark, Finland, Norway, and Sweden in the two-level regression of mathematics achievement score[8] on five different measures of SES, represented both with multiple and single indicators (Table 3.2).

Table 3.1 Results of the measurement invariance testing of SES latent construct measured among the fourth-grade students in TIMSS 2015 across Denmark, Finland, Norway, and Sweden

Invariance level	CFI	RMSEA	SRMR	Δ CFI	Δ RMSEA
Configural	1.000	0.000	0.000	–	–
Metric	0.996	0.035	0.024	0.004	−0.035
Scalar	0.960	0.074	0.066	0.036	0.039

Note. We used the following thresholds for good fit: CFI ≥ .95, RMSEA ≤ 0.08, SRMR ≤ 0.06. For the acceptable threshold, we used: CFI ≥ .90, RMSEA ≤ 0.10, SRMR ≤ 0.10, Δ CFI ≤ −0.01, and Δ RMSEA ≤ 0.05
In RMSEA difference test, Δ RMSEA must be ≤0.05 when testing for metric invariance and ≤ 0.01 when testing for scalar invariance

---

[8] Mathematics achievement score was computed using the IMPUTATION command out of five plausible values given in TIMSS 2015 datasets for Finland, Denmark, Norway, and Sweden.

Table 3.2 Country ranking as per mathematics achievement regression on different measures of socioeconomic background (Two-level SEM)

Country	SES measure	Country ranking[a]	Within level			Between level		
			Coef.	S.E.	R^2	Coef.	S.E.	R^2
Denmark	Latent SES	2	0.34	0.02	12%	0.64	0.07	41%
	Composite SES	1	0.32	0.02	10%	0.60	0.07	35%
	Books	2	0.28	0.02	8%	0.62	0.07	38%
	Mother's education	2	0.18	0.02	3%	0.59	0.08	35%
	Father's education	3	0.17	0.02	3%	0.65	0.08	43%
Finland	Latent SES	1	0.40	0.02	16%	0.58	0.10	34%
	Composite SES	2	0.32	0.02	11%	0.61	0.10	37%
	Books	1	0.25	0.02	6%	0.59	0.13	34%
	Mother's education	1	0.26	0.02	7%	0.49	0.10	24%
	Father's education	2	0.23	0.02	5%	0.60	0.10	36%
Norway	Latent SES	3	0.43	0.02	19%	0.69	0.07	48%
	Composite SES	3	0.35	0.02	12%	0.76	0.07	58%
	Books	3	0.25	0.02	7%	0.75	0.06	57%
	Mother's education	3	0.26	0.03	7%	0.60	0.09	36%
	Father's education	1	0.22	0.02	5%	0.58	0.08	34%
Sweden	Latent SES	4	0.42	0.02	17%	0.82	0.05	67%
	Composite SES	4	0.33	0.02	11%	0.90	0.03	80%
	Books	4	0.28	0.02	8%	0.86	0.04	74%
	Mother's education	4	0.24	0.02	6%	0.83	0.05	69%
	Father's education	4	0.24	0.02	6%	0.68	0.07	46%

Note. The regression coefficients are standardized coefficients
Coef. regression coefficient, *S.E.* standard error, *R^2* percentage of variance in mathematics achievement explained
[a]The country ranking is given as per the regression coefficient for between or school level

We determined the country ranking according to school (between) level regression coefficient estimates, and it is illustrative how, specifically in the cases of Finland and Denmark, the operationalization of SES may have an impact on which country's education system comes out as the most equitable (see Table 3.2). The strength of the relation between SES and mathematics achievement also differed significantly at the individual level depending on which SES measure was used, which confirmed Sirin's (2005) conclusion.

3.3.3   Levels of Analysis: Regression of Achievement on SES

Our next step of analysis was to compare the way results change when applying one-level regression with the TYPE = COMPLEX command versus two-level regression, as well as when performing this analysis within mathematics and science domains.

Table 3.3 demonstrates that regression coefficients are higher in a single-level model, which reflects high ICC or between-school differences and standard error estimates that are too small. Thus, failing to apply two-level regression leads to

Table 3.3 Country ranking as per correlation of achievement with SES in fourth and eighth grades across science and mathematics domains

	Country	Country ranking[a]	One-level regression			Two-level regression Within level			Between level		
			Coef.	S.E.	R²	Coef.	S.E.	R²	Coef.	S.E.	R²
Fourth grade											
Mathematics domain	Denmark	1	0.39	0.02	15%	0.34	0.02	12%	0.64	0.07	41%
	Finland	2	0.42	0.02	17%	0.40	0.02	16%	0.58	0.10	34%
	Norway	4	0.46	0.03	21%	0.43	0.02	19%	0.69	0.07	48%
	Sweden	3	0.52	0.03	27%	0.42	0.02	17%	0.82	0.05	67%
Science domain	Denmark	1	0.40	0.03	16%	0.36	0.03	13%	0.66	0.08	43%
	Finland	2	0.43	0.02	19%	0.43	0.02	18%	0.50	0.09	25%
	Norway	4	0.53	0.03	28%	0.50	0.02	25%	0.75	0.06	56%
	Sweden	3	0.55	0.03	30%	0.45	0.02	20%	0.80	0.05	64%
Eighth grade											
Maths domain	Norway	1	0.48	0.03	23%	0.43	0.03	18%	0.81	0.06	66%
	Sweden	2	0.51	0.03	26%	0.44	0.03	19%	0.82	0.04	67%
Science domain	Norway	1	0.51	0.03	26%	0.46	0.03	22%	0.82	0.06	67%
	Sweden	2	0.54	0.03	29%	0.48	0.03	23%	0.81	0.04	66%

Note. The regression coefficients are standardized coefficients
[a]The country ranking is represented for the within level of two-level regression

overestimation of SES effects at the individual (within) level and underestimation of its effects at the school (between) level. Although the ranking of countries does not change, regression coefficients and variance explained at both within and between levels vary significantly, which confirms that multilevel modelling is important to see inequalities at both individual and contextual levels. A high percent of variance in achievement is explained by school-SES in all Nordic countries.

The regression coefficients remain almost the same for the SES–achievement relationship in both fourth and eighth grades in the science and mathematics domains. Moreover, there is no significant change between the variance of achievement explained by SES in fourth and eighth grades. However, the larger share of achievement variance is explained by SES in the eighth grade at the school level for Norway, which means that school-SES plays a more important role for older students in Norway.

3.3.4 Dispersion of Achievement Scores

Another way to measure equity represented in ILSAs is to look at the dispersion of achievement between students by estimating standard deviation (SD, Table 3.4). According to Espinoza (2007), this approach is argued to measure equality for all, ensuring that all students have comparatively the same educational outcomes. With the century-long tradition of equality being fundamental to justice in the Nordic society, however, it may be challenging to separate equity from equality in education as they may encompass each other (see Chap. 2).

According to Table 3.4, all Nordic countries have comparatively low standard deviations for mean mathematics and science achievement in the fourth grade;

Table 3.4 Country ranking as per mathematics and science achievement variance among students in fourth and eighth grades

Fourth grade								
		Mathematics				Science		
Country	Country ranking	Mean (S.E.)	SD[a]	S.E.	Country ranking	Mean (S.E.)	SD[a]	S.E.
Finland	1	535 (2.0)	67	1.2	1	554 (2.3)	65	1.7
Sweden	2	519 (2.8)	69	1.7	3	540 (3.6)	73	2.5
Norway	3	493 (2.3)	72	2.0	2	493 (2.2)	69	1.6
Denmark	4	539 (2.7)	75	1.6	2	527 (2.1)	69	1.3
Eighth grade								
Norway	1	487 (2.0)	66	1.3	1	489 (2.4)	76	1.8
Sweden	2	501 (2.8)	72	1.9	2	522 (3.4)	86	2.4

Data are from Mullis et al. (2016a, 2016b)
[a]A lower SD indicates that achievement scores are closer to the mean, which reflects small achievement gaps between students. A higher SD reveals more widespread achievement scores and larger achievement gaps

however, the dispersion of achievement increases in eighth grade in the science domain in Norway and Sweden. This dispersion increase corresponds to a higher percentage of science variance explained by SES at the school level in Norway and may be due to a more ethnically diverse student population participating in TIMSS 2015 in Sweden.

3.3.5 Achievement Gaps Between the Highest-SES and Lowest-SES Groups

To define low-, medium-, and high-SES students, we used the composite variable Home Educational Resources derived by TIMSS internationally.[9] This variable contains the number of books at home, the number of home study supports, and parents' highest level of education. It has three categories (i.e., few, some, and many resources), which we used as indicators of low, medium, and high SES, respectively.

From Table 3.5, we can see the order of equitable countries in terms of achievement gaps between low- and high-SES students within the domains of mathematics and science. Computing the gaps in educational outcomes between the groups with high and low levels of SES is one approach to investigating educational equity (Schleicher, 2019). It also can be regarded as estimating the level of equality on average across socioeconomic groups of students (Espinoza, 2007). The analysis shows that Sweden is the least equitable country in the science domain, while Finland is the least equitable country in the mathematics domain. In general, the gap is larger in science than in mathematics.

In Norway, the achievement gap between the high-SES group of students and the low-SES group of students is reduced from fourth to eighth grade by 19 points in

Table 3.5 Achievement gap between low-SES and high-SES groups

Fourth grade										
		Mathematics					Science			
Country	#	Low SES	Medium SES	High SES	Gap[a]	#	Low SES	Medium SES	High SES	Gap[a]
Finland	4	441	525	563	122	2	451	543	581	130
Sweden	2	442	508	554	112	4	430	529	580	150
Norway	3	416	486	531	115	3	406	488	536	130
Denmark	1	473	526	570	97	1	452	515	556	104
Eighth grade										
Norway	1	430	478	519	89	1	416	479	527	111
Sweden	2	449	491	543	94	2	437	510	578	141

Note. Standardized S.E. varied between 0.2 to 1.4
[a]Achievement gap between the low-SES and high-SES groups

---

[9]For more on this variable, please see: http://timssandpirls.bc.edu/timss2015/international-results/timss-2015/science/home-environment-support/home-educational-resources/

science and by 26 points in mathematics. In Sweden, the gap is reduced from fourth to eighth grade by 18 points in mathematics while the achievement gap is only 9 points less in eighth grade than in fourth grade in science.

We acknowledge that our analyses produced a large body of results and hence provide a summary of the findings prior to the discussion.

3.3.6   Summary

SES was metric invariant across the Nordic countries, which means that we can compare the relation between SES and achievement across the countries. We found that how SES is operationalized was important to the ranking of the countries according to the level of educational equity. The latent construct had the strongest relation with student achievement in all countries at the within level, followed by the composite construct and then the single variables (e.g., number of books at home). However, Sweden was consistently the least equitable regardless of how one measures SES. The analytical approach also mattered for the results. Thus, when it came to the two types of regression within the SEM framework, single- versus two-level regression, we found that the within level regression coefficient was higher for the single-level approach for all countries and for both grades and subject domains (except for Finland in fourth-grade science).

Other important game-changers were the subject domain used to measure academic achievement and the grade level (fourth and eighth grades). These factors were analysed in the two-level regression of achievement on SES (at the student and school levels). We determined that the estimates were higher in science than in mathematics for both levels and all countries, except for Finland and Sweden at the between level. Furthermore, the estimates at the school level were higher in Grade 8 than in Grade 4 in Norway but were approximately the same in Sweden. However, at the student level, the estimates remained same in Norway in mathematics in both grades and even dropped by 0.04 in Grade 8 compared to Grade 4 in the science domain.

The ranking of countries according to the level of equity also varied depending on the type of equity measure. Measuring equity as the relation between SES and achievement, as opposed to measuring equity in terms of the variance in achievement (measured by SD), produced different results. For instance, using SDs, Sweden was no longer the country with the lowest level of equity. Moreover, smaller dispersions were associated with higher achievement in Grade 4 except for Norway, although this trend disappeared in Grade 8. The ranking according to SD also varied according to the subject domain, and the dispersion in achievement increased from Grade 4 to Grade 8 in both domains with the exception of Norway in mathematics.

Notably, we reached the opposite conclusion when investigating equity in terms of achievement gaps between low-SES and high-SES groups: the gap was smaller in Grade 8 than in Grade 4 for Norway and Sweden in both mathematics and science domains. Sweden had the largest gap of all the Nordic countries in the science

domain, and the gap between high- and low-SES groups remained quite large in Grade 8 in Sweden despite a small 9-point reduction from Grade 4 to Grade 8. Furthermore, Finland had the largest gap in Grade 4 in the mathematics domain. On the contrary, Denmark had the smallest gap between high- and low-SES groups in Grade 4 in both domains, thus being the first in the equity league table, though that was not the case when equity was measured in terms of the variance in achievement.

3.4   Discussion

Our first important finding was the cross-cultural comparability of the latent variable SES between the Nordic countries. We found metric invariance which reflects that the construct item factor loadings were comparable across these countries. As a result, we know that the relationships (the regression coefficients) were comparable across the countries. Cut-off criteria for evaluating relative fit was not met at the scalar level (Rutkowski & Svetina, 2017), indicating that the means of the latent variable SES were not comparable. This finding provides another perspective to resolving the one major challenge that the ILSAs are facing – the comparability of SES across the heterogenous mass of countries (Rutkowski, von Davier, & Rutkowski, 2013). For instance, the number of books at home is a common SES indicator, but it may not work as an indicator for developing countries simply because most homes cannot afford books or because there are other indicators that more accurately indicate SES in these countries. The number of books at home may thus not be comparable as an indicator of SES between developed and developing countries. Therefore, one possible solution may be to analyse groups of countries with similar cultures rather than to compare all countries within the same analysis.

We further found that the operationalization of SES mattered to the ranking of the countries, which was in line with previous research (Sirin, 2005; Van Ewijk & Sleegers, 2010). Thus, researchers should make clear what type of SES measures they use and compare their findings to previous studies that use the same type of measure. In addition, there is a possible explanation for the higher coefficient of the association between the latent SES construct and achievement rather than that using the composite SES scale. Essentially, this may be due to the degree of bias that common factor models may produce at the structural level (over- or underestimation of structural parameter estimates), which cannot always be identified through model fit (Rhemtulla, van Bork, & Borsboom, 2019). Once again, this showcases that the choice of SES measure influences the inferences, which in turn may have implications for educational policy in Nordic countries.

The equity rankings changed according to the choice between single- and multilevel regression. This finding is to be expected, as the single-level regression captures both variances between schools and between students, while the two-level regression coefficient at the within level explains only the variance between students (Rutkowski et al., 2013). What was interesting, however, was that the difference between the two within-level regression coefficients was larger for Sweden.

One explanation is that more variance in achievement can be explained at the school level in Sweden than in the other countries (OECD, 2012). The most plausible explanation for the differences between schools in Sweden as opposed to the other Nordic countries is the free school choice and the segregation between schools according to ethnicity which has increased since 2006 with some schools having 100% of students with immigrant backgrounds (Beach, Dovemark, Schwartz, & Öhrn, 2013).

Another finding with regard to the level of analysis was that the between-level regression coefficient was higher than that of the within level in the two-level regression. While this finding was in line with previous research (Van Ewijk & Sleegers, 2010), Sweden again came in last with the largest difference between the within- and between-level regression coefficient in Grade 4. Sweden was closely followed by Denmark, while Finland had the smallest difference. These findings indicated that differences between schools relative to the differences between individual students were largest in Denmark and Sweden and smallest in Finland. This was also in agreement with previous research, which determined that Finland and Norway were some of the most equitable countries in the world (OECD, 2019).

When it came to establishing a pattern in equity results across grades, the pattern for the achievement gaps between the high-SES and low-SES students was more pronounced: the gaps were smaller in Grade 8 than in Grade 4 in both subject domains. This finding could indicate that, in Grade 8, school effects play a greater role in reducing the effects of individual SES on achievement, which would be in accordance with previous research (Gustafsson et al., 2018).

The pattern concerning the subject domains pointed to lower levels of equity in science than in mathematics, regardless of how equity was measured and regardless of grade level. However, the results were more extreme in Sweden. For instance, the gap in science achievement between low- and high-SES students was larger in Sweden than in other countries. Language plays a more dominant role in science than in mathematics, and Sweden had the largest group of immigrant students (Gustafsson & Yang Hansen, 2018; Chap. 2). Hence, it could be that this larger gap in science achievement was related to the minority status of the students and their parents.

Upon comparing results between regression coefficients of the SES–achievement association and achievement gaps, we determined that the Nordic countries had small achievement gaps compared with most other countries (Mullis et al., 2016; OECD, 2019). This finding was less prominent when it came to the regression coefficients, which were comparable to many other countries and in line with previous reviews and meta-analyses (Sirin, 2005; Van Ewijk & Sleegers, 2010; White, 1982). One interpretation is that the gap between students, and especially between schools in Nordic countries, was small compared with other countries, but that the proportion of this gap explained by students' home background in the Nordic countries was similar to that of other countries. Therefore, Nordic countries are achieving their standard of *Equality for All*, which Espinoza (2007) described as each student gaining comparatively the same level of academic achievement regardless of background factors. However, these countries still have considerable work to do

in order to ensure that they achieve the equity goal of reducing the significance of parents' SES as a determinant of their child's academic success. This finding bears implications for educational effectiveness policies in the Nordic region.

Our analysis also demonstrated that of the Nordic countries, with the exception of Norway, those countries with the highest percentage of bright students had the smallest dispersion in achievement scores. This finding corresponded to previous research where high performance was associated with high levels of equity (Schleicher, 2018) or consistently low standard deviations (Gustafsson et al., 2018; Kyriakides & Creemers, 2011; Mullis et al., 2016; OECD, 2016, 2018). Norway also belongs to the group of countries with relatively low standard deviations at both stages, but the average student performance has generally been around the international average or lower. One reason could be that Norway has a long egalitarian tradition where the focus has been on lifting the low-performing students, often neglecting high-performing students (Gustafsson et al., 2018). As discussed in the theoretical section, this outcome could be a result of the "zero-sum game" (Rutkowski et al., 2012).

3.4.1   Limitations

Using cross-country large-scale surveys like TIMSS, PISA, and PIRLS introduces some limitations when investigating the question of educational equity, which relate to the groups of students being assessed and the groups of their peers being excluded from the survey design. As an example, the data is usually missing persons displaced by conflict, children in child labour or out-of-school, students attending non-standard forms of education, nomadic populations, students with disabilities or with limited proficiency in the language of assessment, and schools located in remote regions (OECD, 2016; Schuelka, 2013). Although some of these issues are not relevant for Nordic countries, there may still be exclusion from the assessment based on certain disabilities or limited language proficiency, as well as geographical remoteness or small size of schools. Excluding these particular groups of students who may need fairness and inclusion most of all also has consequences for our inferences on equity. Therefore, once a general picture and tendency for equity in schools is established, further exhaustive quantitative and qualitative research is advisable.

Another limitation is that the conclusion on equity in education could not encompass all the Nordic student populations from the eighth grade, as only eighth-grade students from Norway and Sweden participated in TIMSS 2015. However, as our objective was primarily to provide some empirical examples on how the equity league table of Nordic countries changes with different analytical and methodological choices, it may be concluded that this objective has been achieved.

In general, data from ILSAs have cross-sectional designs and hence do not allow for any causal interpretations.

3.5   Concluding Remarks, Implications, and Further Research

In our study, we briefly discussed educational equity within the global and Nordic perspectives, the common measures used to analyse the equitability of education systems, and the consequences of improving equity for one group of students. Following this discussion, we analysed how the equity league table of Nordic countries changes with the different choices a researcher makes throughout the process of empirical inquiry – choices that are not always explicitly stated in the studies on educational equity. Upon reviewing the equity league tables produced by the different measures of SES, the types of analytical approaches (single- versus multi-level regression), various ways of measuring equity (regression coefficients, dispersion in achievement, and achievement gaps between low- and high-SES students), and even different subject domains and grade levels, it is evident that these different approaches produce different results.

Therefore, the main implication of our results is that inferences about the equitability of education in different countries depend on the choices researchers make on measurements and analytical approaches. There is thus a necessity for transparency in reporting results on educational equity. Researchers need this transparency when conducting meta-analyses and reviews, and politicians and other stakeholders need it in order to draw the correct inferences and take appropriate action.

It is important to remember that equity encompasses many goals; for instance, the egalitarian ideal of equity focuses on small achievement gaps between students. However, only reporting on the achievement gaps may not be sufficient to see the complete picture, and the extent to which these gaps depend on, for instance, SES or minority status must also be investigated. Furthermore, analysing different mechanisms that may improve equity in schools such as mediation and moderation, and further research using such approaches is needed (Caro et al., 2014; Gustafsson et al., 2018).

Overall, our results show considerable variance between the Nordic countries, which could be seen as an implication for the validity of the Nordic education model. The differences between the Nordic countries may, in fact, speak against the existence of a general Nordic model. Conversely, from an international perspective, the Nordic countries are still among the most equitable countries in the world (Mullis et al., 2016; OECD, 2016, 2018, 2019). This latter perspective, seen in the view of the similar culture and educational policies of the Nordic countries, may support the concept of a Nordic model. However, while the gaps are small in Nordic countries compared to other countries, the importance of SES is not. Therefore, one may argue that whether or not a Nordic model still holds depends on the lens one uses – a Nordic or a global lens – as well as on how equity is measured and the analytical approaches taken.

In any case, it is dangerous for the Nordic countries to "rest on their laurels", as previous research has indicated that equity is deteriorating in these countries and especially in Sweden (Gustafsson et al., 2018; Hansen & Gustafsson, 2019).

Moreover, our findings show that SES explains quite a large proportion of the gaps between students. Thus, it is important to continue to investigate what we can do differently in schools in order to reduce the relationship between students' home background and their learning outcomes. Educational equity is essential for future prosperity, but it is even more essential to provide teachers, policy makers, politicians, and other educational actors with correct and transparent information so that they make the right decisions for the betterment of all.

References

Bakken, A., & Elstad, J. I. (2012). For store forventninger?: kunnskapsløftet og ulikhetene i grunnskolekarakterer [Great expectations? The Knowledge Promotion reform and inequalities in primary grades]. *NOVA Rapport, 7*(12). Oslo, Norway: Norsk institutt for forskning om oppvekst, velferd og aldring.

Beach, D., Dovemark, M., Schwartz, A., & Öhrn, E. (2013). Complexities and contradictions of educational inclusion: A meta-ethnographic analysis. *Nordic Studies in Education, 33*(4), 254–268.

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.

Berne, R., & Stiefel, L. (1984). *The measurement of equity in school finance: Conceptual, methodological, and empirical dimensions*. Baltimore, MD: Johns Hopkins University Press.

Bernstein, B. B. (1971). *Class, codes and control. Theoretical studies towards a sociology of language* (Vol. 1). London, UK: Routledge & Kegan Paul.

Blossing, U., Imsen, G., & Moos, L. (2014). Nordic schools in a time of change. In U. Blossing, G. Imsen, & L. Moos (Eds.), *The Nordic education model: "A School for All" encounters neoliberal policy* (pp. 1–14). Dordrecht, The Netherlands: Springer.

Bourdieu, P. (1986). The forms of capital. In J. Richardson (Ed.), *Handbook of theory and research for the sociology of education* (pp. 241–258). New York: Greenwood.

Burkam, D. T., & Lee, V. E. (2002). *Inequality at the starting gate: Social background differences in achievement as children begin school*. Washington, DC: Economic Policy Institute.

Burroughs, N., Gardner, J., Lee, Y., Guo, S., Touitou, I., Jansen, K. et al. (Eds.). (2019). *Teaching for excellence and equity. IEA research for education 6*. Springer Open. Retrieved from: https://doi.org/10.1007/978-3-030-16151-4_7.

Byrne, B. M. (2012). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. New York: Routledge.

Caro, D. H., Sandoval-Hernández, A., & Lüdtke, O. (2014). Cultural, social, and economic capital constructs in international assessments: An evaluation using exploratory structural equation modeling. *School Effectiveness and School Improvement, 25*(3), 433–450.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14*(3), 464–504.

Coleman, J. S. (1966). *Equality of educational opportunity*. Washington, DC: Office of Education.

Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology, 94,* 95–120.

Creemers, B. P. M., & Kyriakides, L. (2008). *The dynamics of educational effectiveness*. London, UK: Routledge.

De Lange, J. (2006). Mathematical literacy for living from OECD-PISA perspective. *Tsukuba Journal of Educational Study in Mathematics, 25*, 13–35.

Duru-Bellat, M., & Mingat, A. (2011). Measuring excellence and equity in education: Conceptual and methodological issues. In K. Van den Branden, P. Van Avermaet, & M. Van Houtte (Eds.), *Equity and excellence in education* (pp. 33–50). New York: Routledge.

EFA Steering Committee Technical Advisory Group. (2014). *Towards indicators for a post-2015 education framework*. Montreal, QC: UNESCO Institute for Statistics.

Egelund, N. (2008). The value of international comparative studies of achievement – A Danish perspective. *Assessment in Education: Principles, Policy & Practice, 15*(3), 245–251. https://doi.org/10.1080/09695940802417400

Espinoza, O. (2007). Solving the equity–equality conceptual dilemma: A new model for analysis of the educational process. *Educational Research, 49*(4), 343–363. https://doi.org/10.1080/00131880701717198

Farrell, J. P. (1999). Changing concepts of equality of education: Forty years of comparative education. In R. F. Arnove & C. A. Torres (Eds.), *Comparative education: The dialectic of the global and the local* (pp. 149–177). Lanham, MD: Rowman & Littlefield.

Field, S., Kuczera, M., & Pont, B. (2007). *No more failures: Ten steps to equity in education*. OECD Publishing. https://doi.org/10.1787/9789264032606-en.

Gewirtz, S., & Cribb, A. (2009). *Understanding education. A sociological perspective*. Cambridge, UK: Polity Press.

Gordon, E. W. (1972). Toward defining equality of educational opportunity. In F. Mosteller & D. Moynihan (Eds.), *On equality of educational opportunity* (pp. 423–434). New York: Random House.

Grek, S. (2009). Governing by the numbers: The PISA "effect" in Europe. *Journal of Education Policy, 24*, 23–27. https://doi.org/10.1080/02680930802412669

Guo, X., Lv, B., Zhou, H., Liu, C., Liu, J., Jiang, K., et al. (2018). Gender differences in how family income and parental education relate to reading achievement in China: The mediating role of parental expectation and parental involvement. *Frontiers in Psychology, 9*, 783.

Gustafsson, J. E., Nilsen, T., & Hansen, K. Y. (2018). School characteristics moderating the relation between student socio-economic status and mathematics achievement in grade 8. Evidence from 50 countries in TIMSS 2011. *Studies in Educational Evaluation, 57*, 16–30.

Gustafsson, J. E., & Yang Hansen, K. (2018). Changes in the impact of family education on student educational achievement in Sweden 1988–2014. *Scandinavian Journal of Educational Research, 62*(5), 719–736.

Hansen, K. Y., & 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.

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

Heyneman, S. P., & Lee, B. (2014). The impact of international studies of academic achievement on policy and research. In L. Rutwoski, M. von Davier, & D. Rutwoski (Eds.), *Handbook of international large-scale assessment: Background, technical issues and methods of data analysis* (pp. 37–72). Boca Raton, FL: Chapman & Hall/CRC Press.

Holsinger, D. B., & Jacob, W. J. (Eds.). (2009). *Inequality in education: Comparative and international perspectives* (Comparative Education Research Centre). Hong Kong, China: Springer.

Hopfenbeck, T. N., Lenkeit, J., El Masri, Y., Cantrell, K., Ryan, J., & Baird, J.-A. (2018). Lessons learned from PISA: A systematic review of peer-reviewed articles on the Programme for International Student Assessment. *Scandinavian Journal of Educational Research, 62*(3), 333–353. https://doi.org/10.1080/00313831.2016.1258726

Hox, J., Maas, C. J. M., & Brinkhuis, M. J. S. (2010). The effect of estimation method and sample size in multilevel structural equation modeling. *Statistica Neerlandica, 64*, 157–170.

Hox, J. J., Moerbeek, M., & van de Schoot, R. (2010). *Multilevel analysis: Techniques and applications*. New York: Routledge.

Johnson, W., McGue, M., & Iacono, W. G. (2007). How parents influence school grades: Hints from a sample of adoptive and biological families. *Learning and Individual Differences, 17*, 201–219.

Khine, M. S. (2013). *Application of structural equation modeling in educational research and practice*. Rotterdam, The Netherlands: Sense.

Kim, S. W., Cho, H., & Kim, L. Y. (2019). Socioeconomic status and academic outcomes in developing countries: A meta-analysis. *Review of Educational Research, 89*(6), 875–916.

Kriegbaum, K., & Spinath, B. (2016). Explaining social disparities in mathematical achievement: The role of motivation. *European Journal of Personality, 30*(1), 45–63. https://doi.org/10.1002/per.2042

Kyriakides, L., & Creemers, B. P. M. (2011). Can schools achieve both quality and equity? Investigating the two dimensions of educational effectiveness. *Journal of Education for Students Placed at Risk, 16*(4), 237–254. https://doi.org/10.1080/10824669.2011.610269

Leamer, E. E. (1983). Let's take the con of econometrics. *American Economic Review, 73*, 31–43.

Leung, F. K. S. (2014). What can and should we learn from international studies of mathematics achievement? *Mathematics Education Research Journal, 26*, 579–605.

Liu, H., Van Damme, J., Gielen, S., & Van Den Noortgate, W. (2015). School processes mediate school compositional effects: Model specification and estimation. *British Educational Research Journal, 41*(3), 423–447.

Marks, G. N. (2013). *Education, social background and cognitive ability: The decline of the social*. Abingdon, UK: Routledge.

Marks, G. N., Cresswell, J., & Ainley, J. (2006). Explaining socioeconomic inequalities in student achievement: The role of home and school factors? *Educational Research and Evaluation, 12*(2), 105–128.

Mood, C., Jonsson, J. O., & Bihagen, E. (2012). Socioeconomic persistence across generations: The role of cognitive and non-cognitive processes. In J. Ermisch, M. J'antti, & T. Smeeding (Eds.), *From parents to children. The intergenerational transmission of advantage* (pp. 53–84). New York: Russell Sage Foundation.

Muijs, D. (2012). Methodological change in educational effectiveness research. In C. P. Chapman, P. Armstrong, A. Harris, D. R. Muijs, D. Reynolds, & P. Sammons (Eds.), *School effectiveness and improvement research, policy and practice: Challenging the orthodoxy* (pp. 58–66). Abingdon, UK: Routledge.

Mullis, I. V., Martin, M. O., & Loveless, T. (2016). *20 years of TIMSS: International trends in mathematics and science achievement, curriculum, and instruction*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I. V. S., Martin, M. O., & Foy, P. (2008). *TIMSS 2007 international mathematics report*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2016a). *TIMSS 2015 international results in mathematics*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: http://timssandpirls.bc.edu/timss2015/international-results/

Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2016b). *TIMSS 2015 international results in science*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: http://timssandpirls.bc.edu/timss2015/international-results/

Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2017). *PIRLS 2016 international results in reading*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: http://timssandpirls.bc.edu/pirls2016/international-results/

Muthén, L. K., & Muthén, B. O. (1998–2010). *Mplus user's guide* (6th ed.). Los Angeles: Muthén & Muthén.

Official Norwegian Reports NOU. (2016). *More to gain: Better learning for students with higher learning potential* (Vol. 14). Oslo: Norwegian Ministry of Education and Research.

Opheim, V. (2004). *Equity in education: Country analytical report Norway*. Oslo, Norway: NIFU STEP.

Organisation for Economic Co-operation and Development. (2004). *Equity in education: Students with disabilities, learning difficulties and disadvantages*. Paris: OECD Publishing.

Organisation for Economic Co-operation and Development. (2012). *Equity and quality in education: Supporting disadvantaged students and schools*. OECD Publishing. https://doi.org/10.1787/9789264130852-en.

Organisation for Economic Co-operation and Development. (2013). *PISA 2012 results: Vol. II. Excellence through equity: Giving every student the chance to succeed*. Paris: OECD Publishing.

Organisation for Economic Co-operation and Development. (2016). *PISA 2015 results: Vol. I. Excellence and equity in education*. Paris: OECD Publishing.

Organisation for Economic Co-operation and Development. (2018). *Equity in education: Breaking down barriers to social mobility* (PISA). Paris: OECD Publishing. https://doi.org/10.1787/9789264073234-en.

Organisation for Economic Co-operation and Development. (2019). *PISA 2018 results: Vol. II. Where all students can succeed*. OECD Publishing. https://doi.org/10.1787/b5fd1b8f-en.

Rawls, J. (1999). *A theory of justice* (Revised Edition, 1971). Cambridge, MA: Harvard University Press.

Rhemtulla, M., van Bork, R., & Borsboom, D. (2019). Worse than measurement error: Consequences of inappropriate latent variable measurement models. *Psychological Methods, 25*(1), 30–45. https://doi.org/10.1037/met0000220

Rjosk, C., Richter, D., Hochweber, J., Lüdtke, O., Klieme, E., & Stanat, P. (2014). Socioeconomic and language minority classroom composition and individual reading achievement: The mediating role of instructional quality. *Learning and Instruction, 32*, 63–72.

Rose, P. (2015). Three lessons for educational quality in post-2015 goals and targets: Clarity, measurability and equity. *International Journal of Educational Development, 40*, 289–296.

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.

Rutkowski, D., Rutkowski, L., & Plucker, J. A. (2012). Trends in education excellence gaps: A 12-year international perspective via the multilevel model for change. *High Ability Studies, 23*(2), 143–166. https://doi.org/10.1080/13598139.2012.735414

Rutkowski, L., & Rutkowski, D. (2010). Getting it "better": The importance of improving background questionnaires in international large-scale assessment. *Journal of Curriculum Studies, 42*(3), 411–430.

Rutkowski, L., & Svetina, D. (2017). Measurement invariance in international surveys: Categorical indicators and fit measure performance. *Applied Measurement in Education, 30*(1), 39–51.

Rutkowski, L., von Davier, M., & Rutkowski, D. (Eds.). (2013). *Handbook of international large-scale assessment: Background, technical issues, and methods of data analysis*. Boca Raton, FL: Chapman & Hall/CRC Press.

Schleicher, A. (2018). *World class: How to build a 21st-century school system* (Strong performers and successful reformers in education). Paris: OECD Publishing. https://doi.org/10.1787/4789264300002-en.

Schleicher, A. (2019). *PISA 2018: Insights and interpretations*. Paris: OECD Publishing.

Schuelka, M. J. (2013). Excluding students with disabilities from the culture of achievement: The case of the TIMSS, PIRLS, and PISA. *Journal of Education Policy, 28*(2), 216–230. https://doi.org/10.1080/02680939.2012.708789

Schwippert, K., & Lenkeit, J. (Eds.). (2012). *Progress in reading literacy in national and international context. The impact of PIRLS 2006 in 12 countries*. Münster, Germany: Waxmann.

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

Solstad, K. J. (1997). *Equity at risk. Planned educational change in Norway: Pitfalls and progress*. Oslo, Norway: Scandinavian University Press.

Steinmayr, R., Dinger, F. C., & Spinath, B. (2010). Parents' education and children's achievement: The role of personality. *European Journal of Personality, 24*, 535–550.

Steinmayr, R., Dinger, F. C., & Spinath, B. (2012). Motivation as a mediator of social disparities in academic achievement. *European Journal of Personality, 26*, 335–349.

Stewart, F. (2005). Horizontal inequalities: A neglected dimension of development. In A. B. Atkinson et al. (Eds.), *Wider perspectives on global development* (pp. 101–135). London: Palgrave Macmillan.

Strietholt, R. (2014). Studying educational inequality: Reintroducing normative notions. In R. Strietholt, W. Bos, J. E. Gustafsson, & M. Rosén (Eds.), *Educational policy evaluation* (p. 51). Münster, Germany/New York: Waxman.

Strietholt, R., & Scherer, R. (2017). The contribution of international large-scale assessments to educational research: Combining individual and institutional data sources. *Scandinavian Journal of Educational Research*. https://doi.org/10.1080/00313831.2016.1258729

Telhaug, A. O., Aasen, P., & Mediås, O. A. (2004). From collectivism to individualism? Education as nation building in a Scandinavian perspective. *Scandinavian Journal of Educational Research, 48*, 141–158. https://doi.org/10.1080/0031383042000198558

Telhaug, A. O., Mediås, O. A., & Aasen, P. (2006). The Nordic model in education: Education as part of the political system in the last 50 years. *Scandinavian Journal of Educational Research, 50*, 245–283. https://doi.org/10.1080/00313830600743274

Tomlinson, C. (1999). *The differentiated classroom: Responding to the needs of all learners*. Alexandria, VA: Association for Supervision and Curriculum Development.

Turmo, A. (2004). Scientific literacy and socio-economic background among 15-year-olds—A Nordic perspective. *Scandinavian Journal of Educational Research, 48*(3), 287–305. https://doi.org/10.1080/00313830410001695745

UNESCO Institute for Statistics (UIS); FHI 360 Education Policy and Data Centre; Oxford Policy Management; Research for Equitable Access and Learning (REAL) Centre, University of Cambridge. (2018). *Handbook on measuring equity in education*. Montreal, QC: UNESCO Institute for Statistics.

United Nations Educational, Scientific and Cultural Organization. (2015). *EFA global monitoring report. Education for All 2000–2015: Achievements and challenges*. Paris: UNESCO.

Van den Branden, K., Van Avermaet, P., & Van Houtte, M. (Eds.). (2011). *Equity and excellence in education: Towards maximal learning opportunities for all students* (Vol. 50). New York: Routledge.

Van Ewijk, R., & Sleegers, P. (2010). The effect of peer socioeconomic status on student achievement: A meta-analysis. *Educational Research Review, 5*(2), 134–150.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*(1), 4–70.

White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin, 91*, 461–481.

Witoszek, N., & Midttun, A. (Eds.). (2018). *Sustainable modernity: The Nordic model and beyond*. Abingdon, UK: Routledge.

World Conference on Education for All. (1990). *World declaration on education for all*. New York: WCEFA Inter-Agency Commission.

World Education Forum. (2000). *The Dakar framework for action. Education for all: Meeting our collective commitments*. Paris: UNESCO.

Yang Hansen, K. (2015). *Measuring trend in educational equity across the Nordic education systems between 2000 and 2012: Evidence from OECD PISA studies*. Paper presented at the ECER, Budapest.

Yang, Y. (2003). *Measuring socioeconomic status and its effects at individual and collective levels: A cross-country comparison* (Gothenburg studies in educational science, Vol. 193). Gothenburg, Sweden: Acta Universitatis Gothoburgensis.

Yang, Y., & Gustafsson, J.-E. (2004). Measuring socioeconomic status at individual and collective levels. *Educational Research and Evaluation, 10*(3), 259–288.

Part II
Focus on the Schools and Teachers

Chapter 4
Teaching Culturally Diverse Student Groups in the Nordic Countries—What Can the TALIS 2018 Data Tell Us?

Julius K. Björnsson

Abstract Almost all Nordic classrooms have some or a considerable number of students with a native language different from the language of instruction. Therefore, most Nordic teachers have to address the issues this setting imposes on them. The chapter is concerned with teachers' attitudes and experiences of teaching in a multicultural setting—that is, variations in their perceived self-efficacy in multicultural classrooms. The TALIS study is used to explore these effects and relate teacher experiences with the issues of equity and diversity. Our analysis includes all five Nordic countries. A linear regression approach was used, taking into account the multi-stage sampling in TALIS. The results indicate that general self-efficacy in teaching and not specific multicultural knowledge or experience has the most significant influence on the experienced ability to handle a multicultural setting. This is a somewhat surprising, albeit reassuring, result, as it indicates that a good and trustworthy teacher education and functional general teacher competencies are the most essential ingredients in adequately handling a multicultural classroom.

Keywords TALIS · Self-efficacy · Multicultural classrooms · Nordic countries

Empirical evidence demonstrates that compared to most other regions in the world, the Nordic region has achieved a considerable degree of equity and equality, with relatively small differences between the schools in these countries. However, socio-economic status appears to have a comparably sized effect in the Nordic—as in most other—countries. Still, the Nordic schools seem to be able to counteract this effect and lift their socio-economically disadvantaged students (Agasisti, Avvisati, Borgonovi, & Longobardi, 2018).

Within the context of the Nordic region, the so-called Nordic model has often been discussed in the literature (Imsen, Blossing, & Moos, 2017; Klette, 2018;

J. K. Björnsson (✉)
Department of Teacher Education and School Research, University of Oslo, Oslo, Norway
e-mail: j.k.bjornsson@ils.uio.no

Veenis, 2014) exemplifying the idea of 'school for all.' Moreover, the view that the five Nordic countries have long been considered similar across many historical, economic and cultural strands, in addition to having comparable education systems, provided support in discussing the model. At the same time, under the auspice of neoliberal policies and trends, the Nordic systems have undergone some changes (Lundahl, 2016), which has led to an increase in differences within the Nordic arena. As an effect, it appears that some of the educational systems are less equitable than they used to be (Lundahl, 2016).

In light of these changing conditions in the Nordic countries, coupled with a considerable influx of immigrants and refugees (Karlsdottir, Norlen, Rispling, & Randall, 2018) and increased mobility between countries of a different kind than in the past, inevitable changes have occurred to the education systems, affecting schools and teachers. Therefore, it has become highly relevant to follow both the achievements of these new citizens and how schools and teachers can accommodate and adapt to the changing conditions, especially since current research indicates that schools have a vital role in this process (Gustafsson, Nilsen, & Hansen, 2016). Against this background, this chapter focuses on teachers' attitudes and beliefs about diversity in the classroom and their perception of how well they can teach in a multicultural classroom setting. In particular, the focus is on the factors related to the teaching profession, such as the experience of teaching, professional development and class characteristics, and how these factors affect teachers' perception of their own self-efficacy in a multicultural classroom. All five Nordic countries (i.e., Denmark, Finland, Iceland, Norway and Sweden) are examined using the data from the 2018 cycle of the Organisation for Economic Co-operation and Development (OECD) Teaching and Learning International Study (TALIS) (OECD, 2019a).

4.1 Effective Teachers in Multicultural Classrooms

In general, and now more than ever before, teachers need to have the ability to handle a complicated teaching environment (Kunter et al., 2013). In a situation where students come from culturally different backgrounds and many experience language disadvantages, combined with a job in which teachers also need to cater to the needs of special needs students in their classrooms, teaching competence is certainly being put to the test (Brante, 2009). If a teacher does not manage this complexity well and experiences not being able to handle this increased diversity, equity suffers with potentially long-term severe consequences for the students.

Research has underlined the importance of this ability to handle multicultural issues (Cushner & Brennan, 2007), and it has been argued that teachers' interpersonal competence and ability to adapt to the different ethnic origins of their students is a crucial factor in achieving optimal learning outcomes in a multicultural setting (Wubbels, den Brok, Veldman, & van Tartwijk, 2006). If the teacher is aware of and

has a positive attitude towards students from different cultural backgrounds, he/she appears to be more able to create a positive classroom atmosphere and meet the diverse needs of the students. Therefore, optimal results in multicultural classrooms appear to be dependent on a few essential factors, which include *monitoring and managing student behaviour, creating positive teacher-student (and peer) relationships, teaching for student attention and engagement and the attitudes and specific knowledge about multiculturality the teacher needs to possess.*

The first three factors are usually included in a variety of conceptualisations of teacher effectiveness (e.g., Hamre et al., 2013), while the fourth addresses specifically multicultural issues and attitudes (Wubbels et al., 2006). Hamre et al. (2013) provide one such example of a general model, which comprises emotional support, classroom organisation and instructional support. Each of the three constructs is further broken into several different aspects. Emotional support includes positive climate, negative climate, teacher sensitivity and regard for student perspectives and overcontrol. Classroom organisation consists of behavioural management, instructional learning formats, productivity and chaos. Instructional support comprises concept development, quality of feedback, language modelling and richness of instructional methods (Hamre et al., 2013). Models like this are complex and multifaceted, pulling together many different but related concepts, conceptualisations and behaviours, although most of them do not include multicultural aspects or specify how the general factors interact with multiculturality.

Such models have been linked to student performance and have revealed that teachers that score high on all these aspects tend to have students that perform better and show more progress (Stronge, Ward, & Grant, 2011). Using a hierarchical linear model, Stronge et al. (2011) showed that there were differences between the teachers who were the most and least effective on their students' performance. The effective teachers experienced less student disruptive behaviour and better interactions with their students and had better classroom management and personal qualities. No significant differences were found across different instructional or assessment methods.

Specific instruments have been developed to measure teacher effectiveness and capacity to handle a multicultural situation (Spanierman et al., 2011). Such instruments focus on what teachers do in these situations, illustrating that without a positive attitude to multiculturalism, quality teaching suffers. Therefore, teachers' attitudes towards handling these issues and their perception of their own ability to handle the variety of these diversities—which are common in most classrooms today—are of paramount importance. Research so far indicates that these perceptions and experiences may go hand in hand with teachers' increased competence in teaching minority students and increased positive attitudes towards such student groups (Glock, Kovacs, & Pit-ten Cate, 2019). Moreover, research has shown that teacher beliefs and attitudes influence students. Geerlings, Thijs, and Verkuyten (2019) recently studied how teacher norms about cultural diversity and practices interact to affect students' perspectives and how these effects differed for minority and majority students. The data from the study included Dutch, Turkish-Dutch and Moroccan-Dutch students from the fourth to sixth grades. The results showed that

all students tend to have a more positive attitude towards ethnic groups when they perceive their teacher doing so. This made clear that students in multicultural classrooms are highly influenced by both their teacher's beliefs and installed classroom practices.

While teacher effectiveness can be conceptualised in many ways, most authors include in the definition of teachers' ability to maintain proper classroom management, students' cognitive activation and the fostering of a supportive climate for all students (Fauth, Decristan, Rieser, Klieme, & Büttner, 2014; Kunter et al., 2013). However, teacher attitudes towards and specific knowledge about multicultural issues are most often not included in these definitions. Given the diversity teachers face in their everyday practice, the ability to handle multicultural situations should become part of such conceptualisations. In this way, we ensure and increase our understanding of what makes teachers stay effective in a multicultural setting (Au & Raphael, 2000).

4.2   Equity and Classroom Diversity in the Nordic Countries

Changes in the Nordic school systems with increasing diversity, an increased number of students with diversified cultural backgrounds and the changes in policies in some of the countries—especially Sweden and Norway—may lead to reduced equity and equality in the Nordic schools. Large-scale international studies, such as Trends in International Mathematics and Science Study (TIMSS) and Programme for International Student Assessment (PISA), have shown that diversity increases over time considering the increased numbers of immigrant students and variation in socioeconomic status, perhaps especially in Sweden (OECD, 2018b). These changes appear to be closely linked to policy changes (Lundahl, 2016) and the amount of immigration in each country (Karlsdottir et al., 2018), which also varies considerably.

Handling increased student diversity goes hand in hand with the question of how different education systems are able to cater to the needs of such students and the extent to which they are offered equal chances to succeed within the system. Within this context, equity and equality are two terms that are often somewhat interchangeably used in the education literature (Espinoza, 2007, see Chap. 2 in this volume for a further discussion). The mixing up of the two terms is noticeable even in the OECD's extensive report on equity in education (OECD, 2018a). The report starts by stating how 'equity in education means that schools and education systems provide equal learning opportunities to all students' (p.13), a statement which is clearly about equality, not equity. The report goes on with elaborating that 'equity does not mean that all students obtain equal education outcomes, but rather that differences in students' outcomes are unrelated to their background or to economic and social circumstances over which students have no control'. Therefore, despite its extensiveness, the equity-equality paradigm is observed through the relationship between educational achievement and socioeconomic status, clearly indicating that the

difference between equity and equality requires further elaboration. Espinoza's (2007) conceptualisation deepens and clarifies this dichotomy by viewing the particular contributions schools may have on the equality and equity continuum in a more varied way. From the stance of multicultural classrooms in particular and how well teachers cater to student diversity, very much connects to his idea of access to education quality, in the sense that all students having equal abilities will gain such access (i.e., equity for equal potential) and will not be constrained by coming from diverse backgrounds (i.e., equality of opportunity).

Although labelled in a somewhat different fashion, the idea of access to quality instruction is very much a central component of the TALIS (OECD, 2019a). Within the TALIS framework, two main perspectives can be recognised: first, the perspective of how to increase equity in schools, by integration or fostering equality and inclusion and valuing diversity in the classroom, and second, by evaluating the so-called 'multiculturalism', which means that schools should acknowledge that differences in culture can and will enrich student life (Ainley & Carstens, 2018). Additionally, TALIS is concerned with teaching and learning in socioeconomically diverse groups of students, an aspect strongly connected to the first two.

The TALIS conceptualisation of equity underlines the complexity of the concept, interlinking the issues of equity and cultural diversity, which touch on most—if not all—aspects of teaching and learning. This is especially true in Europe where migration is an ever-increasing factor when considering equity in schools and education and when evaluating the effects of migration on children and their situation (Moskal & Tyrrell, 2015). These two themes, equity and diversity, are naturally closely linked, as equity issues become more important to a larger number of students with increasing heterogeneity. The TALIS 2018 framework further defines different sources of diversity. Among others, these include socioeconomic diversity, cultural diversity and gender. Equity issues connected to changes in all these areas require knowledge about school policies, teaching practices and approaches to teaching, and touch upon most, if not all, aspects of the organisation of teaching and learning.

Important information on the changes over time concerning diverse equality and equity issues across all the Nordic countries can be found in the results from the TALIS. The study focuses on the attitudes and beliefs of teachers and principals concerning many aspects of their profession. Among other things, it explores how well teachers and school leaders experience that they are able and willing to handle the increasing cultural diversity in schools. As to the Nordic countries, the data has shown considerable changes in the composition of the student body in the three completed study cycles and that immigrant students or students from different cultural backgrounds are now a large and significant part of the Nordic classrooms. Table 4.1 illustrates the situation in the Nordic countries concerning having a native language other than the language of instruction, as per the TALIS 2018.

These numbers indicate that most schools have many students who need specialised instruction and additional language support, with Sweden having the highest percentage of such students. The numbers also indicate that there is a number of schools in the Nordic countries where there are no such students. Denmark is an exception, since almost all Danish schools have students that speak a language at

Table 4.1 Percentage of schools with students having a different language than the language of instruction (from the Principal questionnaire in the TALIS 2018)

Country	Percentage of schools with more than 10% of students that speak a different language than the language of instruction	Percentage of schools with no students that speak a different language than the language of instruction
Denmark	25	3
Finland	17	14
Iceland	34	11
Norway	17	14
Sweden	46	13

home that is different than the language of instruction. Therefore, there are considerable differences between the systems observed, and these numbers also indirectly indicate that there is a clustering of immigrant students in certain schools in all five countries.

However, when looking more closely at the teachers in these schools, a somewhat different picture emerges. In 2018, teachers answering the TALIS questionnaire indicated that in almost 90% of the cases, they are teaching a class with some or many special needs students, including those with language difficulties. Furthermore, on average, 77% of the Nordic teachers indicated that they have some experience teaching students with a different cultural background, the highest percentage (86%) being in Sweden (OECD, 2019a). Therefore, it is clear that cultural and other kinds of diversity are widespread in the Nordic school systems and are continually growing more substantial and becoming a feature of most classes in the Nordic schools. A more detailed table based on the teachers' evaluation can be found in the Appendix.

4.3   Aim of the Chapter

The short review above underlines the fact that multicultural attitudes and teacher self-perception in dealing with a multicultural classroom are important aspects of overall teacher efficacy in providing quality instruction to all students. It is therefore of considerable importance to examine how Nordic teachers experience this changing situation and how they perceive themselves in addressing a multicultural setting in the classroom. Given earlier research that Nordic schools in general seem to be able to counteract the negative effects of student diversity (Agasisti et al., 2018), it is essential to investigate these perceptions across all five countries. The TALIS data, with their international perspective, can aid in examining the extent to which such perceptions are uniform or not.

Therefore, the aim of this chapter is to investigate the variations in *self-efficacy in multicultural classrooms* among Nordic teachers and to identify whether particular aspects of teacher background can predict that variation. The chapter explores whether the perception of teachers' self-efficacy in a multicultural setting has a

two-level component—that is, if teachers from different schools are meaningfully different in their perception of their own self-efficacy. Such differences could indicate that different school policies or practices influence teachers in different ways. With the increasing diversity in schools across the Nordic countries, this could be an essential feature of local school policies, leading to differences between schools (Klette, 2018).

4.4 Methods

The current investigation uses data from the TALIS. Implemented in 48 countries, this study was initiated by the OECD and was in its last cycle in 2018. The TALIS provides a detailed questionnaire for teachers and school principals, administered online. The study was conducted twice before the 2018 cycle, in 2008 and 2013. All additional information regarding the study can be found in the TALIS framework (Ainley & Carstens, 2018) and the accompanying technical report (OECD, 2019b).

All questionnaires were administered to lower secondary (ISCED 2) teachers and school principals in all the participating countries, although some countries added the same or similar questionnaires to either or both ISCED 1 (primary schools) and ISCED 3 (upper secondary schools), resulting in three populations in these countries. In this chapter, only data from lower secondary schools (ISCED 2) from Denmark, Finland, Iceland, Norway and Sweden are examined.

The TALIS uses a stratified two-stage probability sampling design (OECD, 2019b), which sampled schools primarily; a sample of teachers was subsequently drawn from each selected school. Information on the number of schools and teachers in the final sample from the five Nordic countries is provided in Table 4.2.

4.4.1 Variables

The TALIS aims to deliver information on teachers' instructional and professional practices, school leadership, teachers' initial education and initial preparation, teacher feedback and development, school climate, job satisfaction and motivation, teacher human resource measures and stakeholder relations, teacher self-efficacy, innovation and, finally, equity and diversity.

In addition to the individual questionnaire items, the TALIS database includes a number of scales and indices, which can be divided into two types. The first are

Table 4.2 The final number of ISCED-2 participating schools and teachers in the Nordic countries

	Denmark	Finland	Iceland	Norway	Sweden
No. of schools	141	148	122	185	180
No. of teachers	2001	2850	1276	4153	2779

simple summary indices, such as the number of years of teaching experience, the number of different teaching assignments, etc. By contrast, the second type comprises more complex indices (i.e., latent variables), constructed with confirmatory factor analysis (CFA), where an integral part of the scale construction is invariance testing. This is a highly important aspect of the scales, as cultural differences can certainly influence the teachers' answers heavily and possible cultural bias must be accounted for before using the constructs for between-country comparisons. All these indices are based on different questions on the same or related themes. They were all tested for adequate model fit and reliability and finally transformed into a standardised format for inclusion in the final database. All the constructs were based on three or more items (OECD, 2019b).

Almost none of the scales constructed for the ISCED-2 sample (with one exception, innovation) are scalar invariant; therefore, their values or averages cannot be directly compared across countries. However, most of them fulfil the criteria for metric invariance, rendering the results from separate analyses from each included country comparable. The analysis performed here adheres to this and was done separately for each of the five Nordic countries.

The initial analysis included multiple constructs. Three are simple summary indices, namely teacher age and teacher experience, measured in number of years, and an index reflecting the sum of diversity in the teacher's class. The latter is a simple construct addressing diversity in the composition of the target class. The teachers were asked about the proportion of students with a first language different from the language of instruction, low academic achievers in the class, special needs students, students with behavioural problems, students from socio-economically disadvantaged homes, academically gifted students, immigrant students and students who are refugees. The sum of all the answers is a measure of the classroom diversity. All other indices belong to the complex group. These include teaching and professional practices, teachers' motivation, feedback and development, teachers' self-efficacy, job satisfaction, work stress and well-being, school climate, equity and diversity and team innovativeness. Table 4.3 provides a comprehensive overview of these constructs. For more details, see the TALIS technical report (OECD, 2019b).

A backward stepwise regression analysis (OLS-ordinary least squares) was used in the initial analyses including all the described variables. At each step, the variables and indices having a non-significant relationship with self-related efficacy in multicultural classrooms (the dependent variable) were removed from the model. The model was rerun for each country until there were 11 variables and indices left, all of which were significantly related to the dependent variable, self-related efficacy (SEFE) in multicultural classrooms, in at least one of the countries observed in the analyses. Table 4.4 provides descriptive information on the constructs included in the final analysis.

Table 4.3 Complex indices included in the initial analysis

Scale	Description
Teaching practices—overall	A composite scale on teaching practices with three subscales, comprising 12 items in total; the subscales include: -clarity of instruction (4 items on how one teaches, what a teacher does), -classroom management (4 items on how the class is managed, rules, listening and managing disruptions) and -cognitive activation (4 items on teaching methods, tasks, critical thinking, group work).
Teacher cooperation—overall	A composite scale comprising 10 items with two subscales: -professional collaboration in lessons (5 items describing cooperation among teachers in classrooms) and -exchange and cooperation among teachers (5 items describing collaboration among teachers in the school).
Teacher motivation and perceptions	A composite scale comprising 10 items, including the following subscales: -personal utility value of teaching (4 items on why teaching was chosen as a career), -social utility value of teaching (3 items on how teaching allows one to benefit society) and -perceptions of value and policy influence (3 items on how others value teachers).
Feedback and development	A composite scale comprising 8 items with two subscales: -effective professional development (4 items on the quality of professional development) and -teachers' need for professional development for diversity (4 items on teaching individual learning, training in special needs teaching and teaching in multicultural settings).
Teacher self-efficacy—overall	A composite scale comprising 12 items with three subscales: -self-efficacy in classroom management (4 items on what the teacher does in the classroom), -self-efficacy in instruction (4 items about questions for the students, assessment strategies, explanations and instructional variety) -self-efficacy in student engagement (4 items on influencing students' beliefs, attitudes and motivation).
Job satisfaction	A composite scale of 13 items with three subscales: -job satisfaction with work environment (4 items on working at school), -job satisfaction with profession (4 items on the advantages and choice to become a teacher) and -satisfaction with target class autonomy (5 items on to which extent the teacher has control over their teaching methods).
Work stress and well-being	A composite scale comprising 12 items, with the following subscales: -workplace well-being and stress (4 items on stress and how it affects other things and well-being generally), -workload stress (5 items on having too much to do in many areas) and -student behaviour stress (3 items describing stress connected to classroom management, difficult students, and being responsible for student achievement).

Table 4.3 (continued)

Scale	Description
School climate	A composite scale of 13 items, with the following subscales: -teachers' perceived disciplinary climate (4 items on what happens in the lessons), -teacher-student relations (4 items on whether the teacher gets along with the students, is interested in them and provides them with help) and -participation among stakeholders (5 items on what the school provides for the collaboration between teachers, parents and students).
Equity and diversity	A composite scale comprising 9 items with two subscales: -self-related efficacy in multicultural classrooms (5 items on how well the teacher perceives their own ability to handle multicultural classrooms) and -diversity practices (4 items on what the teacher specifically does in a multicultural situation).
Team innovativeness	Four items on how the teachers and the school develop new ideas and methods.

Table 4.4 Variables included in the final analysis

		Denmark		Finland		Iceland		Norway		Sweden	
Variable	Type	Avg	SD	Avg	SD	Avg	SD	Avg	SD	Avg	SD
Experiences as a teacher (years)	Simple	15.2	10.0	15.8	9.5	15.0	10.4	14.7	10.5	15.5	9.9
Age (years)	Simple	44.4	10.4	44.8	9.8	46.2	10.5	43.8	11.2	45.7	10.4
Sum of diversity in class	Simple	17.9	4.1	16.9	4.2	17.9	4.0	18.2	3.9	20.0	4.9
Workload stress	Construct	9.2	2.6	9.2	1.4	9.2	2.0	9.2	1.8	9.2	1.9
Teaching practices—overall	Construct	11.5	2.0	11.5	2.0	11.5	2.0	11.5	1.8	11.5	1.9
Teacher-student relations	Construct	13.3	1.8	13.3	1.9	13.3	1.9	13.3	1.8	13.3	1.9
Social utility value of teaching	Construct	12.2	2.1	12.2	2.6	12.2	2.7	12.2	2.3	12.2	2.4
Teacher self-efficacy	Construct	12.7	1.6	12.7	2.24	12.7	2.0	12.7	1.5	12.7	1.9
Needed professional development for diversity	Construct	9.9	1.9	9.9	2.0	9.9	2.2	9.9	2.0	9.9	2.5
Job satisfaction	Construct	12.0	2.1	12.0	2.1	12.0	2.0	12.0	2.0	12.0	2.2
Disciplinary climate	Construct	8.7	1.9	8.7	2.3	8.7	2.0	8.7	2.0	8.7	2.3
Self-related efficacy in multicultural classrooms (dependent variable)	Construct	11.2	1.6	11.2	1.8	11.2	2.4	11.2	1.8	11.2	1.8

Note that the averages are the same for the complex CFA constructs in the table, as they reached only metric invariance and were standardised for each country. However, the SD of the constructs differs between the countries

4.5   Data Analysis

The data used in this study were analysed and handled with SPSS and the IDB Analyzer from IEA (IEA, 2019) using teacher weights and the 100 replicate weights of the TALIS (i.e., Balanced Repeated Replication [BRR]), thus taking into account the two-stage sampling. The backward stepwise regression was done separately for each country. The two-level regression analyses were performed for each Nordic country separately with Mplus 8 (Muthén & Muthén, 1998–2017). Only the variables that had a significant relationship to SEFE in one or more of the five countries were included in the final model.

Missing data. Before any analysis, the data were checked for missingness. For the indices related to a work situation in general, the amount of missing answers was rather low across all the Nordic countries. For example, the amount of missing answers on professional development was 6% on average, while the index on job satisfaction was also about the same. However, when it came to the cultural diversity scales, the situation was very different. For the dependent variable in this analysis, the index of self-rated efficacy in multicultural classrooms, over 29% of the answers on average across the Nordic countries were missing, with Iceland having 40,6% missing and the lowest being Sweden and Finland, with about 23% missing on these questions. Table 4.5 shows the missing percentage of SEFE in each country.

The statistical modelling of the constructs (i.e., Structural Equation Modelling-SEM and Confirmatory Factor Analysis-CFA) takes into account the missing values using a model-based approach to estimating them (OECD, 2019b). Thus, the model makes it possible to use data from all the countries, assuming that the data are missing at random (MAR)—an aspect which should not be overlooked when interpreting the results.

In addition to this, all the simple and complex scales were inspected for missing data in age, gender and teacher experience. Age shows a more substantial amount of missing answers only in Norway in the age group 25–29 and a smaller amount in the 60+ age group in Finland. No differences between age groups are observed for Denmark. As to gender, the number of missing values revealed no significant differences between males and females: 25% of females and 28% of males had missing values across the countries overall. Concerning the total experience as a teacher in these five countries, they all had a similar mean length of experience, about 15.2 years. Variation between the countries was almost non-existent. Furthermore, there is no increase or decrease in the number of missing values according to the length of experience as a teacher; the pattern appears to be mostly random, so the MAR assumption of the SEM modelling appears to be upheld.

Table 4.5 Percentage of missing answers on the index of self-efficacy in multicultural classrooms

Denmark	Finland	Iceland	Norway	Sweden
35.7	22.7	40.6	25.5	22.4

4.6   Results

This section will first describe the two-level regression model and follow with the one-level models. The two-level model was tested first in order to ascertain whether such a model was necessary.

4.6.1   The Two-Level Model

Because of the two-stage sampling employed in the TALIS and the possibility of a school-level influence (i.e., between level variance), a two-level regression analysis was carried out (Geiser, 2013). For this purpose, a null model checking the intra-class correlation (ICC) of the SEFE variable was obtained. This analysis indicates whether teachers' multicultural self-efficacy has a significant between-schools component.

   The data in Table 4.6 indicate an inconsistent level of ICC in the five countries. The ICC is higher than 0.05 only in Sweden and Norway, indicating that only these two countries have significant differences between schools when controlling for variability among the teachers.

   The predictor variables used in a further two-level regression analysis were all the indices and constructs described earlier, with the analysis performed separately for each country. However, none of these variables were significant at the school level, not even in Norway and Sweden when looking at their relationship with SEFE. Therefore, this type of analysis was not pursued further, and a conventional OLS one-level backward stepwise regression analysis was conducted as described earlier.

4.6.2   One-Level Models

The one-level multiple regression analysis performed separately for each country yielded the results shown in Table 4.7. All the TALIS constructs described earlier were initially included in the model. However, in the final analysis, we included only those with a significant relation to SEFE in one or more of the five examined countries. The final analysis estimated the exact same model in all five countries.

   This regression analysis result shows the relation between the teachers' perceived competence to handle multicultural classrooms and students and the variables that contribute to its prediction. All the shown coefficients are significant at

Table 4.6 Intra-class correlations for self-efficacy in multicultural classrooms

	Denmark	Finland	Iceland	Norway	Sweden
ICC	0.038	0.023	0.033	0.073	0.093

Table 4.7 Multiple regression analysis of self-reported efficacy in multicultural classrooms

	Denmark	Finland	Iceland	Norway	Sweden
Experiences as a teacher—total (years)	−0.11 (0.05)	−0.10 (0.04)	−0.11 (0.05)	−0.08 (0.03)	−0.13 (0.03)
Age (years)	ns	0.13 (0.04)	ns	ns	ns
Workload stress	ns	ns	−0.15 (0.04)	−0.05 (0.02)	−0.06 (0.02)
Teaching practices—overall	0.11 (0.04)	0.16 (0.03)	0.09 (0.05)	0.15 (0.02)	0.14 (0.03)
Teacher-student relations	ns	0.06 (0.03)	0.09 (0.04)	0.06 (0.02)	0.07 (0.03)
Social utility value of teaching	0.08 (0.03)	0.10 (0.03)	ns	0.10 (0.02)	0.09 (0.03)
Teacher self-efficacy—overall	0.27 (0.03)	0.23 (0.02)	0.29 (0.04)	0.23 (0.02)	0.24 (0.03)
Needed professional development for diversity	−0.10 (0.03)	ns	ns	−0.09 (0.02)	−0.10 (0.02)
Job satisfaction	ns	0.05 (0.02)	ns	ns	0.05 (0.02)
Disciplinary climate	ns	ns	ns	−0.07 (0.02)	ns
Sum of diversity in class	0.08 (0.03)	0.14 (0.03)	0.08 (0.04)	0.15 (0.02)	0.17 (0.02)
R-square	0.15 (0.02)	0.20 (0.02)	0.19 (0.03)	0.19 (0.02)	0.22 (0.02)

Note. Displayed coefficients are significant at $p < = 0.05$
Standardised regression coefficients, (Standard Error) in parenthesis

the 0.05 level in at least one of the countries, while the non-significant coefficients are not displayed. The shared model explains between 15 and 22% of the total variation in self-efficacy in multicultural classrooms, which is a sizeable proportion. The explanatory value was 15% in Denmark, 20% in Finland, 19% in Iceland and Norway and 22% in Sweden.

When observing the contributions from the different variables independently, the most significant predictor across all the countries is *teachers' overall self-efficacy*, with the markedly highest relation to SEFE. Still, *teaching practices*, the *amount of diversity in the classroom*, and the *length of experience as a teacher* are all variables that have a similar effect in all the countries, allowing for observing some common patterns. It must be noted that *length of experience as a teacher* has a negative relation to the perceived multicultural attitudes.

It is also interesting to note that more of the variables shown in the table are more significant in Norway and Sweden than in the other three countries. In most cases, the regression coefficients depict a 'logical' relationship to the SEFE variable, with rather small differences between the countries. However, there is one exception—length of experience being a teacher—which has a negative relationship to SEFE.

Workload stress has a considerably larger relation to SEFE in Iceland than in Norway and Sweden, while it does not seem to affect teachers' SEFE in Denmark and Finland. Age is another variable that only has a relation to SEFE in Finland but appears to be unrelated to SEFE in all the other countries. Teacher-student relations have a similar effect in four countries, except Denmark. The social utility value of teaching is positively related to SEFE in all the countries, except for Iceland. And, interestingly, the multicultural attitudes of teachers in Finland and Iceland have no relation to their need for more multicultural professional development. Last, job satisfaction is significant only in Finland and Sweden, while Norway is the only country with a negative relationship between SEFE and class disciplinary climate. Therefore, although the coefficients are small, all these variables underline some differences between the countries.

4.7   Discussion

All the descriptive information on equity and diversity presented here indicates that the multicultural attitudes of Nordic teachers are in some respects quite similar, a finding well aligned with the guiding concepts of the Nordic school model (Veenis, 2014). At the same time, clear differences are also observable, especially in Sweden—which might be expected since the country has the highest number of recent immigrants and refugees (Statistics Norway[1]). Surprisingly, Iceland also has a large number of students with potential language barriers. Still, one issue concerning any comparison is whether these large groups of students from different cultures are comparable across these countries at all. This is probably not the case, as Sweden, for example, has a large number of recent refugees from the Middle East, while Iceland has mostly work-related immigrants, primarily from eastern European countries. These differences in the type of immigrants are not reflected in the TALIS data but are a concern and become important when trying to understand the difficulties connected to integrating these students in the Nordic schools (Karlsdottir et al., 2018). This means that even though the Nordic school systems are similar, there are differences in the student body not reflected in the TALIS database. The TALIS questions do not differentiate between, for example, immigrants that are refugees from wars and hardship and those that come from more peaceful circumstances but are immigrating to increase their standard of living. Future studies focusing on these nuances are therefore needed.

When looking at the results from the regression analysis, it is—perhaps not surprisingly—apparent that the one variable having the largest impact is general self-efficacy in teaching. This is a composite index consisting of self-efficacy in classroom management, instruction and student engagement. Therefore, this is a broad measure of teacher self-efficacy but does not include the ability to handle

---

[1] https://www.ssb.no/befolkning/statistikker/innvbef/aar

multicultural classes or attitudes towards multiculturalism. In other words, if a teacher is generally competent and their experience is that they can handle most teaching situations well, they are probably also comfortable in a multicultural setting and will perceive themselves as able to handle such circumstances adequately. This conclusion is partially upheld in the literature, where teachers with high teacher self-efficacy have been shown to do better with minority students that others (Jenkins-Martin, 2014). Furthermore, recent studies indicate the importance of positive attitudes towards multicultural students where this appears to enhance learning (Sela-Shayovitz & Finkelstein, 2020).

General self-efficacy in teaching appears to be highly comparable and similar across the Nordic countries and is the most important variable in the whole regression analysis. Age does not influence this relationship, except in Finland where increased age (and therefore, presumably experience) appears to have a positive effect. However, experiences as a teacher in total (i.e., measured as the number of years) does have a negative and similarly strong relation in all the countries. This is somewhat counter-intuitive—i.e., that a longer experience as a teacher should lead to weaker self-efficacy in handling multicultural classrooms. One explanation could be that older, more experienced teachers did not experience these situations initially in their career, as multicultural classrooms were not very prevalent not so many years ago. Consequently, they might not have a significantly longer experience handling multicultural classrooms than their younger colleagues and perhaps mistrust themselves in this situation or do not like it as much as the earlier conditions. Another possibility is that teachers with a long experience, who have not had special preparation in addressing multicultural settings, experience burnout and an inability to cope with the new complex situation. Current research partially supports this explanation (Dubbeld, de Hoog, den Brok, & de Laat, 2019).

Quality teaching practices have a similar positive relation to self-efficacy in multicultural classrooms across the countries, and student relations as well, although the relationship is weaker. An exception is Denmark, where student relations are not significant in the model, perhaps because of a ceiling effect. Incidentally, the TALIS shows that Denmark has one of the best results in the study concerning student-teacher relations (OECD, 2019a) while also having most classes with multicultural students.

Additionally, the sum of diversity in the classroom appears to have a positive relation to self-efficacy in a multicultural setting across all the countries. The effect is not large, but it does indicate that as teachers get more diverse student groups, they master the situation better and therefore perceive that they can handle a multicultural setting better than those used to smaller diversity. The finding corresponds well with earlier research where exposure to a multicultural situation appears to increase positive attitudes towards this situation (Glock et al., 2019).

However, there are a few inconsistencies in the model, at least from the perspective of a common Nordic model. Workload stress has a negative effect only in three of the five countries. In Finland and Denmark, this effect was found as insignificant, perhaps indicating that the general workload of teachers in these two countries might be lower than in the other three (Carlgren & Klette, 2008). The same goes for

the need for professional development for diversity, which has no relation to self-efficacy in multicultural classrooms in Iceland and Finland. Only a small negative relation was observed in the other three countries. This might be because immigration has historically been mostly low in Iceland and Finland compared to the other three countries. This is considerably different from the situation reported across the OECD countries overall (OECD, 2019a) where teachers, in general, reported a significant need for training in this area.

Job satisfaction has a small positive relation in only two of the countries and disciplinary climate a negative relation only in one country although earlier research indicates that job satisfaction among teachers has a moderately strong relationship to school practices, especially those concerning handling increasing diversity (Aydan, 2016). Therefore, these last variables can perhaps be considered less important than the ones that show a clear relationship to multicultural self-efficacy in all the countries. However, research has again shown that teachers' approaches, attitudes, job satisfaction and efficacy are strongly related (Gutentag, Horenczyk, & Tatar, 2017). In any case, the results presented here shed some light on a number of important differences between the studied countries, variations which might be worthwhile examining in future investigations.

Finally, it is important to reiterate that the model presented here only explains a part of the variation in self-efficacy in multicultural classrooms—about 20% on average—and there are certainly other factors that should be considered in future studies. The explanatory value (i.e., R-square) is considerably lower in Denmark than in the other countries, indicating that the multicultural setting in Denmark is perhaps somewhat different from the other four countries. One possible explanation is that Denmark has a considerably longer history of multiculturalism in education than all the other countries. However, what is different there will be addressed another time. Still, the historical development of the school systems in the Nordic countries during the last decades has not been aligned, and possible explanatory factors could be found in the different ways these educational systems evolved over time.

4.7.1 Limitations of the Study

As mentioned before, the TALIS data has a unique pattern of missing data in the constructs and scales measuring culturally sensitive issues. This high number of missing answers might affect the results presented here, although the scaling model attempts to correct for this (OECD, 2019b). There are a few possible explanations for this high amount of missing answers (e.g., uncertain attitudes about multicultural issues or reluctance or uncertainty in discussing these issues), which, to an extent, marks and restricts the conclusions that can be drawn. In addition, the models explored in this study manage to explain only a part of the variation in self-efficacy in multicultural settings; consequently, the remaining influencing variables should also be identified and studied. Finally, it is also important to mention that the

differences between the Nordic countries underlined here are primarily because some of the variables related to the SEFE did not reach significance in all the countries. Nevertheless, as in any correlational analysis, a non-significant result does not allow us to conclude that there is no relationship present. The absence of evidence can never be evidence of absence.

4.8   Conclusions

One of the main conclusions we can take from the results is that teachers that perceive themselves competent in general and capable of handling most teaching situations in an adequate way (i.e., high teacher efficacy) will probably experience mastery and a higher perception of their own self-efficacy in a multicultural setting as well. This effect appears to strengthen with quality teaching practices. It also seems that having a shorter experience with teaching seems to be an asset, yet further investigation is needed to assess how age and experience mediate the possible views of teachers on multicultural issues in the classroom. Exposure to multicultural classes seems to generate more positive teacher attitudes towards diverse ethnic groups and probably leads to a better class climate and a better learning environment, although some type of burnout in older, more experienced teachers could also be a factor that diminishes the teachers' efficacy in such a setting (Gutentag et al., 2017).

It remains to be seen whether these results strengthen or weaken the concept of a common Nordic school model, as in this simple analysis, there do appear to be considerable differences that do not support a homogeneous pattern in these countries. However, the results still indicate that general high teacher capability and high teacher efficacy should be the essential ingredients in ensuring high equity in the Nordic classrooms. Nonetheless, it is very important to discriminate here between equity and equality; one might suspect that the Nordic countries have done a good job ensuring equality but may have fallen somewhat short in also ensuring equity. This last consideration indicates that minorities and students with different cultural backgrounds could benefit from individualised assistance and instruction. While the practice is established in some schools and classrooms, this is certainly not the case in all of them, thus hindering equity in practice. An awareness of the requirements of an equitable learning environment for all is therefore probably still something that can be improved in all the Nordic countries.

There are, of course, methodological barriers to doing an analysis of this type. The largest one is perhaps the teachers' reluctance to answer questions about multicultural issues, as evidenced by the large number of missing answers to these questions. The over 40% missing data from Iceland evidently supports this, in addition to the significantly large amount of missing answers from the other countries. Teachers do not have any problems answering factual questions about their work, about their schools or their education and professional development. Still, a significant number of them do not answer questions about diversity and multicultural

issues. These teachers certainly include those that have a considerable number of multicultural or minority students, so a further exploration of why they do not answer needs to be undertaken. It is not enough to methodologically and statistically account for missing values in the models employed, as is done with the TALIS scaling methods. These teachers' reluctance to answer questions about culturally sensitive issues and how it affects what happens in the classrooms must be better understood. Do the teachers skip questions they are unsure about, are they reluctant to answer them for some other reason or do they simply not know the answers? This must be explored further, perhaps with more concrete teacher items that are as free from value judgements as possible.

Therefore, the main conclusion of this chapter appears to be that if a teacher is competent in what they do, uses appropriate and effective methods in everyday practice, and is supported by their school and colleagues, they will most probably be able to handle a multicultural situation adequately. In addition, one could add that a good teacher most probably knows that they are doing a good job, something that most probably reflects positively on their students.

Appendices

Appendix 1

Table 4.8 Reports the Percentage of Teachers Who Have a Certain Proportion of Students of Each Type. For Example, 43% of the Teachers in Denmark Report That the First Language of Between 1 and 10% of Their Students is Different from Danish

Table 4.8 Questions about diversity in the class (percentages rounded to whole numbers)

		Denmark	Finland	Iceland	Norway	Sweden
Percentage of students whose first language is different from the language of instruction	None	37	50	23	26	15
	1–10%	43	35	53	51	44
	11–30%	12	9	18	16	21
	31–60%	5	3	3	3	9
	More than 60%	4	3	3	3	11
Percentage of students who are low academic achievers	None	3	6	5	2	3
	1–10%	43	42	45	37	42
	11–30%	39	36	36	46	34
	31–60%	10	8	8	10	12
	More than 60%	5	8	6	4	7

Table 4.8 (continued)

		Denmark	Finland	Iceland	Norway	Sweden
Percentage of students whose first language is different from the language of instruction	None	37	50	23	26	15
	1–10%	43	35	53	51	44
	11–30%	12	9	18	16	21
	31–60%	5	3	3	3	9
	More than 60%	4	3	3	3	11
Percentage of students with special needs	None	11	26	11	9	9
	1–10%	56	48	49	56	51
	11–30%	25	15	28	28	25
	31–60%	5	3	6	4	6
	More than 60%	4	7	6	3	9
Percentage of students with behavioural problems	None	32	17	13	28	19
	1–10%	50	47	52	53	54
	11–30%	14	26	26	15	18
	31–60%	2	7	6	3	4
	More than 60%	2	3	3	1	4
Percentage of students from socio-economically disadvantaged homes	None	28	24	17	26	22
	1–10%	47	48	57	54	48
	11–30%	19	21	20	15	18
	31–60%	5	5	3	3	7
	More than 60%	2	2	2	1	4
Percentage of students who are academically gifted	None	3	9	7	2	6
	1–10%	14	31	35	20	24
	11–30%	28	34	30	26	27
	31–60%	35	20	21	31	30
	More than 60%	21	6	6	21	13
Percentage of students who are immigrants or with a migrant background	None	40	51	24	26	14
	1–10%	38	34	51	50	42
	11–30%	12	10	19	17	23
	31–60%	6	4	3	4	9
	More than 60%	4	2	3	3	11
Percentage of students who are refugees	None	71	79	76	62	42
	1–10%	23	17	21	32	42
	11–30%	4	2	2	4	10
	31–60%	1	1	1	1	3
	More than 60%	1	1	0	1	3

Appendix 2

Following is the Data (Percentages of Teachers Who Provided a Certain Answer) That Goes Into the Index of Self-efficacy in Multicultural Classrooms Table 4.9.

Table 4.9　Q 45 In teaching a culturally diverse class, to what extent can you do the following?

		Denmark	Finland	Iceland	Norway	Sweden
	Not at all	1	0	3	1	1
	To some extent	14	31	36	40	31
Cope with the challenges of a multicultural classroom	Quite a bit	59	51	42	51	51
	A lot	26	17	19	8	17
Adapt my teaching to the cultural diversity of my students	Not at all	3	6	8	2	4
	To some extent	29	55	39	54	44
	Quite a bit	52	29	38	38	40
	A lot	15	10	15	6	13
Ensure that students with and without a migrant background work together	Not at all	2	6	3	1	2
	To some extent	12	25	28	31	28
	Quite a bit	48	41	40	53	45
	A lot	37	28	29	15	25
Raise awareness for cultural differences amongst students	Not at all	2	9	3	4	4
	To some extent	17	47	33	47	38
	Quite a bit	53	31	44	42	42
	A lot	28	12	20	7	16
Reduce ethnic stereotyping amongst students	Not at all	2	4	3	3	3
	To some extent	13	39	25	40	33
	Quite a bit	52	37	47	46	43
	A lot	32	21	25	12	21

Note. All percentages were rounded to whole numbers

References

Agasisti, T., Avvisati, F., Borgonovi, F., & Longobardi, S. (2018). *Academic resilience: What schools and countries do to help disadvantaged students succeed in PISA*. Paris: OECD Publishing.

Ainley, J., & Carstens, R. (2018). *Teaching and Learning International Survey (TALIS) 2018 Conceptual Framework*. (OECD Education Working Papers, No. 187). OECD Publishing.

Au, K. H., & Raphael, T. E. (2000). Equity and literacy in the next millennium. *Reading Research Quarterly, 35*(1), 170–188. https://doi.org/10.1598/rrq.35.1.12

Aydan, O. (2016). The effects of diversity management on job satisfaction and individual performance of teachers. *Educational Research and Reviews, 11*(3), 105–112. https://doi.org/10.5897/ERR2015.2573

Brante, G. (2009). Multitasking and synchronous work: Complexities in teacher work. *Teaching and Teacher Education, 25*(3), 430–436. https://doi.org/10.1016/j.tate.2008.09.015

Carlgren, I., & Klette, K. (2008). Reconstructions of Nordic teachers: Reform policies and teachers' work during the 1990s. *Scandinavian Journal of Educational Research, 52*(2), 117–133. https://doi.org/10.1080/00313830801915754

Cushner, K., & Brennan, S. (Eds.). (2007). *Intercultural student teaching: A bridge to global competence*. Lanham, MD: Rowman & Littlefield Education.

Dubbeld, A., de Hoog, N., den Brok, P., & de Laat, M. (2019). Teachers' attitudes toward multiculturalism in relation to general and diversity-related burnout. *European Education, 51*(1), 16–31. https://doi.org/10.1080/10564934.2017.1401435

Espinoza, O. (2007). Solving the equity-equality conceptual dilemma: a new model for analysis of the educational process. *Educational Research, 49*(4), 343–363. https://doi.org/10.1080/00131880701717198

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*(29), 1–9. https://doi.org/10.1016/j.learninstruc.2013.07.001

Geerlings, J., Thijs, J., & Verkuyten, M. (2019). Preaching and practicing multicultural education: Predicting students' outgroup attitudes from perceived teacher norms and perceived teacher-classmate relations. *Journal of School Psychology, 75*, 89–103. https://doi.org/10.1016/j.jsp.2019.07.003

Geiser, C. (2013). *Data analysis with MPLUS*. New York: The Guilford Press.

Glock, S., Kovacs, C., & Pit-ten Cate, I. (2019). Teachers' attitudes towards ethnic minority students: Effects of schools' cultural diversity. *British Journal of Educational Psychology, 89*(4), 616–634. https://doi.org/10.1111/bjep.12248

Gustafsson, J.-E., Nilsen, T., & Hansen, K. Y. (2016). School characteristics moderating the relation between student socioeconomic status and mathematics achievement in grade 8. Evidence from 50 countries in TIMSS 2011. *Studies in Educational Evaluation, 57*. https://doi.org/10.1016/j.stueduc.2016.09.004

Gutentag, T., Horenczyk, G., & Tatar, M. (2017). Teachers' approaches toward cultural diversity predict diversity-related burnout and self-efficacy. *Journal of Teacher Education, 69*(4), 408–419. https://doi.org/10.1177/0022487117714244

Hamre, B. K., Pianta, R. C., Downer, R. C., DeCoster, J., Mashburn, A. J., Jones, S. M., et al. (2013). Teaching through interactions: Testing a developmental framework of teacher effectiveness in over 4,000 classrooms. *The Elementary School Journal, 113*(4), 461–487. https://doi.org/10.6027/ANP2018-74210.1086/669616

IEA. (2019). *IDB Analyzer*. https://www.iea.nl/data-tools/tools#section-308

Imsen, G., Blossing, U., & Moos, L. (2017). Reshaping the Nordic education model in an era of efficiency. Changes in the comprehensive school project in Denmark, Norway, and Sweden since the millennium. *Scandinavian Journal of Educational Research, 61*(5), 568–583.

Jenkins-Martin, T. (2014). *The role of culturally responsive teaching, multicultural education and teacher self-efficacy: Am I reaching my students?* Doctoral dissertation, UC San Diego.

ProQuest ID: JenkinsMartin_ucsd_0033D_14065. Merritt ID: ark:/20775/bb49845065. https://escholarship.org/uc/item/6791s6vn

Karlsdottir, A., Norlen, G., Rispling, L., & Randall, L. (Eds.). (2018). *State of the Nordic region 2018: Immigration and integration edition*. Nordic Council of Ministers. https://doi.org/10.6027/ANP2018-742

Klette, K. (2018). Individualism and collectivism in Nordic schools: A comparative approach. In N. Witoszek & A. Midttun (Eds), *Sustainable Modernity: The Nordic Model and Beyond* (1st Edn.). London: Routledge. https://doi-org.ezproxy.uio.no/10.4324/9781315195964.

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

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

Moskal, M., & Tyrrell, N. (2015). Family migration decision-making, step-migration and separation: children's experiences in European migrant worker families. *Children's Geographies, 14*(4), 453–467. https://doi.org/10.1080/14733285.2015.1116683

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide (Version 8)*. Los Angeles, CA: Muthén & Muthén.

OECD. (2018a). *Equity in education: Breaking down barriers to social mobility*. Paris: OECD Publishing. https://tinyurl.com/y5pur9sk

OECD. (2018b). *The resilience of students with an immigrant background: Factors that shape well-being*. Paris: OECD Publishing. https://tinyurl.com/yyxarm9b

OECD. (2019a). *TALIS 2018 results (Volume I): Teachers and school leaders as lifelong learners*. Paris: OECD Publishing. https://doi.org/10.1787/1d0bc92a-en

OECD. (2019b). *Talis 2018 Technical Report*. Paris: OECD Publishing. https://tinyurl.com/yyfrwepu

Sela-Shayovitz, R., & Finkelstein, I. (2020). Self-efficacy in teaching multicultural students in academia. *International Journal of Higher Education, 9*(1), 159.

Spanierman, L. B., Euna, O., Heppner, P. P., Neville, H. A., Mobley, M., Wright, C. V., et al. (2011). The multicultural teaching competency scale: Development and initial validation. *Urban Education, 46*(3), 440–464. https://doi.org/10.1177/0042085910377442

Stronge, J. H., Ward, T. J., & Grant, L. W. (2011). What makes good teachers good? A cross-case analysis of the connection between teacher effectiveness and student achievement. *Journal of Teacher Education, 62*(4), 339–355. https://doi.org/10.1177/0022487111404241

Veenis, J. C. (2014). The Nordic education model: 'A school for all' encounters neo-liberal policy (U. Blossing, G. Imsen, and L. Moos, Eds.). *Leadership and Policy in Schools, 16*(4), 634–640. https://doi.org/10.1080/15700763.2016.1278247

Wubbels, T., den Brok, P., Veldman, I., & van Tartwijk, J. (2006). Teacher interpersonal competence for Dutch secondary multicultural classrooms. *Teachers and Teaching, 12*(4), 407–433. https://doi.org/10.1080/13450600600644269

Chapter 5
Exploring Diversity in the Relationships Between Teacher Quality and Job Satisfaction in the Nordic Countries—Insights from TALIS 2013 and 2018

Kajsa Yang Hansen, Jelena Radišić, Xin Liu, and Leah Natasha Glassow

Abstract Equity and quality are the common goals to strive for in the Nordic education systems. Yet the mechanisms through which the separate education systems approach these goals have become more diverse. The chapter provides evidence in support of the different facets of teacher quality, such as self-efficacy, as well as teacher-students relations concerning their importance for teachers' job satisfaction across the Nordic countries. Diversities, however, were also observed. The results from the TALIS 2013 model outlined two subgroups of the Nordic countries with similar mechanisms: the Norway-Sweden and the Denmark-Finland groups. No distinctive group was found in the TALIS 2018 results, producing more country-specific patterns, such as the importance of social utility value for Norway, adverse classroom composition in Sweden or teacher effective professional development positively impacting the personal and social utility values of teachers in Finland. These observed diversities and changing patterns may find their reasons in the gradually dissolved unity of the Nordic model by the different reform actions taken in recent years, such as in the example of Sweden, and in the long-term prerequisites for the teaching profession, where Finland is the country that stands out.

Keywords Teacher quality · Job satisfaction · TALIS · Nordic countries

K. Yang Hansen (✉) · L. N. Glassow
Department of Education and Special Education, University of Gothenburg, Gothenburg, Sweden
e-mail: kajsa.yang-hansen@ped.gu.se

J. Radišić
Department of Teacher Education and School Research, University of Oslo, Oslo, Norway

X. Liu
Department of Educational Studies, Ghent University, Ghent, Belgium

Access to highly qualified, skilled and experienced teachers is viewed as a crucial contributing factor in ensuring quality and equity in education and achieving optimal outcomes for each and every student (e.g., Blömeke, Olsen, & Suhl, 2016; Goe, 2007). However, when student composition and background are accounted for, the achievement gap portrays a different story. Differences between students at-risk, minority students and students in high-poverty areas and those not struggling with any of such difficulties are still noticeable, despite being high on the agenda for many education systems worldwide (Organisation for Economic Co-operation and Development [OECD], 2019a). Even in the Nordic countries, which are viewed as being among the most equitable systems in the world, what were once common patterns in student outcomes (see, e.g., Gustafsson & Blömeke, 2018) are being blurred by a trend of increasing socioeconomic achievement gaps (e.g., Chmielewski, 2019; OECD, 2019a). This could imply that different educational policies and practices may be in play across the Nordic education systems. Although pursuing educational quality and equal opportunity to all remains the common goal, differences in how schools and teachers cater to the needs of different student needs may very well exist. In the process, both schools and teachers may encounter different obstacles; in this context, retaining quality teachers who enjoy their profession and are able to answer to the needs of diverse students remains a constant need.

Teacher quality enters the spotlight every time a question is raised as to how schools ensure the optimal outcomes of their students or provide an optimal learning environment (Darling-Hammond, 2017). However, despite a long tradition in investigating the concept of teacher quality, there is no consensus regarding a comprehensive definition that gathers all its constituents. Instead, the quality includes what a teacher is, has and does, thus encompassing his or her qualifications (e.g., years of experience, specialisation, professional development), characteristics (e.g., professional self-efficacy, values and beliefs) and teaching practices (Goe, 2007). Over the years, many studies have dedicated their efforts on linking the different aspects of teacher quality to student learning outcomes (e.g., Nye, Konstantopoulos, & Hedges, 2004; Scherer & Nilsen, 2016; Zee & Koomen, 2016). However, the results related to these different facets are far from conclusive (Alvunger, Sundberg, & Wahlström, 2017), showing both direct and indirect links (e.g., the relationship between teacher qualifications and student outcomes may be mediated by instructional quality; Reimer et al., 2018). Other strands have centred their efforts on connecting teachers' perceptions of their own professions, such as job satisfaction and working environment, with the quality of student learning and outcomes, focusing primarily on diversity related to student social or migration background (Banerjee, Stearns, Moller, & Mickelson, 2017; Dicke et al., 2020).

Against this background, we investigate how different aspects of teacher quality contribute to job satisfaction. The diversity of the school environments concerning student composition and outcomes are taken into account. In particular, we wish to examine whether the determined mechanisms are alike across the Nordic countries and if the same patterns are consistent over time. The Teaching and Learning International Survey's (TALIS) data from 2013 and 2018 (OECD, 2013a, 2019b) are used for this purpose.

5.1 Diverse Faces of Teacher Quality

Research on teacher quality and effectiveness takes place at the crossroads of somewhat diverse disciplines such as econometrics, psychology and sociology (Reimer, 2019). Within each field, particular contributions may be found in understanding the idea of teacher quality and its impact on students' outcomes. With this in mind, we remain aware of the complex nature of the concept of teacher quality and observe it as the interplay between teachers' qualifications, characteristics and practices of teaching (Goe, 2007).

Along the lines of the sociology of education, Coleman et al.'s report (Coleman et al., 1966) was probably one of the most forceful push-in pieces discussing the impact of schools and teachers on students achievement and the extent to which education systems are responsible for closing the gap between different social groups. Bourdieu's (1990) ideas have also contributed to the discussion; he argued that schools and teachers, because of their direct contact with the students, only assist in the reproduction of already existing inequalities by favouring a particular habitus (i.e., students who come to classrooms with particular individual upbringings and cultural competences).

The field of econometrics, on the other hand, views the education process through the lens of input–output relations, where students' outcomes lie at the end and teachers (with their own experiences and qualifications) are situated at the beginning of the process (Hanushek, 2008). Nevertheless, within such an approach, the characteristics that seem to be the easiest to measure (e.g., teacher qualifications and experience) often contribute the least in explaining the variance in teacher quality (Hanushek & Rivkin, 2012). The education and psychology lenses take another turn, covering a myriad of topics about different aspects of teacher quality and the teaching profession. Among these, the idea of teachers' professional knowledge and practice, starting with Shulman's differentiation (Shulman, 1986; Shulman, 1987), has slowly led to a profound investigation to understand content mastery concerning the subject that one teaches (i.e., content knowledge) and how this translates into particular instructional repertoire (i.e., pedagogical content knowledge; Baumert et al., 2009; König et al., 2016). Although studies show teacher mastery does increase with years of service (Fischer et al., 2018; Nye et al., 2004), both mastery and practice have been linked to student outcomes (e.g., Baumert et al., 2009; Desimone, Smith, & Phillips, 2013; Varghese, Garwood, Bratsch-Hines, & Vernon-Feagans, 2016).

Adjacent to these investigations are the attempts in mapping out teachers' beliefs about teaching and learning (Pajares, 1992), which are seen as essential determinants of teachers' everyday practice (Buehl & Beck, 2015). Among them, constructivist beliefs (i.e., viewing students as active participants in the process of knowledge coconstruction; Berger & Lê Van, 2019) have been associated with higher levels of self-efficacy and instructional practices that are more grounded in constructivism (Nie, Tan, Liau, Lau, & Chua, 2012). A vital contribution to these ideas is found in the work of Blömeke, Gustafsson, and Shavelson (2015), who developed a

competency framework that gathers the aforementioned aspects together with those of the self-related beliefs teachers hold of the profession, their motivation and their practices.

Among the different self-related beliefs, self-efficacy (i.e., teachers' beliefs of their capability to perform particular tasks concerning teaching at a desired level of quality; Dellinger, Bobbett, Olivier, & Ellett, 2008) has been given much attention in the research on teacher quality. Consistently, teacher self-efficacy has been associated with teachers' professional practices (Vieluf, Kuenther, & van de Vijver, 2013; Zee & Koomen, 2016) and student outcomes (Caprara, Barbaranelli, Steca, & Malone, 2006; Zee & Koomen, 2016), as well as overall job satisfaction (Caprara, Barbaranelli, Borgogni, & Steca, 2003; Vieluf et al., 2013) and commitment to the profession (Chesnut & Burley, 2015; Zee & Koomen, 2016). Also, general teacher self-efficacy has been linked to student-specific efficacy, thus affecting the teacher–student relationship (Schwab, 2019). Together with self-efficacy, motivational constructs appear to hold an important position in examining the different facets of teacher quality and their mutual associations (i.e., motivation and practice; Reimer, 2019). Although different theoretical approaches may be used (e.g., Ryan & Deci, 2000; Eccles & Wigfield, 2002), conceptualisations built on Eccles' work are often used because of their value component. For example, Richardson and Watt (2006, 2016) differentiated between personal utility value (i.e., the value teachers place on the personal aspects of a teaching career) and social utility value (i.e., the utility and future outcomes of working with children and adolescents). The latter, social utility value, is seen as the consistent, positive predictor of professional engagement and job satisfaction (Torsney, Lombardi, & Ponnock, 2019).

Although mastery remains linked to teachers' experiences, similar associations are found between teacher professional development and practice (Fischer et al., 2018). It is argued, though, that more effective development programmes provide opportunities for teacher collaboration, focus on content, use affordance of the local context and offer sustained support and active participation in the context of professional learning (Akiba & Liang, 2016; Correnti, 2007; Matsumura, Garnier, & Resnick, 2010; Penuel, Fishman, Yamaguchi, & Gallagher, 2007; Roth et al., 2011).

Collaborative practices (Wang, Chen, Luo, Li, & Waxman, 2018) are also conducive to teacher job satisfaction, that is, how teachers perceive actual job outcomes compared with their desired ones (Griffith, 2004). Besides these, many factors have been linked to teacher job satisfaction (Wang, Li, Luo, & Zhang, 2019): perception of the teachers' self-efficacy (Caprara et al., 2006; Skaalvik & Skaalvik, 2014; Wang et al., 2019; Zee & Koomen, 2016), the teacher–student relationship (Collie, Shapka, & Perry, 2012; Gil-Flores, 2017; Veldman, van Tartwijk, Brekelmans, & Wubbels, 2013), the proportion of students with a lower socioeconomic status (Matsuoka, 2015; Wang et al., 2019) and the organisational culture and working conditions (Banerjee et al., 2017; Liu & Verblow, 2019). Here, the results on the relationship between teachers' demographic characteristics and job satisfaction are inconsistent. For example, some studies demonstrate a positive correlation between years of work experience and satisfaction (Ferguson, Frost, & Hall, 2012; Gil-Flores, 2017), while others provide just the opposite (e.g., Skaalvik & Skaalvik,

2009). In turn, job satisfaction is linked to teachers' occupational well-being, motivation and retention (Dicke et al., 2020), while the educational background of the teacher does not seem to be linked to job satisfaction (Wang et al., 2018).

Taken together, the different faces of teacher quality show interdependence and both direct and indirect associations with student outcomes, student composition and teacher job satisfaction (Dicke et al., 2020). Although a significant number of national-level studies have been conducted (e.g., Fischer et al., 2018), the TALIS data open a new possibility for fruitful cross-country comparisons on the subject (e.g., Liu & Verblow, 2019; Vieluf et al., 2013). At the same time, the data aid in examining the extent to which previously determined relationships hold across different countries and time points (Reimer, 2019). In this way, concrete theoretical assumptions may be tested across different contexts, and the results of such analyses may provide more nuanced insights into these relationships, thus paving the way for more attuned interventions and future investigations.

5.1.1   The Nordic Lens on Equity and Teacher Quality

In the years after World War II, the idea of equity while providing education at large was widespread across numerous education systems in Europe. The idea has accumulated momentum, and it became the foundation of the Nordic model. Under this model, schools ought to be inclusive, comprehensive, with no streaming and a smooth transition between the levels (Blossing, Imsen, & Moos, 2014; Husén, 1989; Imsen, Blossing, & Moos, 2017; Lundahl, 2016). In this model, the state is seen as a device that can provide equal opportunities to all children but not necessarily ensuring the equality of outcomes. Instead, the differences in students' outcomes were expected to be unrelated to their background or socioeconomic circumstances (Espinoza, 2007; OECD, 2018). All in all, during this time, education was seen as an essential device contributing to economic growth, minimising societal differences and promoting social mobility.

With the influx of neoliberal thinking and the economic trends at the end of the 1980s, the Nordic education systems were inevitably influenced by these concepts (Imsen et al., 2017). The neoliberal movement has led to profound debate on the sustainability of the Nordic system (Antikainen, 2006). Meanwhile, it was acknowledged that some significant differences regarding particular policies do exist across the Nordic countries (Volckmar & Wiborg, 2014). In Sweden, the policies included extensive decentralisation and deregulation reforms, the introduction of public-funded, private-run, for-profit and independent schools (Blossing & Söderström, 2014), along with severe marketisation (Lundahl, 2016). These policies left their toll, leading Sweden to lag behind in rankings of the most equitable school systems of the Nordic countries (Imsen et al., 2017). Until now, the Norwegian education policy has withstood its restrictive stance on the privatisation of the school market (Imsen & Volckmar, 2014). Still, it is not immune to accountability practices, which have been gradually introduced (Imsen et al., 2017). In Denmark, the competitive

discourse has become stronger (Rasmussen & Moos, 2014), while in Finland, polarisation between the schools became evident both in the equity of provision (i.e., the unequal distribution of municipality funds) and in the socioeconomic backgrounds of the students (Ahonen, 2014).

The global push towards educational measurement and comparison since the 1990s has introduced more visible accountability practices in all the Nordic countries (Wallenius, Juvonen, Hansen, & Varjo, 2018; Wollscheid & Opheim, 2016). The establishment of quality assurance systems has produced more extensive documentation of the work both the schools and teachers do (Imsen & Volckmar, 2014). This has profoundly influenced how the teachers view their profession, and what they do has become more regulated and scrutinised. Comparisons across the Nordic countries indicate the job satisfaction of teachers in Sweden is the lowest among their Nordic colleagues (Taajamo, 2016), while teachers in Finland strongly believe their profession is valued in society (Reimer, 2019). Overall, substantial variations across the Nordic countries may be found regarding teachers' beliefs of the profession, perceptions of their instructional practices and perceived appreciation. Involvement in different types of professional development activities remains a challenge. The opportunities offered, as well as their variety, do not seem to provide enough of an incentive to the teachers (Taajamo, 2016), although Finland stands out both in the prerequisites for the teaching profession (Aspfors, Hansen, & Ray, 2014) and the long tradition in linking practice with research (Wollscheid & Opheim, 2016).

The ideas of the Nordic model remain the backbone in understanding the purpose of education and the role teachers may have in the education process. However, to fully comprehend diversity and its effects on the potential mechanisms that affect teacher quality, we need to take into account the ongoing processes in each of the Nordic systems, as well as how these may affect the strength and direction of the relationship between teacher quality, job satisfaction and educational outcomes.

5.2   The Present Study

Documenting a comprehensive overview of all the relevant aspects pertinent to teacher quality and views of the profession is beyond the scope of this chapter. Nonetheless, the literature review grounds our work and showcases the line of thinking that guided us in the current analyses. At the same time, we use the affordances of the TALIS data in examining the same type of mechanisms (see Fig. 5.1) across four Nordic countries (i.e., Denmark. Finland, Norway and Sweden) in both 2013 and 2018. In this way, we are also able to follow the extent to which associations in the data are relevant to particular contexts or across them.

In this investigation, we focus on the distinctive mechanisms that are found among several major aspects of teacher quality (i.e., teacher qualifications, professional development, beliefs, practices, self-efficacy) in an attempt to understand diversity in the relationships among them and how each contributes to teachers' job satisfaction. In line with the theoretical review and empirical background presented,
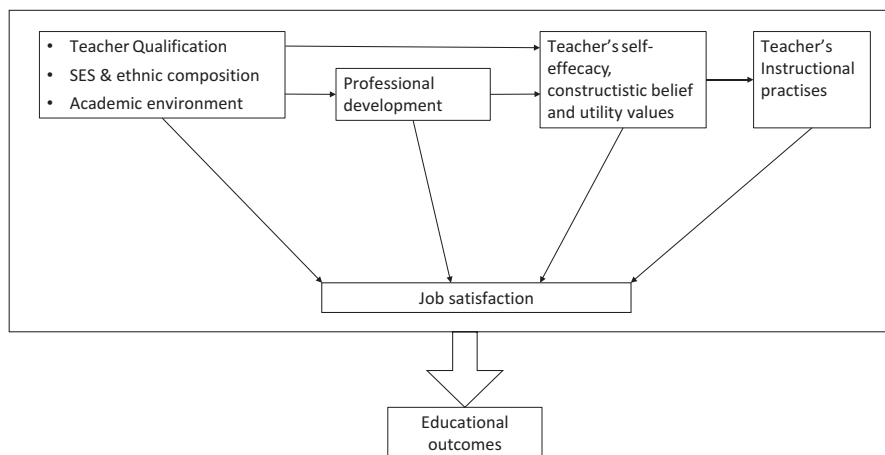
Fig. 5.1 A hypothesis model of the mechanisms among teacher quality, working environment, professional development, self-efficacy and beliefs, teaching practices and job satisfaction

we hypothesise job satisfaction is influenced by the perception of one's self-efficacy (Caprara et al., 2006; Skaalvik & Skaalvik, 2014; Vieluf et al., 2013; Zee & Koomen, 2016; Wang et al., 2019), professional development and collaborative practices (Fischer et al., 2018; Wang et al., 2018) and utility values (Torsney et al., 2019). The influence of teacher qualifications, here combining years of service and education (e.g., Gil-Flores, 2017; Wang et al., 2018), the teacher–student relationship (Collie et al., 2012; Gil-Flores, 2017; Veldman et al., 2013), academic environment in the classroom and the proportion of students with a lower socioeconomic status (SES) in the classroom (Matsuoka, 2015; Wang et al., 2019) also are included in the model. Figure 5.1 shows the hypothetical model that was tested using both TALIS 2013 and 2018 data. In this way, both the direct and mediating effects can be examined. The relationships are tested separately for each of the Nordic countries.

Within the last two TALIS cycles, somewhat differing information concerning teachers' beliefs, values and instructional practices has been collected. Therefore, the hypothesised model is operationalised in a slightly different way across the two. Because TALIS does not collect information on educational outcomes, only part of the hypothesis model in the rectangular frame is tested. However, the model controls for teachers' perceptions of their classroom academic and demographic environments, which are made up of the proportions of students with special needs and those with disadvantaged SES and migration backgrounds.

5.3 Method

5.3.1 Participants

Four Nordic countries (i.e., Denmark, Finland, Norway and Sweden) joined both the TALIS 2013 and 2018 cycles. Information about the samples used in the analyses is provided in Table 5.1 and is displayed by the countries analysed. Additional technical details on the sample may be found in the TALIS technical reports (OECD, 2013b, 2019c).

5.3.2 Variables

The two consecutive TALIS cycles gathered various information about different aspects of the teaching profession and their related characteristics and practice. The following variables were included in the 2013 model.

Teacher's professional self-efficacy is a composite variable encompassing efficacy in classroom management, efficacy in instruction and efficacy in student engagement, gathering 12 items in total. Each item is on a four-point scale with the response categories ranging from 'not at all' to 'a lot'.

The teacher's job satisfaction is made up of two subscales describing their satisfaction with the current work environment and with the teaching profession. Both subscales amount to eight four-point items, with the response alternatives ranging from 'strongly disagree' to 'strongly agree'.

Teacher–student relations is an index measure set on a four-point scale with four items. The response categories include a range from 'strongly disagree' to 'strongly agree' on items focusing on aspects such as whether the teachers and students usually get on well with each other.

The *index of constructivist beliefs* was measured by four items using a four-point scale, with response categories ranging from 'strongly disagree' to 'strongly agree'. The items included inquiring about the perceptions of the role of teachers in facilitating students' inquiry or the best ways students may be learning.

Teacher's effective professional development is a four-item composite score set on a four-point response scale ranging from 'not in any activities' to 'yes, in all

Table 5.1 The number of teachers and schools in the Nordic countries in TALIS 2013 and 2018

	TALIS 2013		TALIS 2018	
	Teachers	Schools	Teachers	Schools
Denmark	1649	148	1853	141
Finland	2739	146	2761	148
Norway	2981	145	3802	185
Sweden	3193	186	2488	180
Total	10562	625	10904	654

activities'. The compound construct focuses on the different opportunities for active learning methods or collaborative learning activities or research with other teachers.

Teacher collaboration is an index measure focusing on the opportunities for collaboration with different stakeholders or activities (e.g., teach jointly as a team in the same class). The six-item response options of the index range are from 'never' to 'once a week or more'.

Teacher qualification is a principle component factor score comprised of the highest level of teacher formal education, completion of teacher training programme and years of work experience.

Classroom composition of SES and migration is a principle component factor score of the percentage students whose first language is not the native language and who are from socioeconomically disadvantaged homes.

Classroom academic environment is a principal component factor score of the percentage of students with special needs, low achievement, behavioural problems and among the less gifted. Table 5.2 provides more details of the constructs used, including where these constructs differ between the 2013 and 2018 cycles.

Table 5.2 Descriptive statistics of the variables in the analysis for TALIS 2013 and TALIS 2018

	Demark		Finland		Norway		Sweden	
TALIS 2013	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Teacher's professional self-efficacy	13.33	1.39	11.86	1.85	11.38	1.47	12.23	1.63
Teacher's job satisfaction	12.60	1.83	12.35	1.78	12.25	1.70	11.40	1.96
Teacher–student relations	14.73	1.79	13.47	1.93	14.11	1.98	14.04	1.81
Teacher's constructive beliefs	13.42	1.74	12.69	1.58	11.72	1.05	11.06	1.37
Teacher's effective professional development	9.23	2.04	8.39	1.70	7.57	1.63	8.38	1.98
Teacher collaboration	10.94	1.63	9.26	1.62	10.30	1.67	10.78	1.48
Teacher qualification	0.06	0.94	0.04	0.92	0.08	0.85	−0.13	1.18
classroom SES and migration composition	−0.15	1.03	−0.19	0.89	0.01	0.86	0.22	1.13
Classroom academic environment	−0.22	1.08	0.11	1.07	0.04	0.82	−0.02	1.02
TALIS 2018	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Teacher's professional self-efficacy	11.19	1.81	11.08	1.74	11.15	1.71	11.17	1.88
Teacher's personal utility values	10.37	2.25	10.43	2.16	9.96	1.95	9.92	2.10
Teacher's social utility values	12.10	2.01	12.01	2.44	11.75	1.68	11.68	2.25
Teacher's job satisfaction	12.08	2.08	11.96	2.18	11.42	2.07	11.98	2.11
Teacher's effective professional development	10.07	1.37	11.45	2.23	9.73	2.44	10.10	2.34
Teaching practice	11.41	2.07	11.34	2.10	11.56	1.48	11.75	1.97
Teacher–student relations	12.75	2.33	12.52	2.51	10.88	2.93	11.87	2.71
Teacher qualification	0.04	0.95	0.00	0.94	0.00	1.00	−0.18	1.13
Classroom SES and migration composition	1.28	3.57	1.21	3.67	4.79	4.97	3.27	4.37
Classroom academic environment	1.33	4.10	1.51	4.28	4.57	6.05	1.74	4.18

In the 2018 model, three new variables were added to the list. These include teacher's personal utility values, teacher's social utility values and teaching practice. Teacher's constructive beliefs were excluded from the variable list.

Teacher's personal utility values is a four-item composite related to the different aspects teachers value to be part of the teaching profession (e.g., teaching offers a steady career path or teaching provides a reliable income). The scale is set on a four-point scale, with the response categories ranging from 'Not important at all' to 'Of high importance'.

Teacher's social utility values also relate to the different aspects teachers may value relative to the teaching profession but from the perspective of the immediate environment and community (e.g., teaching allowed me to benefit the socially disadvantaged). The scale is comprised of four items and set on a four-point range, with response categories ranging from 'Not important at all' to 'Of high importance'.

The final composite scale, *teaching practice,* comprises subscales on the clarity of instruction, cognitive activation and classroom management, with 12 items in total. Response options include the following: 'Never or almost never', 'Occasionally', 'Frequently' and 'Always'. All variables were used and aligned with the TALIS technical manuals (OECD, 2013b, 2019b). For more information on each scale, see the TALIS technical reports (OECD, 2013b, 2019b).

5.3.3   *Analytical Method and Data Analyses*

All analyses were performed in Mplus (Muthén & Muthén, 1998–2017). The FIML option was used to handle missing data. In the current study, a path modelling approach was adopted to examine the mechanism through which teacher characteristics, professional belief and values and teaching practices may affect their job satisfaction. In the conditional model information on teacher experience and specialisation, student socioeconomic and immigration composition in the classroom and classroom academic environment were accounted for. One of the advantages of a path analysis is its ability to estimate the direct effects of an independent variable on a dependent variable, along with being able to estimate an indirect effect from the same independent variable through a mediator on the dependent variable (e.g., Wolfe, 1980).

The path model was specified in light of prior research evidence. We provide a simplified illustration to demonstrate this principle. In Model A, for example, teacher qualification affects the teacher's job satisfaction, which is a direct and total effect with a strength of *a*. However, according to the specified model and prior evidence, teacher qualifications may have an effect on teacher effective professional development; thus, in turn, it can impact teacher's job satisfaction (Model B). In Model B, the total effect of teacher qualification is decomposed into a direct effect from teacher qualification on their job satisfaction *a*' and an indirect effect. The strength of the latter is a product of two direct effects, namely, a direct effect of teacher qualification on teacher's professional development *b* and teacher's

professional development on their job satisfaction *c*. Thus, the total effect of teacher qualification on their job satisfaction is now the sum of the direct and indirect effects, *a' + bc* (e.g., Baron & Kenny, 1986).

In Model B, the mediating effect *bc* may account for part of the total effect between teacher qualification and job satisfaction *a* in Model A, making the direct effect *a'* in Model B smaller than *a*. In a particular situation of full mediation, the mediating effect *bc* may be overlapping entirely with the total effect *a* in Model A. In this case, the direct effect *a'* in Model B is spurious.

5.4 Results

In this section, we focus on the relationship between teachers' professional self-efficacy, constructive beliefs, practices and job satisfaction in TALIS. We explore the relationship separately for TALIS 2013 and 2018, which is followed by a short comparison between the 2013 and 2018 results.

Both the direct effects and indirect effects are shown. The hypothesis model in Fig. 5.2 is used as a common point of departure for all countries and each TALIS cycle. Because there are two types of effects (i.e., direct effect and indirect effect) in the path analysis, the effect of an independent variable on the dependent variable needs to consider both effect types. The operational models for all the Nordic countries in both TALIS cycles are saturated, meaning that no relation between any two factors was left out. However, when presenting the parameter estimates in the path diagrams, only the statistically significant paths are included. Full estimations are provided in the supplementary material. We start by observing these mechanisms across the 2013 cycle for the four studied Nordic countries.
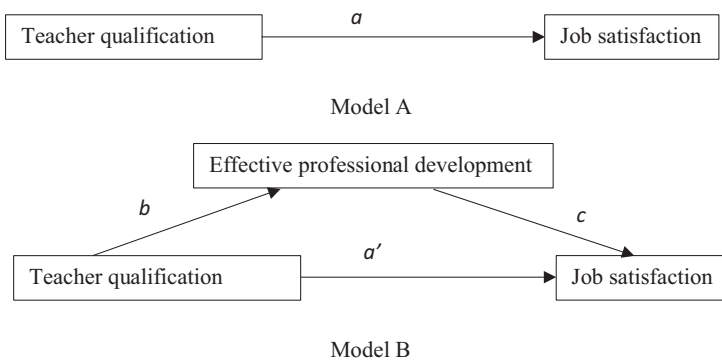
Fig. 5.2 Direct and indirect effects between teacher qualifications, teach effective professional development and teacher job satisfaction

5.4.1 Diverse Mechanisms in the TALIS 2013 Data

In Sweden, all the teacher and teaching-related factors in the operationalised model for the 2013 TALIS data have a significant impact on teachers' job satisfaction. The most substantial total effect on job satisfaction (TJOBSATS) came from the teachers' self-efficacy (TSELEFFS, 0.21), where the direct effect was 0.15 and the indirect effect was 0.07. The overall effect of teachers' effective professional development (TEFFPROS) on their job satisfaction was about 0.15, of which 0.09 was the direct effect and 0.07 was the indirect effect. Albeit statistically significant, the effects of teacher qualification (TQ) and constructivist belief (TCONSBS) were rather small. Significant direct effects of classroom SES and migration composition (SESMIG), teacher–student relations (TSCTSTUDS) and teacher collaboration (TCCOLLS) have also been observed in Sweden, at −0.07, 0.19 and 0.15, respectively. As shown in Fig. 5.3 (top diagram), teachers' qualifications (TQ), the teachers' effective professional development (TEFFPROS), self-efficacy (TSELEFFS) and teachers' constructivist beliefs (TCONSBS) have significantly affected the teachers' job satisfaction both directly and indirectly.

The 2013 model for Norway (Fig. 5.3, lower diagram) shows significant effects, both direct and indirect, from teachers' self-efficacy (TSELEFFS) and effective in-service professional development (TEFFPROS) on their job satisfaction. The highest total effect was 0.28 from self-efficacy. When decomposed, 0.16 went to the direct effect, and 0.09 was the indirect effect. Effective in-service professional development (TEFFPROS) was found to have a substantial effect on teachers' job satisfaction (TJOBSATS), 0.16 in total. This value was contributed to equally from both direct and indirect effect, each being 0.08. Teachers' professional collaboration (TCCOLLS) held the most substantial impact on teachers' job satisfaction at 0.13. Teacher–student relations also has a considerable effect (0.31). Only a small negative direct effect (−0.06) was found for classroom academic environment (CLACDEM) on teachers' job satisfaction.

In Finland, for the contextual factors, classroom SES and ethnic composition (SESMIG) and the classroom academic environment (CLACDEN), only small negative effects were found: −0.07 and − 0.05 respectively. A little indirect effect was observed between teacher effective professional development (TEFFPROS) and job satisfaction (TJOBSATS). Teachers' professional self-efficacy (TSELEFFS) affected their job satisfaction both directly (0.20) and indirectly (0.09). No significant effect was found for the remaining factors in the model (see Fig. 5.4).

In the case of Denmark, Fig. 5.4 indicates that TQ directly influenced the teachers' job satisfaction (TJOBSATS, −0.06). It also significantly mediated the effect of teacher professional self-efficacy (TSELEFFS, 0.05) and teacher–student relations (TSCTSTUDS, 0.01); self-efficacy and teacher professional collaboration (TEFFPROS, 0.01) also affected teachers' job satisfaction (TJOBSATS). However, these indirect effects were rather small. The most substantial direct effect was found in teacher–student relations and their job satisfaction (0.28). Teachers' professional collaboration (TCCOLLS) also was significantly related to their job satisfaction
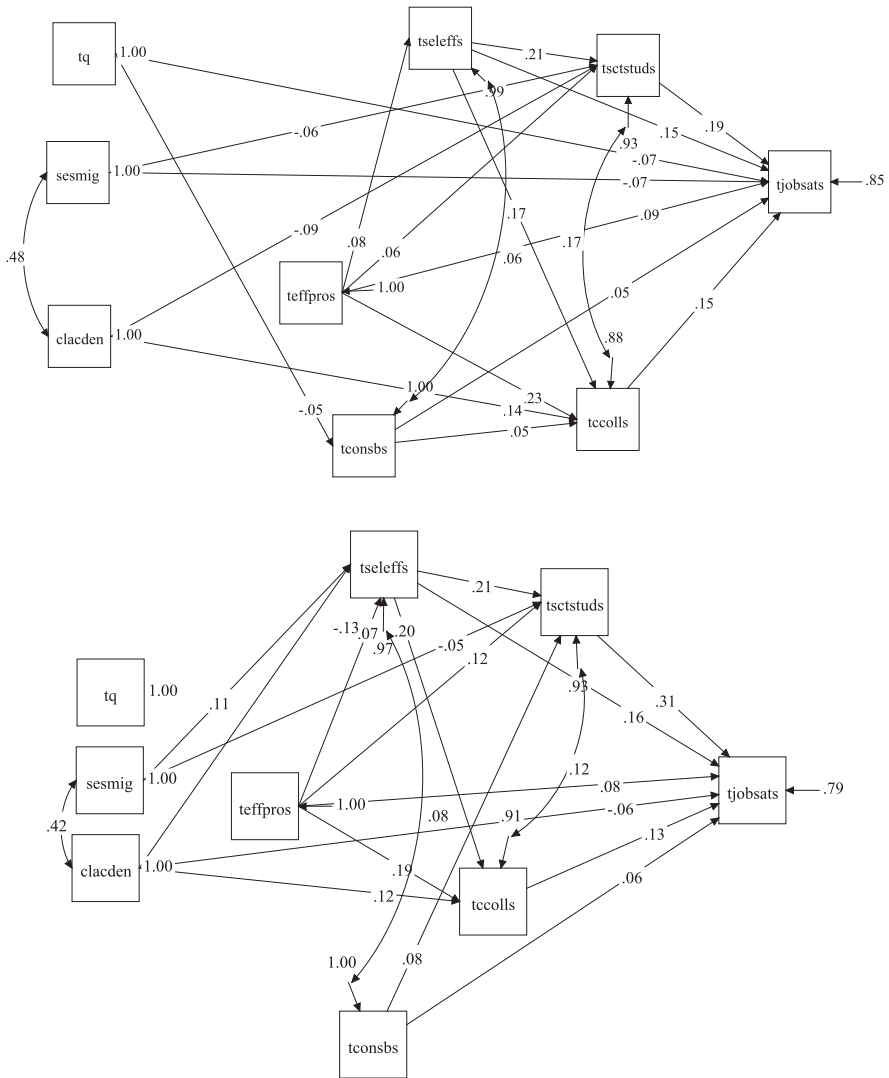
Fig. 5.3 Path diagram for Sweden (top) and Norway (down) in TALIS 2013. (Note: *TSELEFFS* teacher's professional self-efficacy, *TJOBSATS* teacher's job satisfaction, *TSCTSTUDS* teacher–student relations, *TCONSBS* teacher's constructive beliefs, *TEFFPROS* teacher's effective professional development, *TCCOLLS* teacher collaboration, *TQ* teacher qualification, *SESMIG* classroom SES and migration composition, *CLACDEM* classroom academic environment. Only significant paths are shown)

(0.07). For the classroom contextual factors, the SES-ethnic composition was only indirectly related to teachers' job satisfaction (TJOBSATS) through teacher–student relations (TSCTSTUDS, −0.06) and constructivist belief (TCONSBS) and teacher–student relations (0.01). Classroom academic environment was found to have both
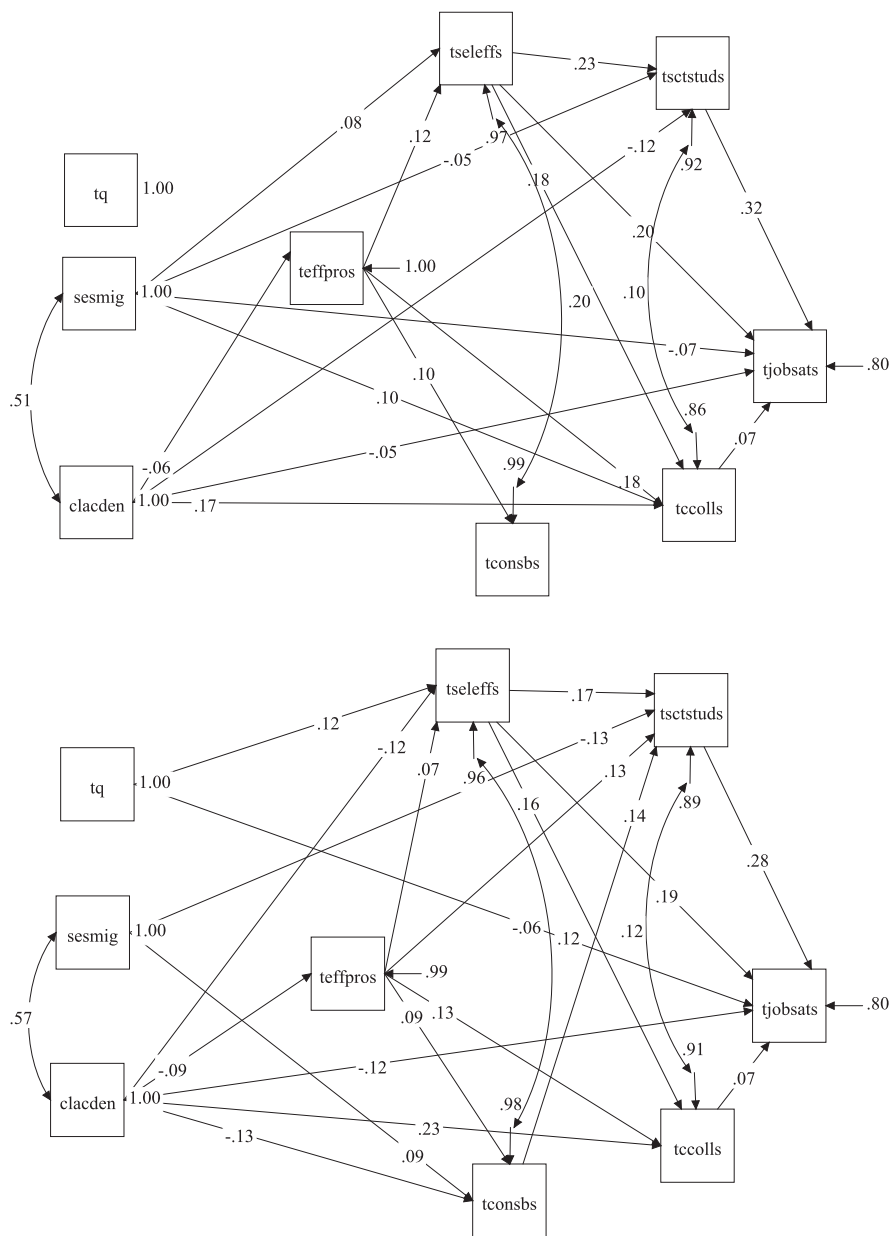
Fig. 5.4 Path diagram for Finland (top) and Denmark (down) in TALIS 2013. (Note: *TSELEFFS* teacher's professional self-efficacy, *TJOBSATS* teacher's job satisfaction, *TSCTSTUDS* teacher–student relations, *TCONSBS* teacher's constructive beliefs, *TEFFPROS* teacher's effective professional development, *TCCOLLS* teacher collaboration, *TQ* teacher qualification, *SESMIG* classroom SES and migration composition, *CLACDEM* classroom academic environment. Only significant paths are shown)

a significant direct and indirect effect on teachers' job satisfaction (TJOBSATS), at −0.12 and − 0.04, respectively. For teachers' characteristics, professional self-efficacy (TSELEFFS) affected job satisfaction both directly (0.19) and indirectly through teacher–student relations and professional collaborations (0.09). Only a small indirect effect was observed from the teachers' constructivist beliefs on their job satisfaction via teacher–student relations (0.04).

Overall, different mechanisms were found across the four Nordic countries in the TALIS 2013 survey. However, some common patterns also were observed. Among them, the teachers' professional self-efficacy was one of the most significant factors affecting teachers' job satisfaction both directly and indirectly via teacher–student relations. Also, the teachers' professional development mediates the effects of their professional self-efficacy and student–teacher relations, which, in turn, affects their job satisfaction. The strongest effect on teachers' job satisfaction came from the teacher–student relations. Teachers' professional collaborations also were found to have a substantial effect, higher in Sweden and Norway than those in Denmark and Finland do. Given these common features, the four Nordic countries can be separated into two groups with similar mechanisms: Norway-Sweden group and Denmark-Finland group.

From Table 5.3, the path model can explain equally the amount of variance in teacher's job satisfaction in Denmark, Finland and Norway, at around 20%, while it performed less well in Sweden (15%). Different amounts of explained variances in other teacher-related factors also indicate the different pathways through which these factors are mediating and affecting job satisfaction. The variation in teachers' effective professional development cannot be attributed to any of the factors in the model in all the Nordic countries and neither can the variance of teachers' constructive beliefs in Norway and Sweden. Please see the supplementary material, appendices A–C for the detailed specification on all the direct, indirect and total effects.

Table 5.3 Explained variance of all the endogenous variables in the path models of the four Nordic countries in TALIS 2013

	Denmark		Finland		Norway		Sweden	
Variables about teacher	R^2 (%)	z-value	R^2 (%)	z-value	R^2 (%)	z-value	R^2 (%)	z-value
Job satisfaction	20	9.37	20	12.30	21	10.05	15	11.25
Teacher–student relations	11	5.47	8	6.97	7	4.96	7	5.84
Professional collaborations	9	4.88	14	8.73	9	3.21	12	9.64
Constructivist belief	2	2.22	1	2.82	0	.90	0	1.57
Professional self-efficacy	4	3.66	3	3.42	3	3.13	1	2.11
Effective professional development	1	1.10	0	1.46	0	.48	0	.72

5.4.2 Diverse Mechanisms in the TALIS 2018 Data

Following the same assumptions grounded in prior research, the hypothesis model was operationalised with the available factors in TALIS 2018. The analyses indicated a high correlation, that is, over 0.90, between the classroom SES-ethnic composition and the academic environment for all Nordic countries. Therefore, only the classroom SES-ethnic composition was kept in the operationalised model. Again, we observe the results for each country separately

In Sweden, the highest direct effect on teacher's job satisfaction (T3JOBSA) was from teacher–student relations (T3STUD, 0.42). The disadvantaged SES-ethnic classroom composition (SESMIG) also had a relatively high direct effect (0.29). The challenging classroom composition may make Swedish teachers feel a sense of fulfilment from their work, thus contributing to their satisfaction. However, it strongly affected teacher–student relations negatively (−0.67), resulting in the mediation effect on job satisfaction (T3JOBSA) via teacher–student relations to be negative (−0.28). It is interesting to observe that teachers' effective professional development (T3EFFPD) was positively related to teachers' personal utility motivation (i.e., teacher profession offers a steady career path, a reliable income/secure job and good schedule, T3PERUT) by 0.35. However, it was found to have no impact on teachers' social utility motivation (i.e., teacher's belief that teaching allows them to influence the development of children and young people, helping disadvantaged and contributing to society, T3SOCUT). Teachers' professional self-efficacy (T3SELF) indirectly affected teachers' job satisfaction (T3JOBSA) through student–teacher relations (0.31), but no significant direct effect was found. No significant direct effect was found for teaching practices (T3TPRA).

As shown in Fig. 5.5, the only significant direct effects on teachers' job satisfaction (T3JOBSA) in Norway were from teacher–student relations (T3STUD) and teachers' social utility motivation (T3SOCUT) at 0.28 and 0.14, respectively. We also observed significant indirect effects of classroom disadvantaged SES-ethnic composition (SESMIG) on teachers' job satisfaction via teacher–student relations (−0.11), teachers' social utility motivation to teach (−0.02) and teachers' effective professional development and teacher–student relations (−0.04). Classroom disadvantaged SES-ethnic composition (SESMIG) directly affected all other teacher-related factors except for the teachers' practice (T3TPRA). The highest direct effect was SESMIG on teachers' professional self-efficacy (−0.80), followed by SESMIG effect on teacher–student relations (−0.41). The direct effects of SESMIG on teachers' personal and social utility motivations were also substantial at −0.20 and − 0.16, respectively. However, no relationship was found between classroom disadvantaged SES-ethnic composition (SESMIG) and job satisfaction (T3JOBSA). Effective professional development (T3EFFPD) positively affected teacher–student relations (T3STUD, 0.34), with no significant mediation effect on job satisfaction (T3JOBSA). It is worth noticing that Norwegian teachers' qualifications positively (TQ) affected their teaching practices (T3TPRA, 0.11).
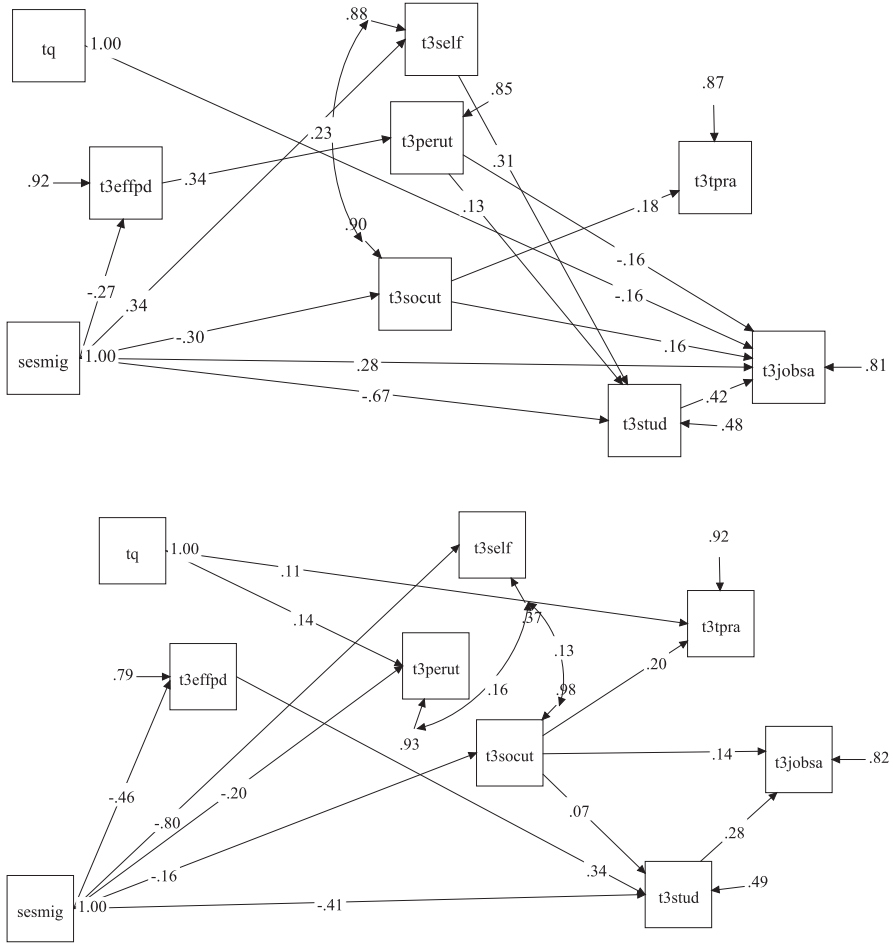
Fig. 5.5 Path diagram for Sweden (top) and Norway (down) in TALIS 2018. (Note: *T3EFFPD* effective professional development, *T3PERUT* personal utility value, *T3SOCUT* social utility value, *T3STUD* teacher–student relations, *T3TPRA* teaching practices, *T3JOBSA* job satisfaction, *T3SELF* teacher self-efficacy, *TQ* teacher qualification, *SESMIG* classroom students SES and migration background composition. Only significant paths are shown)

In the case of Finland, the most substantial direct effect on teachers' job satisfaction (T3JOBSA) was from teacher–student relations (T3STUD, 0.21), and the effect of teachers' social utility motivation (T3SOCUT) to teach was also substantial (0.16). Disadvantaged classroom SES and ethnic composition (SESMIG) was found to have rather strong negative influences on teachers' self-efficacy (T3SELF, −0.58), effective professional development (T3EFFPD, −0.44), teacher–student relations (T3STUD, −0.50) and personal utility motivation to teach (T3PERUT, −0.18) (Fig. 5.6).
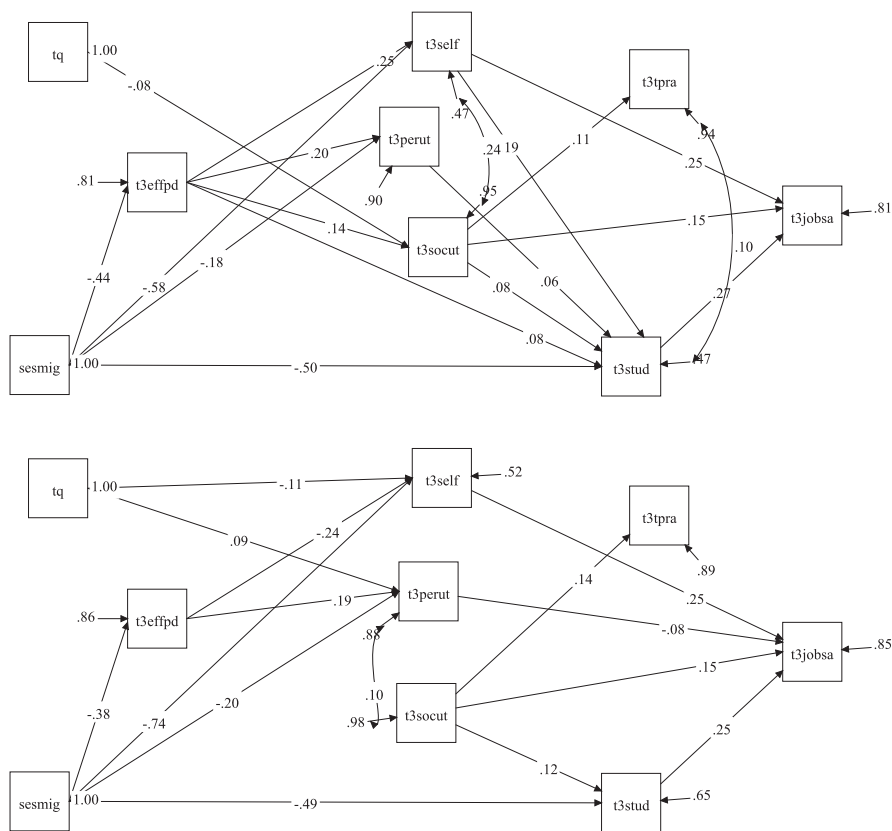
Fig. 5.6 Path diagram for Finland (top) and Denmark (down) in TALIS 2018. (Note: *T3EFFPD* effective professional development, *T3PERUT* personal utility value, *T3SOCUT* social utility value, *T3STUD* teacher–student relations, *T3TPRA* teaching practices, *T3JOBSA* job satisfaction, *T3SELF* teacher self-efficacy, *TQ* teacher qualification, *SESMIG* classroom students SES and migration background composition. Only significant paths are shown)

The analysis also revealed a total negative indirect effect of perceived disadvantaged classroom SES-ethnic composition (SESMIG) on teacher job satisfaction (T3JOBSA) by −0.36. The most significant indirect effect of SESMIG on T3JOBSA was via the teachers' professional self-efficacy (−0.15) and student–teacher relations (−0.14). Other indirect effects between the two (SESMIG on T3JOBSA) were via teacher practices (T3TPRA), teacher social utility motivation (T3SOCUT), effective professional development (T3EFFPD) and teacher–student relations (T3STUD). However, these indirect effects, despite their significance, were minimal. TQ was also observed to have small indirect effects on job satisfaction (T3JOBSA) through teacher–student relations (T3STUD) and their social utility motivation (T3SOCUT).

In Denmark, teachers' professional self-efficacy (T3SELF) and their relation with students (T3STUD) have a significant and positive direct impact on their job

satisfaction, both being at 0.25. Teachers' personal utility motivation to teach (T3PERUT) was found to have a small negative effect (−0.08), while teachers' social utility motivation to teach (T3SOCUT) had a positive impact (0.15), almost twice as large as their personal utility motivation. No significant effects were found for teachers' effective professional development (T3EFFPD), teaching practices (T3TPRA), TQ and classroom disadvantaged SES-ethnic composition (SESMIG) on their job satisfaction (T3JOBSA).

Classroom socioeconomic and ethnic composition (SESMIG) affected most of the teacher-related factors negatively. The highest effect was found on teachers' professional self-efficacy (−0.75), followed by the effect on teacher–student relations (−0.49). The effects of the classroom socioeconomic and ethnic composition (SESMIG) on teachers' effective professional development (T3EFFPD) and teachers' personal utility motivation (T3PERUT) were − 0.20 and − 0.38, respectively. The classroom socioeconomic and ethnic composition (SESMIG) also was found to have a significant indirect effect on teachers' job satisfaction (−0.27).

A common feature revealed in the analysis of the TALIS 2018 data is the positive direct effects of teacher–student relations and teachers' social utility motivation to teach on job satisfaction in all the Nordic countries. Sweden held the most substantial impact of the teacher–student relationship on job satisfaction (0.42), and the effect of the rest of the Nordic countries was very similar, around 0.25. The direct effect of social utility motivation, on the other hand, as about the same level in all the four Nordic countries, approximately 0.15. We also found a positive impact of teachers' social utility motivation to teach on their teaching practices, with Norway and Sweden being higher than those of Denmark and Finland. Strong adverse effects were observed of the disadvantaged classroom SES-ethnic composition and teacher–student relations, ranging from −0.67 in Sweden to −0.41 in Norway and on teachers' effective professional development, ranging from −0.46 in Norway to −0.27 in Sweden. SESMIG significantly affected teacher's professional self-efficacy. However, the effect was highly negative in Norway, Denmark and Finland but positive in Sweden.

Diversities in the mechanisms, however, were also revealed in the analysis. For example, Swedish teacher's job satisfaction and effective professional development were affected positively by the disadvantaged classroom SES-ethnic composition. In contrast, teacher qualification affected their job satisfaction negatively. In Denmark, the teachers' effective professional development positively impacted their personal utility motivation and was negatively related to their professional self-efficacy. However, the opposite or no effect was found in other countries. Please see the supplementary material, appendices D–F for the detailed specification on all the direct, indirect and total effects.

Table 5.4 shows the explained variance for all the dependent variables in the model. For the outcome variable job satisfaction, the proposed mechanism in the model was not fully reflected in the operationalised model in TALIS 2018. On average, around 18% variance in job satisfaction was accounted for by the model in the Nordic countries. This may imply that additional factors and mechanisms need to be considered for in teachers' job satisfaction. The proposed model explained a large

Table 5.4 Explained variance of all the endogenous variables in the path models of the four Nordic countries in TALIS 2018

Variables about teacher	Denmark		Finland		Norway		Sweden	
	R^2 (%)	z	R^2 (%)	z	R^2 (%)	z	R^2 (%)	z
Job satisfaction	15	2.35	19	4.73	18	2.64	19	4.22
Teacher–student relations	35	3.20	54	8.95	51	8.48	52	10.79
Teacher's self-efficacy	48	3.19	53	8.57	63	8.73	12	1.06
Personal utility motivation to teach	12	2.75	11	2.17	7	2.59	15	2.77
Social utility motivation to teach	2	0.69	5	1.98	2	1.48	10	1.85
Effective professional development	14	1.40	19	2.75	21	3.75	8	1.27
Teaching practices	11	1.21	6	2.21	8	1.69	13	2.53

amount of the variance in teachers' professional self-efficacy in Norway (63%), Finland (54%) and Denmark (48%). However, in Sweden, only 12% of the differences can be attributed to the factors in the model, which was not significant. The same pattern, but to a much less extent, was found for teacher's effective professional development. Here, the amount of explained variance in Denmark and Sweden was not significant. The model explained a significant amount of the variances in all teacher-related variables in Finland. For the explained variance in teacher's social utility motivation to teach and teaching practices, Denmark and Norway had a small and nonsignificant amount.

Comparing the results from TALIS 2013 and 2018, the single factor that consistently affects teachers' job satisfaction is teacher–student relations. This effect is the largest in all four Nordic countries and TALIS cycles. However, teacher–student relations was significantly related to classroom SES-ethnic composition and teachers' professional self-efficacy. Furthermore, the TALIS 2013 analysis revealed the importance of teachers' professional self-efficacy for most of the other teacher-related factors in all Nordic countries. However, this is not the case in TALIS 2018. The hypothesised model seems to be proved as true by the TALIS 2013 data in all the four Nordic countries, yet it worked less well for the TALIS 2018 data, especially for Denmark and Sweden.

5.4.3 Discussion

The idea of teacher quality and how teachers matter to students' well-being and outcomes has provoked a mass investigation that has spread over several decades and across disciplines. Although no unified definitions have been found, the concept of teacher quality embraces teachers' qualification, characteristics and the practices of teaching (Goe, 2007). Over the years, different aspects of teacher quality have been investigated, showing mutual interdependence (e.g., Liu & Verblow, 2019; Fischer et al., 2018; Zee & Koomen, 2016; Wang et al., 2019) and a link with student learning and outcomes (e.g., Caprara et al., 2006; Nye et al., 2004; Zee &

Koomen, 2016) and job satisfaction (e.g. Caprara et al., 2003; Vieluf et al., 2013). The latter has especially come to the fore in an era when retaining quality teachers has become challenging when teachers are met with more and more demands to adapt their own teaching relative to the needs of students with diverse social or migration backgrounds (Banerjee et al., 2017; Dicke et al., 2020). With this in mind, we investigated how the different aspects of teacher quality contribute to job satisfaction in connection to varied school environments relative to student composition and outcomes. In particular, we examined whether the mechanisms hold across the Nordic countries and if the same patterns are consistent over time (Reimer, 2019).

The comparative stance represented an essential facet of the current study. Analysing four Nordic countries has allowed us to observe systems in which the schools ought to be inclusive and comprehensive, while the teachers are seen as essential contributors in providing equal opportunities to all children (Blossing et al., 2014; Imsen et al., 2017; Lundahl, 2016). Thus, teacher quality is understood as instrumental in balancing equity across the education system. However, over the last two decades, even in the Nordic countries, an influx of accountability measures and marketing practices has introduced some changes, influencing how teachers view their profession and job satisfaction (Reimer, 2019; Taajamo, 2016). Capturing these factors were our focus.

Across the countries, we have observed that both uniform and diverse patterns were found relative to the relationship between teacher quality and job satisfaction. Comparing the results from TALIS 2013 and 2018, the single factor that consistently affects teachers' job satisfaction was the teacher–student relations. The effect on this was the largest in all four Nordic countries and both TALIS cycles. Prior research has also indicated relevant links between job satisfaction and the overall teacher–student relationship (Collie et al., 2012; Gil-Flores, 2017; Veldman et al., 2013).

Conversely, in the results related to the TALIS 2013 data, the factor teacher–student relations was significantly associated with classroom SES-ethnic composition and teachers' professional self-efficacy. The former has been reported in several studies, that is, a decrease in job satisfaction is affected by an increasing proportion of students from socioeconomically disadvantaged homes (Matsuoka, 2015; Wang et al., 2019). Schwab (2019) has also demonstrated how general self-efficacy is valuable in understanding the teacher–student relations: the higher the teachers' general self-efficacy is, the higher their student-specific self-efficacy will be. The latter, as Schwab reported, was lower for students from the special needs spectrum (i.e., learning, behavioural and emotional disorders). TALIS 2018 data also indicate that job satisfaction and teacher–student relations are highly affected by disadvantaged classroom academic environment. Given that job satisfaction is directly associated with occupational well-being, motivation and staying in the profession (Dicke et al., 2020), it is critical to provide the teachers with support in addressing the diversity they come across in the classrooms in a more sustained manner.

In all the Nordic countries, the analysis of TALIS 2013 data revealed the importance of teachers' self-efficacy for their collaboration activities with other teachers and their behaviour and attitude towards their students' learning and well-being. In

turn, these factors together effectively affect teachers' job satisfaction. The finding very much links to a plethora of research on the importance of self-efficacy for teachers' professional practices and commitment to the profession (Chesnut & Burley, 2015; Vieluf et al., 2013; Zee & Koomen, 2016). Even though no student outcome was included in TALIS, it has been shown that these chained effects can enhance students' academic performance (e.g., Bandura, 1977; Klassen & Tze, 2014). However, the pattern was only partially confirmed in the TALIS 2018 analysis. One explanation for this absence could be in the data 'unavailability'; that is, the two teacher-related factors in the TALIS 2013 model (i.e., teacher's constructivist beliefs and teacher collaboration) are absent from the TALIS 2018 models. These scales were replaced by teacher's personal and social utility motivation to teach and teaching practices. Another reason may be the somewhat lower explanatory power of the TALIS 2018 model. The correlation between classroom SES-ethnic composition and classroom academic environment was exceptionally high, leading to the exclusion of the classroom academic environment construct to avoid multicollinearity issues. Consequently, the model structure of the two TALIS cycles was not identical, and the estimation and interpretation of the interrelationships among the factors estimated from the model might be differentiated.

Although our results provide corroborating evidence in support of student–teacher relations or self-efficacy as affecting teachers' job satisfaction, diversities were also observed across the Nordic countries. The result patterns from the TALIS 2013 model outlined two subgroups of Nordic countries with similar mechanisms: the Norway-Sweden group and the Denmark-Finland group. This distinction is lost in the 2018 results, leading to more diverging and country-specific patterns, such as the importance of social utility value for Norway, adverse classroom composition in Sweden or teachers' effective professional development positively impacting personal and the social utility values of teachers in Finland. These observed diversities and changing patterns may be because of the gradually dissolved unity of the Nordic model by different reform actions taken in recent years, such as in Sweden (Lundahl, 2016), as well as in the long-term prerequisites for the teaching profession, where Finland stands out (Aspfors et al., 2014).

5.4.4   Limitation and Further Research

The nature of the data used in the current study holds both advantages and disadvantages. Although the data provide solid grounds for a comparative perspective, the data are cross-sectional. This means that even when observing information from different cycles, that is, 2013 and 2018, we cannot consider this to be a longitudinal investigation because different teachers within a country partake in each cycle. Nonetheless, the data do allow for conclusions on trends or a lack of these on established relationships within and across countries.

Second, with each TALIS cycle, a more robust data set has been built, offering more and more varied scales on the different aspects of teacher quality. With this in

mind, some of the constructs used in the presented models are there in both the 2013 and 2018 data, while few were novel to the 2018 cycle (e.g., personal utility value). Although the use of the same variables in the 2013 and 2018 models would offer opportunities for a direct comparison between the models, we opted for a more comprehensive view that would not limit our investigation merely to the constructs available in both cycles. Our assumption was that the approach would allow for a more nuanced view of the essential mechanisms contributing to teachers' job satisfaction.

In this round of our investigation, we opted for one-level models. We were guided by the idea that such an approach would foster more focused studies at a later stage that could involve the exploration of school-level factors pertinent to the particular direct and indirect effects established in this step. The results for the outcome variable in the operationalised model in TALIS 2018 (e.g., around 18% variance in job satisfaction being accounted for by the model in the Nordic countries) support this line of thinking.

5.5   Conclusions

In this chapter, we have investigated how the different aspects of teacher quality may affect job satisfaction, here in connection to diverse school environments relative to student composition and outcomes. The extent that the determined mechanisms apply across the Nordic countries and if the same patterns are consistent over different time points became the second focus of the study. Although common values are shared across the Nordic arena (Blossing et al., 2014; Imsen et al., 2017; Lundahl, 2016) and some of these are mirrored in the results of the current study (i.e., patterns in 2013 data), these also point to some diverse practices and ideas pertinent to individual countries (e.g., Aspfors et al., 2014; Wollscheid & Opheim, 2016). The latter are especially noticeable in the observed mechanism for 2018, indicating the presence of more diversified practices across the Nordic countries. Although equity and quality are still the common goals that these countries are striving to achieve, the mechanisms through which each education system approaches these goals have become more diverse. Both the changing patterns and differences could originate in the steadily dissolved unity of the Nordic model, here affected by the different reform actions taken in recent years. Sweden is a clear example of the latter with its extensive decentralisation and deregulation reforms, while Finland stands out with its long-term prerequisites for the teaching profession. The current evidence (i.e., the importance of social utility value for Norway, adverse classroom composition in Sweden or teacher effective professional development positively impacting personal and social utility values of teachers in Finland) warrants a continuation of the investigation into these distinctive patterns, with the possible inclusion of additional factors and mechanisms from the school level.

Appendices

Appendix A: Standardized Direct Effects Among Variable in the Path Analysis for the Four Nordic Countries in TALIS 2013

	Denmark		Finland		Norway		Sweden	
	Estimate	Z	Estimate	Z	Estimate	Z	Estimate	Z
TJOBSATS ON TSELEFFS	**.19**	**7.04**	**.20**	**8.45**	**.16**	**6.88**	**.15**	**7.76**
TJOBSATS ON TEFFPROS	.01	.48	.03	1.21	**.08**	**3.68**	**.09**	**3.81**
TJOBSATS ON TCONSBS	.03	1.01	.00	−.18	**.06**	**2.13**	**.05**	**2.49**
TJOBSATS ON TCCOLLS	**.07**	**2.80**	**.07**	**3.58**	**.13**	**6.11**	**.15**	**7.86**
TJOBSATS ON TSCTSTUDS	**.28**	**11.49**	**.32**	**14.69**	**.31**	**12.46**	**.19**	**10.10**
TJOBSATS ON TQ	**−.06**	**−3.29**	−.02	−.85	−.04	−1.56	**−.07**	**−4.19**
TJOBSATS ON SESMIG	−.04	−1.17	**−.07**	**−3.09**	−.04	−1.84	**−.07**	**−2.65**
TJOBSATS ON CLACDEN	**−.12**	**−3.63**	**−.05**	**−2.14**	**−.06**	**−2.23**	−.04	−1.79
TCCOLLS ON TSELEFFS	**.16**	**5.53**	**.18**	**8.48**	**.20**	**5.31**	**.17**	**7.60**
TCCOLLS ON TCONSBS	.03	1.15	.03	1.35	−.04	−1.28	**.05**	**2.08**
TCCOLLS ON TEFFPROS	**.13**	**4.28**	**.18**	**7.70**	**.19**	**5.32**	**.23**	**10.35**
TCCOLLS ON TQ	.00	.11	−.03	−1.40	.03	1.23	.04	1.54
TCCOLLS ON SESMIG	.02	.55	**.11**	**3.10**	−.01	−.30	.04	1.46
TCCOLLS ON CLACDEN	**.23**	**5.32**	**.17**	**6.73**	**.12**	**3.45**	**.14**	**6.44**
TSCTSTUD ON TSELEFFS	**.17**	**6.46**	**.23**	**10.28**	**.21**	**6.56**	**.21**	**10.34**
TSCTSTUD ON TCONSBS	**.15**	**4.83**	.03	1.46	**.08**	**2.79**	.05	1.92

	Denmark		Finland		Norway		Sweden	
	Estimate	Z	Estimate	Z	Estimate	Z	Estimate	Z
TSCTSTUD ON TEFFPROS	**.13**	**4.65**	.05	1.65	**.12**	**3.52**	.06	2.28
TSCTSTUD ON TQ	−.04	−1.87	.03	1.11	.01	.42	.02	.88
TSCTSTUD ON SESMIG	**-.13**	**−3.82**	**−.05**	**−2.00**	**−.05**	**−1.98**	**−.06**	**−2.15**
TSCTSTUD ON CLACDEN	−.04	−1.13	**−.12**	**−4.47**	.00	−.15	**−.09**	**−3.24**
TSELEFFS ON TEFFPROS	**.07**	**2.24**	**.12**	**5.62**	**.07**	**2.82**	**.08**	**3.59**
TSELEFFS ON TQ	**.12**	**4.16**	.03	1.40	.06	1.79	.04	1.93
TSELEFFS ON SESMIG	−.05	−1.37	**.08**	**3.05**	**.11**	**4.19**	.02	1.08
TSELEFFS ON CLACDEN	**−.12**	**−3.35**	.04	1.80	−.13	−5.32	−.02	−1.00
TCONSBS ON TEFFPROS	**.09**	**2.79**	**.10**	**4.41**	.04	1.37	.04	1.72
TCONSBS ON TQ	.03	.95	.04	1.57	−.02	−.59	**−.05**	**−2.03**
TCONSBS ON SESMIG	**.09**	**3.13**	.02	1.03	.02	.79	.03	1.16
TCONSBS ON CLACDEN	**−.13**	**−3.68**	.03	1.40	−.05	−1.58	−.01	−.30
TEFFPROS ON TQ	−.01	−.49	.04	1.64	.00	−.10	.03	1.46
TEFFPROS ON SESMIG	.04	.88	.04	1.21	.02	.54	.00	.13
TEFFPROS ON CLACDEN	**−.09**	**−2.14**	**−.06**	**−2.18**	.01	.44	.00	−.11

Appendix B: Standardized Total Direct and Indirect Effects Among Variable in the Path Analysis for the Four Nordic Countries in TALIS 2013

	Danmark		Finland		Norway		Sweden	
Effects from TQ to TJOBSATS	est.	z	est.	z	est.	z	est.	z
Total	−.04	−1.63	.00	.21	−.02	−.79	**−.05**	**−2.79**
Direct	**−.06**	**−3.29**	−.02	−.85	−.04	−1.56	**−.07**	**−4.19**
Total indirect	.02	1.559	.02	1.90	.02	1.29	**.02**	**2.09**
Effects from SESMIG	est.	z	est.	z	est.	z	est.	z
Total	**−.08**	**−2.13**	**−.05**	**−1.98**	−.02	−1.07	**−.07**	**−2.33**
Direct	−.04	−1.17	**−.07**	**−3.09**	−.04	−1.84	**−.07**	**−2.65**
Total indirect	**−.04**	**−2.34**	.02	1.43	.02	1.18	.00	.18
Effects from CLACDEN	est.	z	est.	z	est.	z	est.	z
Total	**−.16**	**−4.43**	**-.07**	**−2.74**	**−.08**	**−2.69**	−.04	−1.68
Direct	**−.12**	**−3.63**	**−.05**	**−2.14**	**−.06**	**−2.23**	−.04	−1.79
Total indirect	**−.04**	**−2.62**	−.02	−1.50	−.02	−1.35	.00	.04
Effects from TEFFPRO	est.	z	est.	z	est.	z	est.	z
Total	**.08**	**2.69**	**.09**	**3.59**	**.16**	**5.32**	**.15**	**6.34**
Direct	.01	0.48	.03	1.21	**.08**	**3.68**	.09	3.81
Total indirect	**.07**	**5.22**	**.06**	**4.63**	**.08**	**5.03**	.07	7.04
Effects from TSELEFF	est.	z	est.	z	est.	z	est.	z
Total	**.24**	**8.85**	**.29**	**12.46**	**.25**	**9.05**	**.21**	**11.01**
Direct	**.19**	**7.04**	**.20**	**8.45**	**.16**	**6.88**	**.15**	**7.76**
Total indirect	**.06**	**5.93**	**.09**	**9.44**	.089	8.09	.07	9.95
Effects from TCONSBS	est.	z	est.	z	est.	z	est.	z
Total	**.07**	**2.60**	.01	.45	**.08**	**3.17**	.07	3.31
Direct	.025	1.010	−.00	−.18	**.06**	**2.13**	.05	2.49
Total indirect	**.044**	**4.342**	.01	1.65	.02	1.78	**.02**	**2.39**

Appendix C: Detailed Indirect Effect in the Operationalized Model in All Four Nordic Countries in TALIS 2013

	Denmark		Finland		Norway		Sweden	
	est	z	est	z	est	z	est	z
TQ on TSELEFFS on TJOBSATS	**.02**	**3.45**	.01	1.35	.01	1.82	.01	1.87
TQ on TSCTSTUD on TJOBSATS	−.01	−1.82	.01	1.10	.00	.42	.00	.89
TQ on TCONSBS on TJOBSATS	.00	.67	.00	−.18	.00	−.59	.00	−1.51
TQ on TEFFPROS on TJOBSATS	.00	−.34	.00	1.02	.00	−.11	.00	1.36

	Denmark		Finland		Norway		Sweden	
	est	z	est	z	est	z	est	z
TQ on TCCOLLS on TJOBSATS	.00	.11	.00	−1.33	.00	1.23	.01	1.58
TQ on TEFFPROS on TSELEFFS on TJOBSATS	.00	−.47	.00	1.56	.00	−.11	.00	1.39
TQ on TSELEFFS on TSCTSTUD on TJOBSATS	**.01**	**3.51**	.00	1.37	.00	1.72	.00	1.82
TQ on TCONSBS on TSCTSTUD on TJOBSATS	.00	.95	.00	1.06	.00	−.54	.00	−1.50
TQ on TEFFPROS on TSCTSTUD on TJOBSATS	.00	−.48	.00	1.11	.00	−.11	.00	1.12
TQ on TEFFPROS on TCONSBS on TJOBSATS	.00	−.42	.00	−.18	.00	−.10	.00	1.08
TQ on TSELEFFS on TCCOLLS on TJOBSATS	**.00**	**2.37**	.00	1.31	.00	1.70	.00	1.86
TQ on TCONSBS on TCCOLLS on TJOBSATS	.00	.83	.00	1.00	.00	.54	.00	−1.58
TQ on TEFFPROS on TCCOLLS on TJOBSATS	.00	−.48	.00	1.50	.00	−.10	.00	1.42
TQ on TEFFPROS on TSELEFFS on TSCTSTUD on TJOBSATS	.00	−.47	.00	1.59	.00	−.11	.00	1.38
TQ on TEFFPROS on TCONSBS on TSCTSTUD on TJOBSATS	.00	−.49	.00	1.02	.00	−.10	.00	.91
TQ on TEFFPROS on TSELEFFS on TCCOLLS on TJOBSATS	.00	−.47	.00	1.46	.00	−.11	.00	1.32
TQ on TEFFPROS on TCONSBS on TCCOLLS on TJOBSATS	.00	−.43	.00	.92	.00	.10	.00	1.00
SESMIG on TSELEFFS on TJOBSATS	−.01	−1.28	**.02**	**2.98**	**.02**	**3.73**	.00	1.09
SESMIG on TSCTSTUD on TJOBSATS	**−.04**	**−3.53**	**−.02**	**−1.96**	−.02	−1.91	**−.01**	**−2.12**
SESMIG on TCONSBS on TJOBSATS	.00	.96	.00	−.18	.00	.71	.00	1.07
SESMIG on TEFFPROS on TJOBSATS	.00	.49	.00	.93	.00	.52	.00	.13
SESMIG on TCCOLLS on TJOBSATS	.00	.53	**.01**	**2.21**	.00	−.30	.01	1.42
SESMIG on TEFFPROS on TSELEFFS on TJOBSATS	.00	.83	.00	1.14	.00	.51	.00	.13
SESMIG on TSELEFFS on TSCTSTUD on TJOBSATS	.00	−1.33	**.01**	**2.89**	**.01**	**3.74**	.00	1.05
SESMIG on TCONSBS on TSCTSTUD on TJOBSATS	**.00**	**2.63**	.00	.84	.00	.80	.00	.96
SESMIG on TEFFPROS on TSCTSTUD on TJOBSATS	.00	.88	.00	1.02	.00	.52	.00	.13
SESMIG on TEFFPROS on TCONSBS on TJOBSATS	.00	.64	.00	−.18	.00	.50	.00	.13

(continued)

	Denmark		Finland		Norway		Sweden	
	est	z	est	z	est	z	est	z
SESMIG on TSELEFFS on TCCOLLS on TJOBSATS	.00	−1.28	.00	2.17	**.00**	**3.09**	.00	1.05
SESMIG on TCONSBS on TCCOLLS on TJOBSATS	.00	.98	.00	.79	.00	−.59	.00	1.01
SESMIG on TEFFPROS on TCCOLLS on TJOBSATS	.00	.85	.00	1.09	.00	.53	.00	.13
SESMIG on TEFFPROS on TSELEFFS on TSCTSTUD on TJOBSATS	.00	.83	.00	1.13	.00	.50	.00	.13
SESMIG on TEFFPROS on TCONSBS on TSCTSTUD on TJOBSATS	.00	.82	.00	.94	.00	.46	.00	.13
SESMIG on TEFFPROS on TSELEFFS on TCCOLLS on TJOBSATS	.00	.82	.00	1.08	.00	.50	.00	.14
SESMIG on TEFFPROS on TCONSBS on TCCOLLS on TJOBSATS	.00	.70	.00	.81	.00	−.46	.00	.13
CLACDEN on TSELEFFS on TJOBSATS	**−.02**	**−3.33**	.01	1.74	**−.02**	**−4.07**	.00	−1.00
CLACDEN on TSCTSTUD on TJOBSATS	−.01	−1.14	−.04	−4.14	.00	−.15	**−.02**	**−3.12**
CLACDEN on TCONSBS on TJOBSATS	.00	−1.00	.00	−.18	.00	−1.27	.00	−.30
CLACDEN on TEFFPROS on TJOBSATS	.00	−.49	.00	−.99	.00	.43	.00	−.11
CLACDEN on TCCOLLS on TJOBSATS	**.02**	**2.85**	**.01**	**3.27**	**.02**	**3.11**	**.02**	**5.32**
CLACDEN on TEFFPROS on TSELEFFS on TJOBSATS	.00	−1.42	.00	−1.86	.00	.42	.00	−.11
CLACDEN on TSELEFFS on TSCTSTUD on TJOBSATS	**−.01**	**−3.01**	.00	1.79	**−.01**	**−3.75**	.00	−.98
CLACDEN on TCONSBS on TSCTSTUD on TJOBSATS	−.01	−2.75	.00	1.00	.00	−1.30	.00	−.29
CLACDEN on TEFFPROS on TSCTSTUD on TJOBSATS	**.00**	**−2.09**	.00	−1.42	.00	.44	.00	−.11
CLACDEN on TEFFPROS on TCONSBS on TJOBSATS	.00	−.94	.00	.18	.00	.42	.00	−.11
CLACDEN on TSELEFFS on TCCOLLS on TJOBSATS	**.00**	**−2.08**	.00	1.53	**.00**	**−2.76**	.00	−.97
CLACDEN on TCONSBS on TCCOLLS on TJOBSATS	.00	−1.06	.00	1.04	.00	.90	.00	−.29
CLACDEN on TEFFPROS on TCCOLLS on TJOBSATS	.00	−1.86	.00	−1.83	.00	.43	.00	−.11
CLACDEN on TEFFPROS on TSELEFFS on TSCTSTUD on TJOBSATS	.00	−1.44	.00	−1.91	.00	.44	.00	−.11

	Denmark		Finland		Norway		Sweden	
	est	z	est	z	est	z	est	z
CLACDEN on TEFFPROS on TCONSBS on TSCTSTUD on TJOBSATS	.00	−1.49	.00	−1.15	.00	.44	.00	−.11
CLACDEN on TEFFPROS on TSELEFFS on TCCOLLS on TJOBSATS	.00	−1.42	.00	−1.65	.00	.42	.00	−.11
CLACDEN on TEFFPROS on TCONSBS on TCCOLLS on TJOBSATS	.00	−.92	.00	−.98	.00	−.39	.00	−.11
TEFFPROS on TSELEFFS on TJOBSATS	**.01**	**2.09**	**.02**	**4.47**	**.01**	**2.57**	.01	.37
TEFFPROS on TSCTSTUD on TJOBSATS	**.04**	**4.08**	.02	1.67	**.04**	**3.28**	.01	2.21
TEFFPROS on TCONSBS on TJOBSATS	.00	1.02	.00	−.18	.00	1.25	.00	1.42
TEFFPROS on TCCOLLS on TJOBSATS	**.01**	**2.32**	**.01**	**3.33**	**.02**	**4.15**	**.03**	**6.23**
TEFFPROS on TSELEFFS on TSCTSTUD on TJOBSATS	**.00**	**2.10**	**.01**	**5.09**	**.01**	**2.83**	**.00**	**3.30**
TEFFPROS on TCONSBS on TSCTSTUD on TJOBSATS	**.00**	**2.79**	.00	1.41	.00	1.16	.00	1.16
TEFFPROS on TSELEFFS on TCCOLLS on TJOBSATS	.00	1.77	**.00**	**2.66**	**.00**	**2.14**	**.00**	**2.89**
TEFFPROS on TCONSBS on TCCOLLS on TJOBSATS	.00	1.05	.00	1.17	.00	−.91	.00	1.26
TSELEFFS-TSCTSTUD on TJOBSATS	**.05**	**5.50**	**.07**	**8.59**	**.06**	**6.50**	**.04**	**7.34**
TSELEFFS on TCCOLLS on TJOBSATS	**.01**	**2.63**	**.01**	**3.10**	**.03**	**3.65**	**.03**	**5.63**
TCONSBS-TSCTSTUD on TJOBSATS	**.04**	**4.39**	.01	1.44	**.03**	**2.87**	.01	1.89
TCONSBS on TCCOLLS on TJOBSATS	.00	1.09	.00	1.27	−.01	−1.23	**.01**	**2.04**

Appendix D: Standardized Direct Effects Among Variable in the Path Analysis for the Four Nordic Countries in TALIS 2018

	Denmark		Finland		Norway		Sweden	
	Est.	z	Est.	z	Est.	z	Est.	z
T3JOBSA ON TQ	−.01	−.36	−.03	−.87	−.05	−.65	**−.16**	**−3.26**

(continued)

	Denmark		Finland		Norway		Sweden	
	Est.	z	Est.	z	Est.	z	Est.	z
T3JOBSA ON SESMIG	.13	.82	.14	.95	−.17	−.93	**.29**	**2.38**
T3JOBSA ON T3SELF	**.25**	**2.17**	**.25**	**2.72**	.05	.26	.04	.63
T3JOBSA ON T3PERUT	**−.08**	**−2.38**	.03	.44	−.02	−.35	**−.16**	**−3.17**
T3JOBSA ON T3SOCUT	**.15**	**4.19**	**.15**	**3.30**	**.14**	**3.92**	**.16**	**3.41**
T3JOBSA ON T3EFFPD	.01	.08	−.08	−1.21	−.13	−1.30	.01	.14
T3JOBSA ON T3TPRA	−.03	−.61	.04	.54	−.01	−.12	−.11	−1.80
T3JOBSA ON T3STUD	**.25**	**4.32**	**.27**	**5.68**	**.28**	**2.04**	**.42**	**8.33**
T3TPRA ON TQ	−.02	−.44	.08	1.53	**.11**	**2.21**	.10	1.58
T3TPRA ON SESMIG	.12	.78	.00	.01	.08	.67	.26	1.85
T3TPRA ON T3EFFPD	−.02	−.18	−.07	−.73	−.14	−1.72	−.03	−.20
T3TPRA ON T3SELF	.36	1.67	.17	1.30	.10	.94	.12	.73
T3TPRA ON T3PERUT	.02	.49	.11	1.55	.04	1.01	−.08	−1.17
T3TPRA ON T3SOCUT	**.14**	**2.77**	**.11**	**2.24**	**.20**	**4.26**	**.18**	**2.98**
T3STUD ON TQ	.02	0.44	.02	.79	.05	.80	−.05	−1.42
T3STUD ON SESMIG	**−.49**	**−4.53**	**−.50**	**−8.43**	**−.41**	**−5.97**	**−0.67**	**−9.76**
T3STUD ON T3EFFPD	−.01	−.09	**.08**	**2.15**	**.34**	**5.52**	.04	.58
T3STUD ON T3SELF	.16	1.10	**.19**	**3.16**	.04	.49	**.31**	**3.89**
T3STUD ON T3PERUT	−.06	−1.61	**.06**	**2.56**	.10	1.59	**.13**	**1.96**
T3STUD ON T3SOCUT	**.12**	**2.79**	**.08**	**3.42**	**.08**	**2.13**	.03	.59
T3SELF ON TQ	**−.11**	**−2.35**	−.07	−1.45	−.04	−.54	.08	1.09
T3SELF ON SESMIG	**−.75**	**−7.95**	**−.58**	**−8.46**	**−.80**	**−15.34**	**.34**	**1.97**
T3SELF ON T3EFFPD	**−.24**	**−2.71**	**.25**	**3.44**	−.01	−.11	.14	1.46
T3PERUT ON TQ	**.09**	**2.50**	−.01	−.27	.14	2.57	−.14	−1.71
T3PERUT ON SESMIG	**−.20**	**−2.70**	**−.18**	**−1.98**	**−.20**	**−3.26**	−.08	−.85
T3PERUT ON T3EFFPD	**.19**	**2.48**	**.20**	**2.46**	.04	.49	**.35**	**4.29**
T3SOCUT ON TQ	−.02	−.58	**−.08**	**−2.60**	−.04	−.58	−.11	−1.67
T3SOCUT ON SESMIG	.13	1.31	−.10	−1.55	**−.16**	**−2.78**	**−.30**	**−3.63**
T3SOCUT ON T3EFFPD	.06	.81	**.14**	**2.21**	−.08	−1.01	-.07	−.72
T3EFFPD ON TQ	−.01	−.08	.00	.01	.04	.60	.08	.86
T3EFFPD ON SESMIG	**−.38**	**−2.82**	**−.44**	**−5.51**	**−.46**	**−7.52**	**−.27**	**−2.40**

Appendix E: Standardized Total Direct and Indirect Effects Among Variables in the Path Analysis for the Four Nordic Countries in TALIS 2018

	Denmark		Finland		Norway		Sweden	
Effects from TQ to T3JOBSA	Estimate	z	Estimate	z	Estimate	z	Estimate	z
Total	−.05	−1.40	−.06	−1.45	−.04	−.63	**−.19**	**−3.66**
Total indirect	−.04	−1.74	−.03	−1.51	.00	.11	−.02	−.64
Direct	−.01	−.36	−.03	−.87	−.05	−.65	**−.16**	**−3.26**
Effects from SESMIG to T3JOBSA								
Total	−.14	−1.03	**−.22**	**−2.14**	**−.34**	**−3.80**	−.01	−.11
Total indirect	**−.27**	**−3.50**	**−.36**	**−5.23**	−.16	−1.30	**−.30**	**−3.98**
Direct	.13	.82	.14	.95	−.17	−.93	**.29**	**2.38**
Effects from T3EFFPD To T3JOBSA								
Total	−.07	−.59	.05	.73	−.05	−.51	.01	.10
Total indirect	−.08	−1.19	**.13**	**5.75**	.08	1.65	.00	.00
DIRECT	.01	.08	−.08	−1.21	−.13	−1.30	.01	.14
Effects from T3PERUT to T3JOBSA								
Total	**−.10**	**−2.97**	.05	.71	.01	.12	−.10	−1.63
Total indirect	−.02	−1.51	**.02**	**2.24**	.03	1.08	.06	1.85
Direct	**−.08**	**−2.38**	.03	.44	−.02	−.35	**−.16**	**−3.17**
Effects from T3SOCUT to T3JOBSA								
Total	**.17**	**4.95**	**.17**	**3.97**	**.16**	**3.97**	**.15**	**2.90**
Total indirect	.03	1.70	**.02**	**2.91**	.02	1.25	-.01	−.30
Direct	**.15**	**4.19**	**.15**	**3.30**	**.14**	**3.92**	**.16**	**3.41**
Effects from T3SELF to T3JOBSA								
Total	**.28**	**2.38**	**.31**	**3.92**	.06	.31	.15	1.57
Total indirect	.03	.71	**.06**	**2.34**	.01	.46	**.12**	**2.57**
Direct	**.25**	**2.17**	**.25**	**2.72**	.05	.26	.04	.63

Appendix F: Detailed Indirect Effect in the Operationalized Model in All Four NORDIC Countries in TALIS 2018

Path	Denmark		Finland		Norway		Sweden	
Effects from TQ to T3JOBSA	Est.	z	Est.	z	Est.	z	Est.	z
Specific indirect								
TQ – T3SELF – T3JOBSA	−.03	−1.57	−.02	−1.16	.00	−.23	.00	.55
TQ – T3PERUT – T3JOBSA	−.01	−1.71	.00	−.20	.00	−.36	.02	1.54
TQ – T3SOCUT – T3JOBSA	.00	−.58	**−.01**	**−1.99**	−.01	−.57	−.02	−1.37
TQ – T3EFFPD – T3JOBSA	.00	−.05	.00	−.01	−.01	−.58	.00	.13
TQ – T3TPRA – T3JOBSA	.00	.31	.00	.56	.00	−.12	−.01	−1.01
TQ – T2STUD – T3JOBSA	.01	.45	.01	.83	.01	.81	−.02	−1.37
TQ – T3EFFPD – T3SELF – T3JOBSA	.00	.08	.00	.01	.00	−.09	.00	.45
TQ – T3EFFPD – T3PERUT – T3JOBSA	.00	.08	.00	.01	.00	−.29	−.01	−.79
TQ – T3EFFPD – T3SOCUT – T3JOBSA	.00	−.08	.00	.01	.00	−.46	.00	−.53
TQ – T2SELF – T3TPRA – T3JOBSA	.00	.62	.00	−.54	.00	.12	.00	−.51
TQ – T3PERUT – T3TPRA – T3JOBSA	.00	−.40	.00	−.21	.00	−.11	.00	−.81
TQ – T3SOCUT – T3TPRA – T3JOBSA	.00	.39	.00	−.62	.00	.11	.00	1.16
TQ – T3EFFPD – T3TPRA – T3JOBSA	.00	−.08	.00	−.01	.00	.10	.00	.21
TQ – T3SELF – T3STUD – T3JOBSA	.00	−.99	.00	−1.59	.00	−.38	.01	.98
TQ – T3PERUT – T3STUD – T3JOBSA	.00	−1.33	.00	−.27	.00	1.04	−.01	−1.16
TQ – T3SOCUT – T3STUD – T3JOBSA	.00	−.60	**.00**	**−2.21**	.00	−.47	.00	−.51
TQ – T3EFFPD – T3STUD – T3JOBSA	.00	.05	.00	.01	.00	.55	.00	.51
TQ – T3EFFPD – T3SELF – T3TPRA – T3JOBSA	.00	−.08	.00	.01	.00	.07	.00	−.55
TQ – T3EFFPD – T3PERUT – T3TPRA – T3JOBSA	.00	.07	.00	.01	.00	−.10	.00	.72
TQ – T3EFFPD – T3SOCUT – T3TPRA – T3JOBSA	.00	.07	.00	.01	.00	.10	.00	.49
TQ – T3EFFPD – T3SELF – T3STUD – T3JOBSA	.00	.08	.00	.01	.00	−.10	.00	.76
TQ – T3EFFPD – T3PERUT – T3STUD – T3JOBSA	.00	.08	.00	.01	.00	.31	.00	.81
TQ – T3EFFPD – T3SOCUT – T3STUD – T3JOBSA	.00	−.08	.00	.01	.00	−.43	.00	−.39

Path	Denmark		Finland		Norway		Sweden	
Effects from TQ to T3JOBSA	Est.	z	Est.	z	Est.	z	Est.	z
Effects from SESMIG to T3JOBSA								
Specific indirect								
SESMIG – T3SELF – T3JOBSA	**−.19**	**−2.16**	**−.15**	**−2.25**	−.04	−.26	.01	.52
SESMIG – T3PERUT – T3JOBSA	.02	1.76	−.01	−.39	.00	.34	.01	.87
SESMIG – T3SOCUT – T3JOBSA	.02	1.24	−.01	−1.14	**−.02**	**−2.31**	**−.05**	**−2.33**
SESMIG – T3EFFPD – T3JOBSA	.00	-.08	.04	1.27	.06	1.33	.00	−.14
SESMIG – T3TPRA – T3JOBSA	.00	−.64	.00	.01	.00	−.12	−.03	−1.49
SESMIG – T3STUD – T3JOBSA	**−.13**	**−2.70**	**−.14**	**−4.46**	**−.11**	**−2.23**	**−.28**	**−5.97**
SESMIG – T3EFFPD – T3SELF – T3JOBSA	.02	1.05	**−.03**	**−3.04**	.00	.09	.00	−.51
SESMIG – T3EFFPD – T3PERUT – T3JOBSA	.01	1.48	.00	−.50	.00	.33	.02	1.68
SESMIG – T3EFFPD – T3SOCUT – T3JOBSA	.00	−.75	**−.01**	**−2.52**	.01	.99	.00	.71
SESMIG – T3SELF – T3TPRA – T3JOBSA	.01	.67	.00	−.53	.00	.12	.00	−.69
SESMIG – T3PERUT – T3TPRA – T3JOBSA	.00	.41	.00	−.37	.00	.11	.00	−.66
SESMIG – T3SOCUT – T3TPRA – T3JOBSA	.00	−.49	.00	−.69	.00	.12	.01	1.31
SESMIG – T3EFFPD – T3TPRA – T3JOBSA	.00	−.19	.00	.36	.00	−.11	.00	−.20
SESMIG – T3SELF – T3STUD – T3JOBSA	−.03	−.99	**−.03**	**−2.40**	−.01	−.45	.04	1.41
SESMIG – T3PERUT – T3STUD – T3JOBSA	.00	1.42	.00	−1.67	−.01	−1.17	−.01	−.70
SESMIG – T3SOCUT – T3STUD – T3JOBSA	.00	1.18	.00	−1.19	.00	−1.33	.00	−.56
SESMIG – T3EFFPD – T3STUD – T3JOBSA	.00	.09	**−.01**	**−2.19**	**−.04**	**−2.12**	−.01	−.54
SESMIG – T3EFFPD – T3SELF – T3TPRA – T3JOBSA	.00	−.55	.00	−.51	.00	−.07	.00	.68
SESMIG – T3EFFPD – T3PERUT – T3TPRA – T3JOBSA	.00	.40	.00	−.48	.00	.11	.00	−.93
SESMIG – T3EFFPD – T3SOCUT – T3TPRA – T3JOBSA	.00	.39	.00	−.54	.00	−.11	.00	−.64
SESMIG – T3EFFPD – T3SELF – T3STUD – T3JOBSA	.00	.89	−.01	−1.64	.00	.11	−.01	−1.10
SESMIG – T3EFFPD – T3PERUT – T3STUD – T3JOBSA	.00	1.35	.00	−1.38	.00	−.40	−.01	−1.33
SESMIG – T3EFFPD – T3SOCUT – T3STUD – T3JOBSA	.00	−.84	**.00**	**−2.21**	.00	.83	.00	.47
Effects from T3EFFPD To T3JOBSA								
Specific indirect								

(continued)

Path	Denmark		Finland		Norway		Sweden	
Effects from TQ to T3JOBSA	Est.	z	Est.	z	Est.	z	Est.	z
T3EFFPD – T3SELF – T3JOBSA	−.06	−1.32	.06	3.11	.00	−.09	.01	.55
T3EFFPD – T3PERUT – T3JOBSA	−.02	−1.63	.01	.48	.00	−.33	**−.06**	**-2.40**
T3EFFPD – T3SOCUT – T3JOBSA	.01	.75	**.02**	**2.88**	−.01	−1.01	−.01	−.69
T3EFFPD – T3TPRA – T3JOBSA	.00	.19	.00	−.35	.00	.11	.00	.21
T3EFFPD – T3STUD – T3JOBSA	.00	−.09	**.02**	**2.58**	**.10**	**2.13**	.02	.60
T3EFFPD – T3SELF – T3TPRA – T3JOBSA	.00	.60	.00	.50	.00	.07	.00	−.65
T3EFFPD – T3PERUT – T3TPRA – T3JOBSA	.00	−.41	.00	.47	.00	−.11	.00	.93
T3EFFPD – T3SOCUT – T3TPRA – T3JOBSA	.00	−.41	.00	.52	.00	.11	.00	.60
T3EFFPD – T3SELF – T3STUD – T3JOBSA	−.01	−.95	.01	1.74	.00	−.11	.02	1.40
T3EFFPD – T3PERUT – T3STUD – T3JOBSA	.00	−1.35	.00	1.55	.00	.40	.02	1.81
T3EFFPD – T3SOCUT – T3STUD – T3JOBSA	.00	.81	.00	2.13	.00	−.85	.00	−.45
Effects from T3PERUT to T3JOBSA								
Specific indirect								
T3PERUT – T3TPRA – T3JBSA	.00	−.41	.00	.43	.00	−.11	.01	.97
T3PERUT – T3STUD – T3JOBSA	−.01	−1.56	**.02**	**2.68**	.03	1.07	.06	1.83
Effects from T3SOCUT to T3JOBSA								
Specific indirect								
T3SOCUT – T3TPRA – T3JOBSA	.00	−.54	.00	.62	.00	−.11	−.02	−1.65
T3SOCUT – T3STUD – T3JOBSA	**.03**	**2.46**	**.02**	**3.44**	.02	1.33	.01	.59
Effects from T3SELF to T3JOBSA								
Specific indirect								
T3SELF – T3TPRA – T3JOBSA	−.01	−.66	.01	.53	.00	−.12	−.01	−.62
T3SELF – T3STUD – T3JOBSA	.04	1.02	**.05**	**2.34**	.01	.45	**.13**	**3.82**

References

Ahonen, S. (2014). A School for All in Finland. In: Blossing U., Imsen G., Moos L. (eds) *The Nordic Education Model. Policy Implications of Research in Education* (Vol 1). Dordrecht: Springer. https://doi.org/10.1007/978-94-007-7125-3_5.

Akiba, A., & Liang, L. (2016). Effects of teacher professional learning activities on student achievement growth. *The Journal of Educational Research, 109*(1), 99–110. https://doi.org/10.1080/00220671.2014.924470

Alvunger, D., Sundberg, D., & Wahlström, N. (2017). Teachers matter – But how? *Journal of Curriculum Studies, 49*(1), 1–6. https://doi.org/10.1080/00220272.2016.1205140

Antikainen, A. (2006). In search of the Nordic model in education. *Scandinavian Journal of Educational Research, 50*(3), 229–243. https://doi.org/10.1080/00313830600743258

Aspfors, J., Hansen, S.-E., & Ray, J. (2014). Stability, structure and development. Features constituting Finnish teacher education. *Scuola Democratica, 12*(3), 1–12. https://doi.org/10.12828/75800

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191–215. https://doi.org/10.1037/0033-295X.84.2.191

Banerjee, N., Stearns, E., Moller, S., & Mickelson, R. A. (2017). Teacher job satisfaction and student achievement: The roles of teacher professional community and teacher collaboration in schools. *American Journal of Education, 123*(2), 203–1241. https://doi.org/10.1086/689932

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173. https://doi.org/10.1037/0022-3514.51.6.1173

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., et al. (2009). 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

Berger, J.-L., & Lê Van, K. (2019). Teacher professional identity as multidimensional: Mapping its components and examining their associations with general pedagogical beliefs. *Educational Studies, 45*(2), 163–1181. https://doi.org/10.1080/03055698.2018.1446324

Blömeke, S., Gustafsson, J.-E., & Shavelson, R. J. (2015). Beyond dichotomies: Competence viewed as a continuum. *Zeitschrift für Psychologie, 223*(1), 3–13. https://doi.org/10.1027/2151-2604/a000194

Blömeke, S., Olsen, R. V., & Suhl, U. (2016). Relation of student achievement to the quality of their teachers and instructional quality. In T. Nilsen & J.-E. Gustafsson (Eds.), *Teacher quality, instructional quality and student outcomes* (IEA Research for Education 2) (pp. 21–50). Cham, Switzerland: Springer.

Blossing, U., Imsen, G., & Moos, L. (2014). Nordic schools in a time of change. In U. Blossing, G. Imsen, & L. Moos (Eds.), *The Nordic education model. Policy implications of research in education* (Vol. 1, pp. 1–114). Dordrecht, The Netherlands: Springer.

Blossing, U., & Söderström, Å. (2014). A school for every child in Sweden. In U. Blossing, G. Imsen, & L. Moos (Eds.), *The Nordic education model. Policy implications of research in education* (Vol. 1, pp. 17–134). Dordrecht, The Netherlands: Springer.

Bourdieu, P. (1990). Structures, habitus, practices. In *The logic of practice* (pp. 52–65). Cambridge, UK: Polity.

Buehl, M. M., & Beck, J. (2015). The relationship between teachers' beliefs and practices. In H. Fives & M. G. Gregorie (Eds.), *International handbook of research on teachers' beliefs* (pp. 66–184). New York: Routledge.

Caprara, G. V., Barbaranelli, C., Borgogni, L., & Steca, P. (2003). Efficacy beliefs as determinants of teachers' job satisfaction. *Journal of Educational Psychology, 95*(4), 821–832. https://doi.org/10.1037/0022-0663.95.4.821

Caprara, G. V., Barbaranelli, C., Steca, P., & Malone, P. S. (2006). Teachers' self-efficacy beliefs as determinants of job satisfaction and students' academic achievement: A study at the school level. *Journal of School Psychology, 44*(6), 473–1490. https://doi.org/10.1016/j.jsp.2006.09.001

Chesnut, S. R., & Burley, H. (2015). Self-efficacy as a predictor of commitment to the teaching profession: A meta-analysis. *Educational Research Review, 15*, 1–16. https://doi.org/10.1016/j.edurev.2015.02.001

Chmielewski, A. K. (2019). The global increase in the socioeconomic achievement gap, 1964 to 2015. *American Sociological Review, 84*(3), 517–544. https://doi.org/10.1177/0003122419847165

Coleman, J. S., Campbell, E., Hobson, C., McPartland, J., Mood, A., Weinfeld, F., et al. (1966). *Equality of educational opportunity*. Washington, DC: US Government Printing Office.

Collie, R. J., Shapka, J. D., & Perry, N. E. (2012). School climate and social-emotional learning: Predicting teacher stress, job satisfaction, and teaching efficacy. *Journal of Educational Psychology, 104*(4), 1189–11204. https://doi.org/10.1037/a0029356

Correnti, R. (2007). An empirical investigation of professional development effects on literacy instruction using daily logs. *Educational Evaluation and Policy Analysis, 29*(4), 262–1295. https://doi.org/10.3102/0162373707309074

Darling-Hammond, L. (2017). Teacher education around the world: What can we learn from international practice? *European Journal of Teacher Education, 40*(3), 1–19. https://doi.org/10.1080/02619768.2017.1315399

Dellinger, A. B., Bobbett, J. J., Olivier, D. F., & Ellett, C. D. (2008). Measuring teachers' self-efficacy beliefs: development and use of the TEBS-Self. *Teaching and Teacher Education, 24*(3), 751–1766. https://doi.org/10.1016/j.tate.2007.02.010

Desimone, L., Smith, T. M., & Phillips, K. (2013). Linking student achievement growth to professional development participation and changes in instruction: A longitudinal study of elementary students and teachers in title I schools. *Teachers College Record, 115*(5), 1–146.

Dicke, T., Marsh, H. W., Parker, P. D., Guo, J., Riley, P., & Waldeyer, J. (2020). Job satisfaction of teachers and their principals in relation to climate and student achievement. *Journal of Educational Psychology, 112*(5), 1061–1073. https://doi.org/10.1037/edu0000409

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.

Espinoza, O. (2007). Solving the equity–equality conceptual dilemma: a new model for analysis of the educational process. *Educational Research, 49*(4), 343–363. https://doi.org/10.1080/00131880701717198.

Ferguson, K., Frost, L., & Hall, D. (2012). Predicting teacher anxiety, depression, and job satisfaction. *Journal of Teaching and Learning, 8*(1), 27–142. https://doi.org/10.22329/jtl.v8i1.2896

Fischer, C., Fishman, B., Dede, C., Eisenkraft, A., Frumin, K., Foster, B., et al. (2018). Investigating relationships between school context, teacher professional development, teaching practices, and student achievement in response to a nationwide science reform. *Teaching and Teacher Education, 72*, 107–1121. https://doi.org/10.1016/j.tate.2018.02.011

Gil-Flores, J. (2017). The role of personal characteristics and school characteristics in explaining teacher job satisfaction. *Revista de Psicodidáctica, 22*(1), 16–22. https://doi.org/10.1387/RevPsicodidact.15501

Goe, L. (2007). *The link between teacher quality and student outcomes: A research synthesis*. Washington, DC, USA: National Comprehensive Center for Teacher Quality. Retrieved from: http://www.gtlcenter.org/sites/default/files/docs/LinkBetweenTQandStudentOutcomes.pdf.

Griffith, J. (2004). Relation of principal transformational leadership to school staff job satisfaction, staff turnover, and school performance. *Journal of Educational Administration, 42*(3), 333–356. https://doi.org/10.1108/09578230410534667

Gustafsson, J.-E., & Blömeke, S. (2018). Development of school achievement in the Nordic countries during half a century. *Scandinavian Journal of Educational Research, 62*(3), 386–1406. https://doi.org/10.1080/00313831.2018.1434829

Hanushek, E. A. (2008). Education production functions. In P. Newman (Ed.), *The new Palgrave dictionary of economics*. London: Palgrave Macmillan.

Hanushek, E. A., & Rivkin, S. G. (2012). The distribution of teacher quality and implications for policy. *Annual Review of Economics, 4*(1), 131–157. https://doi.org/10.1146/annurev-economics-080511-111001

Husén, T. (1989). The Swedish school reform—Exemplary both ways. *Comparative Education, 25*(3), 345–1355. https://doi.org/10.1080/0305006890250311

Imsen, G., Blossing, U., & Moos, L. (2017). Reshaping the Nordic education model in an era of efficiency. Changes in the comprehensive school project in Denmark, Norway, and Sweden since the millennium. *Scandinavian Journal of Educational Research, 61*(5), 568–1583. https://doi.org/10.1080/00313831.2016.1172502

Imsen, G., & Volckmar, N. (2014). The Norwegian school for all: Historical emergence and neoliberal confrontation. In U. Blossing, G. Imsen, & L. Moos (Eds.), *The Nordic education model. Policy implications of research in education* (Vol. 1, pp. 35–155). Dordrecht, The Netherlands: Springer.

Klassen, R. M., & Tze, V. M. C. (2014). Teachers' self-efficacy, personality, and teaching effectiveness: A meta-analysis. *Educational Research Review, 12*, 59–76. https://doi.org/10.1016/j.edurev.2014.06.001

König, J., Lammerding, S., Nold, G., Rohde, A., Strauß, S., & Tachtsoglou, S. (2016). Teachers' professional knowledge for teaching English as a foreign language: Assessing the outcomes of teacher education. *Journal of Teacher Education, 67*(4), 320–337. https://doi.org/10.1177/0022487116644956

Liu, Y., & Verblow, J. (2019). The operation of distributed leadership and the relationship with organisational commitment and job satisfaction of principals and teachers: A multi-level model and meta-analysis using the 2013 TALIS data. *International Journal of Educational Research, 96*, 41–55. https://doi.org/10.1016/j.ijer.2019.05.005

Lundahl, L. (2016). Equality, inclusion and marketisation of Nordic education: Introductory notes. *Research in Comparative and International Education, 11*(1), 3–12. https://doi.org/10.1177/0022487116644956

Matsumura, L. C., Garnier, H. E., & Resnick, L. B. (2010). Implementing literacy coaching: The role of school social resources. *Educational Evaluation and Policy Analysis, 32*(2), 249–272. https://doi.org/10.3102/0162373710363743

Matsuoka, R. (2015). School socioeconomic context and teacher job satisfaction in Japanese compulsory education. *Educational Studies in Japan: International Yearbook, 9*, 41–54.

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide* (8th ed.). Los Angeles: Muthén & Muthén.

Nie, Y., Tan, G. H., Liau, A. K., Lau, S., & Chua, B. L. (2012). The roles of teacher efficacy in instructional innovation: Its predictive relations to constructivist and didactic instruction. *Educational Research for Policy and Practice, 12*(1), 67–77. https://doi.org/10.1007/s10671-012-9128-y

Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis, 26*(3), 237–257. https://doi.org/10.3102/01623737026003237

OECD. (2013a). *TALIS 2013 Conceptual framework*. OECD Publishing. https://www.oecd.org/education/school/TALIS20ConceptualFramework_FINAL.pdf

OECD. (2013b). *TALIS 2013 technical report*. OECD Publishing. http://www.oecd.org/education/school/TALIS-technical-report-2013.pdf

OECD. (2018). *Equity in education: Breaking down barriers to social mobility*. PISA, OECD Publishing. https://doi.org/10.1787/9789264073234-en

OECD. (2019a). *PISA 2018 results (volume I): What students know and can do?* OECD Publishing. https://doi.org/10.1787/5f07c754-en.

OECD. (2019b). *TALIS 2018 results (volume I): Teachers and school leaders as lifelong learners*. TALIS, OECD Publishing. https://doi.org/10.1787/1d0bc92a-en

OECD. (2019c). *TALIS 2018 technical report*. OECD Publishing. https://www.oecd.org/education/talis/TALIS_2018_Technical_Report.pdf

Pajares, F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research, 62*(3), 307–332. https://doi.org/10.3102/00346543062003307

Penuel, W. R., Fishman, B., Yamaguchi, R., & Gallagher, L. P. (2007). What makes professional development effective? Strategies that foster curriculum implementation. *American Educational Research Journal, 44*(4), 921–958. https://doi.org/10.3102/0002831207308221

Rasmussen, A., & Moos, L. (2014). A school for less than all in Denmark. In U. Blossing, G. Imsen, & L. Moos (Eds.), *The Nordic education model. Policy implications of research in education* (Vol. 1, pp. 57–75). Dordrecht, The Netherlands: Springer.

Reimer, D. (2019). Teachers and teacher education: A call for a renewed sociological research agenda. *International Studies in Sociology of Education, 28*(2), 90–109. https://doi.org/10.108 0/09620214.2019.1601582

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

Richardson, P. W., & Watt, H. M. (2006). Who chooses teaching and why? Profiling characteristics and motivations across three Australian universities. *Asia-Pacific Journal of Teacher Education, 34*(1), 27–56. https://doi.org/10.1080/13598660500480290

Richardson, P. W., & Watt, H. M. G. (2016). Factors influencing teaching choice: Why do future teachers choose the career? In J. Loughran, & M. L. Hamilton (Eds.), *International Handbook of Teacher Education* (Vol. 2, pp. 275–304). Springer. https://doi. org/10.1007/978-981-10-0369-1_8

Roth, K. J., Garnier, H. E., Chen, C., Lemmens, M., Schwille, K., & Wickler, N. I. Z. (2011). Video-based lesson analysis: Effective science PD for teacher and student learning. *Journal of Research in Science Teaching, 48*(2), 117–148. https://doi.org/10.1002/tea.20408

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

Scherer, R., & Nilsen, T. (2016). The relations among school climate, instructional quality, and achievement motivation in mathematics. In T. Nilsen & J.-E. Gustafsson (Eds.), *Teacher quality, instructional quality and student outcomes* (pp. 51–80). Springer. https://doi. org/10.1007/978-3-319-41252-8_3

Schwab, S. (2019). Teachers' student-specific self-efficacy in relation to teacher and student variables. *Educational Psychology, 39*(1), 4–18. https://doi.org/10.1080/01443410.2018.1516861

Shulman, L. S. (1986). Those who understand: A conception of teacher knowledge. *American Educator, 10*(1), 9–15. http://www.jstor.org/stable/1175860

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*(1), 1–23. https://doi.org/10.17763/haer.57.1.j463w79r56455411

Skaalvik, E. M., & Skaalvik, S. (2009). Does school context matter? Relations with teacher burnout and job satisfaction. *Teacher and Teacher Education, 25*(3), 518–524. https://doi. org/10.1016/j.tate.2008.12.006

Skaalvik, E. M., & Skaalvik, S. (2014). Teacher self-efficacy and perceived autonomy: Relations with teacher engagement, job satisfaction, and emotional exhaustion. *Psychological Reports, 114*(1), 68–77. https://doi.org/10.2466/14.02.PR0.114k14w0

Taajamo, M. (2016). Teachers' professional development in Nordic countries. In *Northern Lights on PISA and TALIS* (pp. 101–133). Nordic Council of Ministers.

Torsney, B. M., Lombardi, D., & Ponnock, A. (2019). The role of values in pre-service teachers' intentions for professional engagement. *Educational Psychology, 39*(1), 19–37. https://doi. org/10.1080/01443410.2018.1504892

Varghese, C., Garwood, J. D., Bratsch-Hines, M., & Vernon-Feagans, L. (2016). Exploring magnitude of change in teacher efficacy and implications for students' literacy growth. *Teaching and Teacher Education, 55*, 228–239. https://doi.org/10.1016/j.tate.2016.01.011

Veldman, I., van Tartwijk, J., Brekelmans, M., & Wubbels, T. (2013). Job satisfaction and teacher–student relationships across the teaching career: Four case studies. *Teaching and Teacher Education, 32*, 55–65. https://doi.org/10.1016/j.tate.2013.01.005

Vieluf, S., Kuenther, M., & van de Vijver, F. J. R. (2013). Teacher self-efficacy in cross-cultural perspective. *Teaching and Teacher Education: An International Journal of Research and Studies, 35*, 92–103. https://doi.org/10.1016/j.tate.2013.05.006

Volckmar, N., & Wiborg, S. (2014). A social democratic response to market-led education policies: Concession or rejection? In U. Blossing, G. Imsen, & L. Moos (Eds.), *The Nordic education model. Policy implications of research in education* (Vol. 1, pp. 117–131). Dordrecht, The Netherlands: Springer.

Wallenius, T., Juvonen, S., Hansen, P., & Varjo, J. (2018). Schools, accountability and transparency—Approaching the Nordic school evaluation practices through discursive institutionalism. *Nordic Journal of Studies in Educational Policy, 4*(3), 133–143. https://doi.org/10.1080/20020317.2018.1537432

Wang, K., Chen, Z., Luo, W., Li, Y., & Waxman, H. (2018). Examining the differences between the job satisfaction of STEM and non-STEM novice teachers with leaving intentions. *EURASIA Journal of Mathematics, Science and Technology Education, 14*(6), 2329–2341. https://doi.org/10.29333/ejmste/89516

Wang, K., Li, Y., Luo, W., & Zhang, S. (2019). Selected factors contributing to teacher job satisfaction: A quantitative investigation using 2013 TALIS data. *Leadership and Policy in Schools*. https://doi.org/10.1080/15700763.2019.1586963

Wolfle, L. M. (1980). Strategies of path analysis. *American Educational Research Journal, 17*(2), 183–209. https://doi.org/10.3102/00028312017002183

Wollscheid, S., & Opheim, V. (2016). Knowledge brokering initiatives in education – A systematic map of the Nordic countries. *Nordic Journal of Studies in Educational Policy, 2016*(1), 31111. https://doi.org/10.3402/nstep.v2.31111

Zee, M., & Koomen, H. M. T. (2016). Teacher self-efficacy and its effects on classrooms processes, student academic adjustment, and teacher well-being: A synthesis of 40 years of research. *Review of Educational Research, 86*(4), 981–1015. https://doi.org/10.3102/0034654315626801

Chapter 6
Digital Inclusion in Norwegian and Danish Schools—Analysing Variation in Teachers' Collaboration, Attitudes, ICT Use and Students' ICT Literacy

Anubha Rohatgi, Jeppe Bundsgaard, and Ove E. Hatlevik

Abstract The capability to use digital technologies in an appropriate way has become a fundamental requirement of everyday life and wide adoption of digital technologies has gained a firm footing into the educational systems. Equity is a central goal in the Nordic model and ICT integration policies are warranted at the national level along with massive improvements in ICT infrastructures. The schools in their efforts towards realizing this objective have to integrate digital technology in teaching and learning in such a way that all children are given opportunities to participate in work, life and society. It is thus of interest to study the extent of digital inclusion, by examining the variation in computer and information literacy of students both within and between schools by addressing access and use of ICT in instruction among teachers. Data for the present study comes from 138 schools from Norway (2436 students, 1653 teachers) and 110 schools from Denmark (1767 students, 728 teachers) who took part in the International Computer and Information Literacy Study in 2013. Using a multilevel approach, variations at both levels in student computer and information literacy score and teacher collaboration in ICT use were examined. The results indicate that availability of digital technologies is a significant contributor towards student ICT achievement and teacher collaboration in both countries. There are small differences in computer and literacy score between the schools, while significant variations are noted between the students. Additionally,

A. Rohatgi (✉)
Department of Teacher Education and School Research, University of Oslo, Oslo, Norway
e-mail: anubha.rohatgi@ils.uio.no

J. Bundsgaard
Århus University, Aarhus, Denmark

O. E. Hatlevik
Oslo Metropolitan University, Oslo, Norway

teachers' attitudes are found to contribute significantly towards collaboration between teachers.

In light of digital inclusion, the successful and appropriate integration of information and communication technologies (ICT) in instruction has been acknowledged as a fundamental requirement across education systems worldwide. The manifestation of digital inclusion brings about equality/inclusion in strengthening the digital literacy required for educational achievement, future employment and social and economic development (Cha et al., 2011; Erstad, 2015; Livingston & Helsper, 2007; OECD, 2015). However, although digital inclusion keeps track of fast-changing and varied digital technologies, inclusion for all citizens still poses a challenge. This digital divide, which produces a participation gap, can be attributed to factors such as quality of ICT resources, extent of ICT usage, personal abilities/skills and variations in opportunities in terms of the frequency and complexity of tasks involving ICT (European Commission, 2013; Fraillon, Ainley, Schulz, Friedman, & Duckworth, 2019; Fraillon, Ainley, Schulz, Friedman, & Gebhardt, 2014; Hawkins & Oblinger, 2006).

In the same manner, despite in-depth investments in ICT resources and better ICT access, ensuring that all students and teachers make ideal use of ICT remains a challenge for educators and authorities. Notable variations in ICT use and proficiency, attitudes towards ICT and levels of achievement are still visible in ICT research (Fraillon et al., 2014; Vanderlinde, Aesaert, & Van Braak, 2014). The current situation resonates with the concerns raised in the past two decades – that students may experience different access to ICT (Pedró, 2007) and that a digital divide could appear (Scheerder, van Deursen, & van Dijk, 2017). Digital divides are related to the socio-economic background and the cultural differences between students in addition to the variation in cultural conditions between schools concerning how ICT is used in teaching and learning. To some extent, schools can be expected to reduce the digital divide by trying to ensure that both students and teachers receive equal opportunities to acquire ICT skills and benefit from ICT integration and high-quality digital teaching materials in the subjects rather than by merely amassing more ICT resources in the school (Bremholm & Bundsgaard, 2019; Gorski, 2002). However, further research is required on this topic.

The integration of ICT in schools does not by itself lead to more innovative practices (Bundsgaard, Pettersson, & Puck, 2014; Cuban, 2013). To create a more innovative teaching practice, teachers need to change a number of aspects of their thinking about teaching and learning, their planning and organisation of teaching and learning and the roles of both themselves and the students in everyday

classroom practices. Following this line of thought, teacher collaboration is vital because new practices emerge and grow from teamwork, cooperation and networking (Fredriksson, Jedeskog, & Plomp, 2008).

In the last 10–20 years, there have been rapid changes related to digital technology, visible both in the education system and in society. This advent of digital components has posed some new difficulties and challenges to fulfilling the idea of 'School for All', which is one of the building blocks in the structure of the education system of the Nordic countries (Buchholtz, Stuart, & Frønes, 2020). The introduction of digital technologies to the education system has led to concerns regarding whether and to what extent this introduction could lead to a digital divide (Dybkjær & Christensen, 1994; Warschauer, 2002). Attempts to bridge the digital divide by providing massive ICT resources (Gorski, 2009) do not guarantee that students also experience mastery in digital technologies. Within the Nordic model, in contrast to digital equality (i.e. all students and schools receive the same resources), digital equity as a qualitative property concerning justice allows for the targeted distribution of technology and support so that no child is left behind. Digital equity involves giving all students equal access and opportunities to develop their holistic ICT proficiency both within and outside the classroom.

From a government policy view in the Nordic countries, high-level ICT investments in education have been made. The efforts also include a revision of the curricula in Nordic countries, in general, in a manner where digital competence encompasses not only the competent use of digital tools but also broader societal issues and critical aspects in digital inclusion (Krumsvik, 2008).

Inequities in terms of the opportunities that the students have to learn and achieve are to be counteracted by providing sound ICT infrastructure and high-quality teaching and learning. Concerning digital equity, ICT resources are equally distributed among schools in both Norway and Denmark. However, the information collected on the quality of the current ICT resources or on how well the teachers can use ICT resources in their own teaching is still limited.

Schools in both Norway and Denmark are entitled to national elementary funding for ICT integration towards fulfilling the goal of achieving digital equality in national policies. However, individual variations reflecting diversity have been noted in the number of resources installed in different municipalities, thereby creating some formal barriers or 'inequality' (Volckmar, 2019).

Although great efforts are put into increasing the levels of ICT infrastructure, the evidence in empirical research about the positive influences of ICT on teaching, learning or teachers' professional development is limited (Cox et al., 2003; Ward & Parr, 2010). Reiterating Espinoza's (2007) thoughts on addressing inequality with changing procedures, schools would benefit by equipping teachers with better digital skills so that they can transfer these skills as part of their own teaching. Moreover, research on ICT in schools supports the notion that ICT tools for communication, information and collaboration can aid in enhancing school outcomes and the effectiveness of both the teaching and personal learning of teachers (Kozma, 2009). As such, a large body of research has dealt with the specific ICT competencies needed by teachers in their role as educators (Pettersson, 2018). It is reasonable to say that

teachers play an essential role in ICT integration and the implementation of necessary technology tools in instruction (Davis, Eickelmann, & Zaka, 2013; Pettersson, 2018).

The International Computer and Information Literacy Study (ICILS), designed by the International Association for the Evaluation of Educational Achievement (IEA), has measured the international differences in students' computer and information literacy (CIL) in Grade 8 (or its national equivalent). ICILS, in addition to student achievement, has collected contextual information at the student, teacher and school levels. One of the findings in the ICILS 2013 noted that ICT use in lessons was rather limited in most participating countries, except Denmark, although teachers showed positive attitudes towards ICT in teaching (Fraillon et al., 2014). In terms of the pedagogical aspects of ICT, teachers face new demands on a regular basis in their pursuit of acquiring new skills and pedagogical practices. For instance, teachers' ICT use for communication and information-sharing purposes is instrumental in strengthening certain ICT skills and expertise, but this type of use alone is not automatically sufficient for integrating ICT in pedagogical practices; thus, ICT needs to be incorporated into teacher education and professional development (Hatlevik, 2017).

Norwegian and Danish schools aim for all students to have the opportunity to develop themselves and their abilities. ICT integration policies are not only directed towards institutional levels in terms of improving infrastructures and resources but also directed towards supporting ICT integration in instructional practices within the organisation. However, ensuring digital inclusion can be a dilemma if there are major differences in teachers' pedagogical usage of ICT technologies both within and between schools. As mentioned earlier, teachers significantly influence their students' opportunities for equality and the extent to which students can reach their individual potential and attain the highest possible outcomes.

In the current chapter, we focus on digital inclusion by assessing the differences between schools in Norway and Denmark in relation to factors such as teachers' access to ICT, their use of ICT in instruction and their attitudes towards ICT. The data for our study is obtained from the ICILS 2013 cycle, and we use this data to examine the traces of digital inclusion in Norwegian and Danish schools. We also try to connect the responses from teachers at the school to student outcomes.

6.1 Theoretical Background

6.1.1 Digital Inclusion and the Use of ICT in Teaching Practices

In today's digital society, having access to Internet services and ICT devices in addition to opportunities for training and support for ICT integration are considered as defining elements for being digitally competent. As defined in *Building the digitally*

inclusive framework for digitally inclusive communities, 'Digital inclusion is the ability of individuals and groups to access and use information and communication technologies (ICT)' (IMLS et al., 2011, p. 1). This definition is extended as 'digital inclusion encompasses not only access to Internet but also the availability of hardware and software; relevant content and services; and training for the digital literacy skills required for effective use of information and communication technologies' (p. 1). In other words, for a teacher to become digitally competent, they would require not only access to ICT in terms of both quantity and quality of resources but also to accumulate wide and effective experience in ICT use. Digital equity is yet another concept often understood as a part of the digital inclusion route towards goals set for enhancing social and economic equity (Gorski, 2002; OECD, 2015). In the earlier definitions of digital inclusion, the dichotomy of ICT users vs. ICT non-users concerning the digital divide was widely considered. For instance, inequalities regarding ICT access and use have been shown to be dependent on both age and socioeconomic status (SES) but not as much on gender (Livingston & Helsper, 2007). In fact, the understanding of digital inclusion in recent studies encompasses not only gradations in both access and use of ICT technologies but also the attitudes and motivations of ICT users (Robinson et al., 2015).

Education systems worldwide acknowledge that teachers are the cornerstone in schools and are responsible for system-wide implementation (Hargreaves & Fullan, 2012; Hattie, 2009). This encourages the widespread adoption of ICT in schools aimed at the development of ICT skills across the entire teaching profession. In general, the process of ICT integration is targeted through increased ICT resources, curriculum priorities and teachers' professional development in schools. Research, however, has shown mixed reports regarding the relationship between the availability of ICT resources, ICT implementation in instruction, teachers' attitudes and teachers' professional development (Fraillon et al., 2014, 2019). Several studies have focused on teachers' pedagogical use of digital technologies in teaching and instruction (González-Sanmamed, Sangrà, & Muñoz-Carril, 2017; Prestridge, 2017) and the multidimensionality of ICT use in the classroom (Donnelly, McGarr, & O'Reilly, 2011). The results have highlighted a common characteristic among many European countries: Teachers seem to demonstrate a rather modest use of ICT for teaching purposes (Gill, Dalgarno, & Carlson, 2015; Haydn, 2014; Tondeur et al., 2015; Wastiau et al., 2013). In contrast, differences between European countries have also been noted. For example, Danish teachers report more frequent ICT use in teaching than Norwegian teachers or teachers from other countries (Fraillon et al., 2014). Recent research also reports differences between teachers regarding their attitudes towards ICT and what they believe about successful ICT use as part of their teaching practices (Haydn, 2014). Investments have been made in infrastructure, but these are insufficient. Providing training for selected teachers is necessary so that these teachers can be local supports for their colleagues.

6.1.2   *Digital Equality and Teacher Collaboration*

Equality, according to Corson (2001), implies sameness in general treatment. The concept of 'equality for all' mirrors that of equality of opportunity for all with the goal of ensuring that all individuals have the same amount of, and access to, resources without any political, legal, economic or social constraints (Espinoza, 2007). Using educational attainment as the output angle in light of digital inclusion, 'equality' means that all teachers have the same opportunities to use and master digital technology. In addition, equality means that each student receives the same opportunity to obtain the highest possible individual outcome (Ainscow & Miles, 2008; Espinoza, 2007). The digital divide highlights the opposite of digital equality in educational opportunities for all students, making it even more important to ensure digital equality in schools. Regarding teachers, this includes eliminating inequities as they attempt to learn to effectively use ICT coupled with the provision of access to ICT resources.

The equal distribution of ICT resources and other infrastructures represents a quantitative level of equality, whereas the concept of equity can be understood as the qualitative factor of providing 'just opportunities' for enhancing ICT competence and improving school outcomes (Espinoza, 2007). As part of the compensatory approach towards school effectiveness, it can be argued that teachers who collaborate in their ICT use not only improve their competence and ICT self-efficacy but also compensate for a lack of well-distributed resources or compensate for individual student characteristics (e.g. learning challenge) in their endeavour for equity. In other words, each student should benefit from a teacher possessing better ICT skills as part of within-school factors concerning policies and practices (Ainscow, Dyson, Goldrick, & West, 2016). Ainscow et al. (2016) further elaborated that 'the starting point for strengthening the capacity of a school to respond to learner diversity should be with the sharing of existing practices through collaboration amongst staff and joint practice development' (p. 149). Through this 'just distribution' of developing ICT skills in students and teachers, propositions of achieving equity can be envisaged.

Collaboration between teachers can facilitate the exploitation of both existing and new technologies in instructional practices and is an efficient tool for professional development (Bacigalupo & Cachia, 2011; Fogarty & Pete, 2010; McCormick, 2004). In addition to cultivating ICT use among their students as part of new literacy frameworks, teachers are regularly the 'learners' of new ICT and related tasks. Importantly, the precursor for optimal ICT implementation is when teachers experience a personal need for using ICT and feel digitally competent in their ability to effectively use ICT in instructional practice (Ward & Parr, 2010). In-house training and adoption of ICT-related practices within schools contribute to the development of teachers' own ICT competence and support the improvement of a student-oriented pedagogical approach (Drent & Meelissen, 2008; Egeberg et al., 2012; Fraillon et al., 2014; Wang, Hsu, Reeves, & Coster, 2014). However, the situation is dependent upon how much ICT is used in terms of time and access and upon how

well it is implemented as part of within-school teacher collaboration (Chapman & Fullan, 2007; Lindqvist, 2015).

6.1.3 Computer and Information Literacy (CIL)

Various terms are used to describe students' digital capabilities (Ala-Mutka, 2011) – for example, digital competence (Calvani, Fini, Ranieri, & Picci, 2012), ICT literacy (Erstad, 2006), digital literacy (Mioduser, Nachmias, & Forkosh-Baruch, 2008), CIL (Fraillon et al., 2014), twenty-first century skills (Binkley et al., 2012) and digital skills (Zhong, 2011). These terms describe successful ICT use as an independent and transversal learning area in addition to traditional subjects. They also encompass the combination of certain aspects of digital technologies (e.g. ICT, Internet and computer information) and the capability to benefit from the adopting digital technologies (e.g. skill, competence and literacy; Ferrari, 2012).

In the ICILS 2013 assessment framework, CIL is defined as the ability 'to use computers to investigate, create, and communicate in order to participate effectively' in various areas of life (Fraillon, Schulz, & Ainley, 2013, p. 17). Further, CIL is characterised by two overarching strands that are divided into seven content categories. Strand one is entitled *collecting and managing information*. This strand includes a practical understanding of how to use a computer and the capability to find and critically evaluate online information. Strand two of the framework, entitled *producing and exchanging information*, deals with the aspects of participating, producing and publishing using a computer as a tool. This strand comprises communication, safe use of information, secure use of information and transforming and creating digital information.

6.1.4 The Context of ICT in Norway and Denmark

Digital technology and digital inclusion have been on Norway's national education agenda for many years. At the end of the 1980s, ICT entered Norwegian secondary schools as an elective subject, and since the mid-1990s, national plans have included ICT in schools (Erstad, Kløvstad, Kristiansen, & Søby, 2005). There was also a focus on technology in the plan from 1996–1999 (Ministry of Education and Research, later in text MER, 1996), which included sub-areas such as 'learn to use', technical infrastructure, organisation and teacher education. Further, the national plan for 2000–2003 emphasised the educational use of ICT in schools (MER, 2000). During 2004–2008, the national ambition was to develop the digital competence of students and teachers (MER, 2004). This program overlapped with a curriculum reform, as the capability to use digital tools and resources was one of the five basic competence areas for all students (MER, 2006). In 2012, a framework outlining four areas of competence – *search and process*, *produce*, *communicate* and *digital*

responsibility – for digital skills was presented (The Norwegian Directorate for Education and Training, 2012). These four areas form the fundamental aspects of digital competence that teachers are expected to incorporate into their teaching to facilitate ICT literacy and ensure digital inclusion. As an equity aspect of the national educational plan, the policies state that every student should receive the same opportunity in a uniform school system. Nevertheless, the pedagogical use of ICT for teaching and learning varies between and within schools (Hatlevik & Christophersen, 2013; Hatlevik & Gudmundsdottir, 2013; Hatlevik, Guðmundsdóttir & Loi, 2015).

In Denmark, the integration of ICT in education has been on the national agenda for many years (Caeli & Bundsgaard, 2019). In the 1960s, the first Danish professor of computer science, Peter Naur, spoke in favour of creating a subject with a focus on both the critical understanding of the role of computers in society and the practical skills in the development of computer systems. In the 1970s, a subject was envisioned and ready to be introduced in schools, but a shift in the government stopped it. A similar subject was taught as an elective in the 1980s, computers were acquired, and numerous experiments using computers in teaching and learning were performed. In the 1990s, many government-initiated projects and experiments were conducted, the first wave of broad acquisition of hardware for schools took place, and schools began to become connected to the Internet through the so-called Sektornet, which was owned and maintained by the Ministry of Education until 2014 and provided connection to the Internet for educational institutions in Denmark. In the 2000s, a government funding scheme called ICT and Media in the Public Schools (ITMF or *IT og medier i Folkeskolen* in Danish) resulted in many local research and development projects concerning integrating ICT in teaching and learning. At the same time, massive investments were made in hardware, especially laptops for students and teachers and interactive whiteboards. Around 2010, the government funded laptops for all students in Grade 3 and supported the development of digital learning platforms that were expected to cover complete subjects. In particular, many municipalities and schools began investing in tablets (mostly iPads) for the students and teachers. From 2012 to 2017, the government and the Association of the Municipalities agreed to support the development of learning management platforms among other things. Schools were provided with funding for the acquisition of learning materials, with 50% of the expenses paving the way for the massive development of ICT and leading to the widespread use of ICT in everyday teaching and learning (Bremholm & Bundsgaard, 2019; Bundsgaard, Bindslev, Caeli, Pettersson, & Rusmann, 2019).

6.1.5   The Present Study

Under the broad definition of digital inclusion, this study addresses the diversity in teachers' use, access and attitudes towards ICT. To our knowledge, the assessment of variations in teacher variables in Norway and Denmark using a comparative

analysis approach is rather limited. Previous research has indicated that school-level characteristics, such as school ICT infrastructure/resources and policies related to ICT use, influence the extent to which teachers promote ICT integration in instruction. Moreover, teachers' positive attitudes towards ICT use and their ability to provide support to and receive support from colleagues are highlighted as important in the literature. Keeping this background in mind, we posited four hypotheses (H1–H4) in our study.

The first hypothesis (H1) relates to the variation in teachers' access to ICT, their use of ICT and their ICT attitudes:

H1 In both Norway and Denmark, there is variation between schools concerning teachers' self-reported ICT access, ICT use and ICT attitudes.

It is important that teachers experience equal opportunities to develop their ICT competence. Prior results have shown a positive association between school ICT resources and ICT integration (Fraillon et al., 2014, 2019). However, despite the availability of all-encompassing ICT resources, teachers' backgrounds (e.g. gender and age) play a central role in ensuring successful ICT implementation.

The second hypothesis (H2) aims to study the variation between teachers' background variables, their attitudes and their collaborative practices:

H2 Teachers' backgrounds and their ICT experiences, including a perceived lack of resources, can explain the variation in their teaching with ICT, their self-efficacy, their emphasis on developing ICT capabilities and their collaboration using ICT.

Collaboration between teachers using ICT is an essential characteristic of successful ICT use for teaching purposes. Furthermore, ICT resources play an important role in enhancing collaboration. Thus, our third hypothesis (H3) states the following:

H3 Teachers' backgrounds, ICT resources, ICT use and attitude variables (self-efficacy and views about ICT use) predict their collaboration with colleagues in the use of ICT.

Finally, to our knowledge, few studies have examined what teachers report about their ICT practices in relation to the digital achievement of the students. One could assume a positive relationship between what the teachers do and think on the one hand and the digital proficiency of the students on the other. This led to our fourth hypothesis (H4):

H4 Teachers' ICT use, attitudes towards ICT and perceived collaboration with colleagues predict variation in students' CIL.

6.2   Methods

6.2.1   International Computer and Information Literacy Study (ICILS) 2013

The ICILS 2013 collected data from both students and teachers across 21 participating education systems (Fraillon et al., 2013, 2014). A stratified two-stage probability cluster sampling design was used for school sample selection for all ICILS countries (Meinck, 2015). Both the students and teachers were randomly sampled from the selected schools, and the students participated in a CIL test in a computer-based environment in addition to completing a self-report questionnaire (including information about the students' background). For each student, only a subset of CIL items from a larger pool was administered to compensate for time constraints, with the intention of measuring students' broad CIL.

The ICILS assesses students' CIL using a purpose-designed computer-based test environment. The test comprises tasks (with many small tasks and one large task in each module) based on real-life themes. A proficiency scale describing four competence levels was developed based on a synthesis of typical elements of CIL content and item difficulties. Item Response theory was used to pair the scaled difficulty of each item with the item descriptor (Fraillon et al., 2014, p. 72). To estimate the standard errors possible for the derived statistical procedures (e.g. regression analysis), a plausible value method was used to derive five probable CIL achievement scores for each student, which were imputed based on the estimated latent student ability and responses to the background questionnaire. The ICILS 2013 data has been made publicly available by the IEA.[1]

Teacher participation in the study was voluntary. Teachers received a link to an online self-report questionnaire designed to be answered in about 30 min. For some questions, the teachers were asked to respond to the items about their background along with their views and attitudes in relation to a randomly selected reference class.

6.2.2   Study Sample

Data for the present study were obtained from the Norwegian and Danish samples. Both education systems are guided by the ambition for equalisation, for equal opportunities and that the school can counteract digital diversity among the students. In Norway, the sample comprised 2436 students and 1653 teachers in 138 schools; in Denmark, 1767 students and 728 teachers from 110 schools formed our sample. Because many teachers did not respond to all items, overall, samples

---

[1] https://www.iea.nl/data-tools/repository

comprising 1183 teachers in Norway and 722 teachers in Denmark were included in our analysis. The Norwegian sample comprised 63% female and 37% male participants, whereas 59% female and 41% male participants were included in the Danish sample. The teachers in Norway and Denmark were teaching two or more subjects. The majority of the teachers in Norway (68%) and Denmark (81%) taught test language or a foreign language subject.

6.2.3   Measures

To address our hypotheses, we used several constructs from the teacher data file, whereas the student CIL scores were obtained from the student data file. Teacher gender (coded as 0 for male and 1 for female), teacher age in actual years and teacher experience with ICT (T_EXPT) were used as background questions. Three options – 'Never' as (1), 'Fewer than two years' as (2) and 'Two years or more' as (3) – were used to code for how long the teachers had been using computers for teaching purposes. In questions related to ICT, teachers' ICT use, attitudes and views, the individual indices were scaled using IRT and Warm's weighted likelihood estimates (WLE). The scales presented in Table 6.1 were transformed to a mean of 50 points and a standard deviation of 10 points across participating countries. For details on the measures and scaling procedures, we kindly refer to Fraillon, Schulz, Friedman, Ainley, and Gebhardt (2015), Schulz and Ainley (2015) and Schulz and Friedman (2015).

6.2.4   Analytical Approaches

The information about variation between schools was extracted using the intraclass correlation (ICC; Geiser, 2012; Hox, 2013). The ICC provides a measure of between-school variation (how similar the groups are) in the outcome that is accounted for by the schools (McCoach & Adelson, 2010). In addition, we used multiple regression techniques to investigate the relative strengths of the association of the factors and multilevel structural equation modelling (SEM) on our data.

All analyses were conducted in the statistical package M*plus* 8.3 (Muthén & Muthén, 1998–2015). School identity was used as the cluster variable, and total teacher weight (TOTWGTT) was used in the M*plus* option for WEIGHT. To evaluate the fit of the structural equation models, common guidelines were applied (i.e. CFI ≥ .95, TLI ≥ .95, RMSEA ≤ .08 and SRMR ≤ .10) for an acceptable model fit (Marsh, Hau, & Grayson, 2005). The problem of missing data was resolved by data imputation. M*plus* uses multiple imputation (MI) for missing data using the full information maximum likelihood (FIML) approach. We used the robust maximum likelihood estimation (MLR), which accounted for the clustering of students in schools by correcting the standard errors in M*plus*.

Table 6.1 Measures from the ICILS 2013 used for the current study

Scale	# Items	Item stimulus	Item example	Response categories
Teachers' use of specific ICT applications. (T_USEAPP)	14	Use of ICT practices and activities in their teaching in their reference class.	'Word processors or presentation software'	4-point scale (from 'never' to 'in every or almost every lesson')
Teachers' use of ICT for learning at school. (T_USELRN)	13	Use of ICT for learning at school.	'Working on short assignments (i.e. within one week)'	3-point scale (from 'never' to 'often')
Teachers' use of ICT for teaching practices at school. (T_USETCH)	11	Use of ICT in a set of teaching practices.	'Presenting information through direct class instruction'	3-point scale (from 'never' to 'often')
Teachers' ICT self-efficacy. (T_EFF)	14	How confident they felt in diverse ICT-related tasks by themselves.	'Producing a letter using a word-processing program'	3-point scale ('I know how to do this', 'I could work out how to do this' and 'I do not think I could do this')
Teachers' emphasis on teaching ICT skills. (T_EMPH)	12	Emphasising the development of ICT-based capabilities in students in their reference class.	'Accessing information efficiently'	4-point scale (from 'strong emphasis' to 'no emphasis')
Teachers' positive views on using ICT in teaching and learning. (T_VWPOS)	8	Views on the positive outcomes of using ICT in teaching and learning.	'Enables students to access better sources of information'	4-point scale (from 'strongly agree' to 'strongly disagree')
Teachers' negative views on using ICT in teaching and learning. (T_VWNEG)	7	Views on the adverse outcomes of using ICT in teaching and learning.	'Only encourages copying material from published internet sources'	4-point scale (from 'strongly agree' to 'strongly disagree')
Teachers' perspectives on the lack of computer resources in school. (T_RESRC)	6	Perceptions on the lack of computer resources in school.	'My school does not have sufficient ICT equipment (e.g. computers)'	4-point scale (from 'strongly agree' to 'strongly disagree')
Collaboration between teachers in using ICT. (T_COLICT)	6	Perceptions on the collaborative practices with ICT use during teaching.	'I systematically collaborate with colleagues to develop ICT-based lessons based on the curriculum'	4-point scale (from 'strongly agree' to 'strongly disagree')

(continued)

Table 6.1 (continued)

Scale	# Items	Item stimulus	Item example	Response categories
CIL-achievement score (in plausible value (PV)).	5	CIL scores in ICILS 2013 had a mean of 500 points and a standard deviation of 100 points.	PV1CIL to PV5CIL (5 likely CIL proficiencies for students that attained each score)	

Note. Higher index values indicate higher frequency of use or higher levels of collaboration, except in the case of T_VWNEG and T_RESRC. See the supplementary material for details

6.3   Results

Based on our theoretical assumptions, we introduced four hypotheses regarding the use of ICT in school instruction and teacher collaboration. In this section, we first present the descriptive statistics highlighting the characteristics of the variables used in this study (Table 6.2), particularly reliability (Cronbach's alpha), indicating the internal consistency between the items in a scale. In the second section (Tables 6.3, 6.4, 6.5, and 6.6), the results of successive analyses are presented. Table 6.5 presents the results of the multiple regression analyses with collaboration as the dependent variable, addressing H3, whereas Table 6.6 presents the results of the multiple regression analysis with CIL as the dependent variable, addressing H4.

6.3.1   Summary of Scale Reliabilities, the Means and Standard Deviations

The reliabilities of the scales and descriptive statistics of the constructs in our study were examined before proceeding with other analyses. Regarding the scales' reliability (Table 6.2), almost all scales showed acceptable values above 0.80. Given that the means and standard deviations were internationally set at $M = 50$ and $SD = 10$, respectively, the Norwegian and Danish data do not show ceiling or floor effects.

6.3.2   Variation in Teachers' Self-Reported ICT Access, ICT Use and Their Attitudes (H1)

To study the variation between schools, ICC values were generated for the variables of concern in our study. Table 6.3 presents the results for the two countries.

Higher ICC values indicate a high degree of heterogeneity between schools (Geiser, 2012). In our results, the ICC values were low (ICC < 0.05) for the majority

Table 6.2 Scale Reliabilities and Descriptive Statistics for the Variables in Norway and Denmark

Variables	Norway			Denmark		
	α	M	SD	α	M	SD
Teachers' experience in using ICT for teaching purposes (T_EXPT)	–	1.92	0.30	–	1.95	0.22
Teachers' perspectives on the lack of computer resources (T_RESRC)	0.79	51.35	8.18	0.80	50.91	8.62
Teachers' perspectives on collaboration between teachers in using ICT (T_COLICT)	0.71	44.78	7.83	0.76	45.37	9.01
Teachers' use of ICT applications in teaching (T_USEAPP)	0.81	50.52	6.66	0.86	53.07	6.95
Teachers' use of ICT for learning at school (T_USELRN)	0.81	52.24	7.40	0.86	54.67	7.19
Teachers' use of ICT for teaching at school (T_USETCH)	0.89	51.21	7.49	0.89	53.56	7.51
Teachers' perceived self-efficacy in using ICT at school (T_EFF)	0.83	51.57	8.17	0.82	53.26	7.86
Teachers' emphasis on developing ICT-based capabilities (T_EMPH)	0.95	51.33	7.52	0.96	52.76	7.58
Teachers' positive views on using ICT in teaching and learning (T_VWPOS)	0.81	49.28	8.19	0.82	51.13	8.62
Teachers' negative views on using ICT in teaching and learning (T_VWNEG)	0.78	43.92	9.32	0.77	42.14	10.01

Note. T_EXPT is not used as a scale. All other scales are WLE = weighted mean likelihood estimate (Warm, 1989). SD = standard deviation, α = Cronbach's alpha

Table 6.3 Intraclass Correlation (ICC) for Teachers' Self-reported ICT Access, ICT Use and their ICT Attitudes in Norway and Denmark

Variables	Norway ICC	Denmark ICC
Teachers' experience in using ICT for teaching purposes (T_EXPT)	0.018	0.012
Teachers' perspectives on the lack of computer resources (T_RESRC)	0.280*	0.307*
Teachers' perspectives on collaboration between teachers in using ICT (T_COLICT)	0.090*	0.103*
Teachers' use of ICT applications in teaching (T_USEAPP)	0.018	0.072
Teachers' use of ICT for learning at school (T_USELRN)	0.057	0.060
Teachers' use of ICT for teaching at school (T_USETCH)	0.055	0.080
Teachers' perceived self-efficacy in using ICT at school (T_EFF)	0.007	0.043
Teachers' emphasis on developing ICT-based capabilities (T_EMPH)	0.011	0.047
Teachers' positive views on using ICT in teaching and learning (T_VWPOS)	0.020	0.026
Teachers' negative views on using ICT in teaching and learning (T_VWNEG)	0.033	0.038

Note. *p < .05, **p < .01

Table 6.4 Explained variance in different constructs using teachers' gender, age, experience with ICT and perceived lack of ICT resources for various purposes, attitudes and collaboration

	Norway	Denmark
	Beta (SE)	Beta (SE)
Teachers' use of ICT for learning at school		
Intercept	*6.62 (0.34)**	*0.97 (0.02)**
Gender	0.01 (0.04)	0.03 (0.05)
Age	−0.13 (0.03)**	0.03 (0.04)
Experience	0.19 (0.04)**	0.11 (0.05)*
Perceived lack of resources	−0.05 (0.05)	−0.13 (0.04)**
R-SQUARE	0.05 (0.02)	0.03 (0.02)*
Teachers' use of ICT applications in teaching		
Intercept	*6.91 (0.43)**	*7.33 (0.55)**
Gender	−0.02 (0.04)	−0.06 (0.05)
Age	−0.14 (0.03)**	0.02 (0.04)
Experience	0.21 (0.04)**	0.12 (0.04)**
Perceived lack of resources	−0.02 (0.05)	−0.13 (0.04)**
R-SQUARE	0.05 (0.02)**	0.04 (0.02)*
Teachers' use of ICT for teaching at school		
Intercept	*6.53 (0.41)*	*7.32 (0.47)**
Gender	0.05 (0.04)	0.01 (0.05)
Age	−0.15 (0.03)**	−0.03 (0.05)
Experience	0.18 (0.03)**	0.10 (0.04)*
Perceived lack of resources	−0.05 (0.05)	−0.16 (0.04)**
R-SQUARE	0.05 (0.01)*	0.03 (0.01)*
Teachers' perceived self-efficacy in using ICT at school		
Intercept	*7.86 (0.29)*	*7.49 (0.48)**
Gender	−0.10 (0.04)**	−0.29 (0.04)**
Age	−0.45 (0.04)**	−0.18 (0.04)**
Experience	0.10 (0.03)**	0.12 (0.03)**
Perceived lack of resources	−0.05 (0.03)	0.11 (0.05)*
R-SQUARE	0.20 (0.04)**	0.13 (0.03)**
Teachers' emphasis on developing ICT-based capabilities		
Intercept	*6.374 (0.31)**	*6.589 (0.60)**
Gender	0.13 (0.04)**	0.03 (0.04)
Age	−0.10 (0.03)**	0.07 (0.04)
Experience	0.167 (0.03)**	0.06 (0.04)
Perceived lack of resources	−0.057 (0.09)	−0.10 (0.04)**
R-SQUARE	0.05 (0.01)**	0.19 (0.01)
Teachers' views on collaboration between teachers		
Intercept	*7.13 (0.30)**	*6.55 (0.49)**
Gender	−0.00 (0.03)	0.08 (0.04)*
Age	0.07 (0.03)	0.08 (0.05)
Experience	0.05 (0.03)	0.05 (0.04)
Perceived lack of resources	−0.31 (0.04)**	−0.38 (0.04)**
R-SQUARE	0.11 (0.02)*	0.16 (0.03)**

Note. $*p < .05$, $**p < .01$

Table 6.5 Variations in teachers' views on collaboration in using ICT

	Norway Beta (SE)	Denmark Beta (SE)
Intercept (collaboration in using ICT)	*4.37 (0.54)*	*6.55 (0.49)***
Gender of teacher	−0.03 (0.03)	−0.05 (0.04)
Age of teacher	0.14 (0.04)**	0.08 (0.04)
Teacher experience with ICT	−0.01 (0.03)	0.01 (0.03)
Perceived lack of resources (ICT)	−0.26 (0.03)**	−0.33 (0.04)**
Teachers' use of ICT applications in teaching	−0.08 (0.06)	0.06 (0.09)
Teachers' use of ICT for learning at school	−0.02 (0.07)	0.13 (0.08)
Teachers' use of ICT for teaching at school	0.27 (0.09)**	−0.01 (0.08)
Teachers' perceived self-efficacy in using ICT at school	0.04 (0.04)	0.08 (0.04)
Teachers' emphasis on developing ICT-based capabilities	0.05 (0.07)	0.05 (0.06)
Teachers' positive views on using ICT in teaching and learning	0.15 (0.04)**	0.28 (0.05)**
Teachers' negative views on using ICT in teaching and learning	−0.01 (0.04)	0.13 (0.05)*
R-SQUARE	0.20 (0.03)**	0.31 (0.04)**

Note. *$p < .05$, **$p < .01$

Table 6.6 Multiple regressions with CIL as the dependent variable

	Norway Beta (SE)	Denmark Beta (SE)
Intercept (students' CIL score)	*20.39 (2.0)***	*18.49 (1.6)***
Gender of teacher	−0.02 (0.06)	−0.04 (0.05)
Age of teacher	0.06 (0.05)	0.10 (0.06)
Teacher experience with ICT	−0.02 (0.07)	−0.01 (0.04)
Perceived lack of resources (ICT)	−0.20 (0.08)*	−0.14 (0.07)*
Teachers' use of ICT applications in teaching	−0.21 (0.14)	−0.23 (0.11)*
Teachers' use of ICT for learning at school	0.19 (0.15)	0.30 (0.13)*
Teachers' use of ICT for teaching at school	0.06 (0.13)	0.09 (0.11)
Teachers' perceived self-efficacy in using ICT at school	−0.01 (0.06)	0.06 (0.07)
Teachers' emphasis on developing ICT-based capabilities	−0.08 (0.13)	−0.15 (0.11)
Teachers' positive views on using ICT in teaching and learning	−0.01 (0.08)	−0.02 (0.05)
Teachers' negative views on using ICT in teaching and learning	0.10 (0.06)	−0.04 (0.06)
Teachers' perspectives on collaboration between teachers in using ICT	−0.10 (0.08)	−0.08 (0.06)
R-SQUARE	0.06 (0.03)*	0.065 (0.03)

Note. * $p < .05$, ** $p < .01$

of the measures of access to ICT, use of ICT and ICT attitudes in both countries (see Table 6.2). This means that our assumption about variation in access, use, and attitudes did not hold true for use and attitude.

There were, however, some exceptions, revealing that the assumptions in H1 were valid for the lack of ICT resources and collaboration. In both Norway and Denmark, variations were found in teachers' views on the lack of ICT resources between schools (ICC = 0.28 and ICC = 0.307, respectively), indicating that almost

30% of the variation for this construct was found between schools. Variation between schools for the construct regarding views on teacher collaboration in using ICT was approximately 10% (ICC = 0.09 and ICC = 0.103, respectively).

Further, the ICC values from the Norwegian sample were slightly above 0.05 for the variables use of ICT for learning and use of ICT for teaching. In Denmark, the ICC values were between 0.06 and 0.08 for the variables use of ICT application, use of ICT for learning and use of ICT for teaching. This shows little variation across schools, which does not support H1. The small amount of variation between schools can be considered a problem for our statistical analyses. However, from the equity perspective, less variation between schools contradicts H1, and this can be used as an argument to support the claim of high degrees of equality between schools.

6.3.3 Variation in Teacher Self-Efficacy, Developing ICT Capabilities and Their Collaboration (H2)

In an attempt to study equality in teachers' experiences and collaborative practices in the frame of ICT, H2 addressed variations in teachers' teaching with ICT, their ICT self-efficacy, their collaboration with other teachers in using ICT and their emphasis on developing ICT-based capabilities using background variables in regression analyses. Table 6.4 presents the results for the two countries in terms of the beta values and standard errors.

In Norway, both age and experience with ICT showed a significant contribution to variation in the three different uses of ICT constructs (Table 6.4). However, the levels of explained variations were low (around 5%). Meanwhile, in Denmark, teachers' experience with ICT and their perceptions of the lack of ICT resources seemed to contribute to variation in the use of ICT for teaching at school. Regarding teacher self-efficacy, in both Norway and Denmark, gender (being male), age (being younger) and more experience with ICT significantly contributed to variation in teacher self-efficacy. Furthermore, in Denmark, teachers' perceptions of the lack of ICT resources seemed to have contributed to diversification. The explained variation was 20% in Norway and 13% in Denmark. Gender (being male), age (being younger) and more experience with ICT were also significant contributors to variations in teachers' emphasis on developing ICT-based capabilities in Norway. In contrast, in Denmark, only the perceived lack of ICT resources significantly contributed to the variance. The explained variance was 5% in Norway and 19% in Denmark.

While examining teachers' views on collaboration practices in ICT use, age and experience with ICT were not found to be significant predictors. Gender showed a weak contribution in the case of Denmark. Overall, teachers' views on the lack of ICT resources were a significant contributor in both Norway and Denmark. The explained variation was 11% in Norway and 16% in Denmark.

In both countries, H2 held for teachers' perceived self-efficacy in using ICT at school and teachers' views on collaboration between teachers. H2 also held for Danish teachers' emphasis on developing ICT-based capabilities, meaning that H2

did not have support when examining variation in teachers' use of ICT for learning at school, teachers' use of ICT applications in teaching and teachers' use of ICT for teaching at school.

6.3.4 Teacher Collaboration Predicts ICT Use and Teachers' Positive Views (H3)

Multiple regression analyses with collaboration as the dependent variable for the two countries were individually performed. The results are displayed in Table 6.5. All independent variables were simultaneously entered into the regression.

As seen in Table 6.5, no regular patterns were visible among the predictors for collaboration between teachers in either country. However, the perceived lack of ICT resources at school and teachers' positive views (for ICT use in instruction) played a significant role and had a relatively stable predictive power for teacher collaboration in both countries. The standardised regression coefficient weights in the case of Denmark were higher than those in the case of Norway for perceived lack of resources ($\beta = -0.26$ vs. $\beta = -0.33$) and teachers' positive views ($\beta = 0.15$ vs. $\beta = 0.28$). These fell into the medium effect size category (Cohen, 1988).

The age of the teacher ($\beta = 0.14$) and the use of ICT for teaching at school ($\beta = 0.27$) were significant contributors to the explained variance in Norway. In addition, teachers' negative views on using ICT in teaching and learning were substantial contributors in the case of Denmark ($\beta = 0.15$). The indicators under consideration provided different explanations, as reflected by the explained variances of the regression model. The model for Norway explained 20% of the variance compared with the model for Denmark, which had a variance of 31%. These findings support the assumption in H3 that there are variables and concepts that can explain the variance in teachers' collaboration using ICT.

6.3.5 Variation in CIL Score Using Teacher Variables (H4)

In our attempt to explain the variation in students' CIL scores using teacher variables through H4, were aggregated at the school level in this analysis. The results of the regression analyses for the two countries, with CIL score as the dependent variable and where the independent variables were simultaneously entered, are presented in Table 6.6 in terms of the beta values and their standard errors.

Overall, significant results ($p < 0.05$) were observed only for teachers' perceived lack of resources in both countries. The standardised regression coefficient weights in Norway were higher than those in Denmark for perceived lack of resources ($\beta = -0.20$ vs. $\beta = -0.14$). Two use variables, teachers' use of ICT applications in teaching and teachers' use of ICT for learning at school, contributed to the explained

variance in Denmark with values of $\beta = -0.23$ and $\beta = 0.30$, respectively. Concerning these two beta values, providing a clear explanation of why one was positive and the other was negative is difficult.

All the other regression coefficients were not statistically significant, leaving us with a low value of the explained variation in the CIL scores in both countries. From an equality perspective, we have identified variation in the CIL scores on the individual level; however, it does not seem that the difference between teachers' use of ICT and attitudes can explain sufficient variation. Another way to interpret this finding is that 'use of ICT' alone by teachers in a school does not necessarily lead to equality. Overall, these results do not support the assumption in H4 that teachers' use of ICT, their attitudes and perceived collaboration with colleagues can explain the variation in students' CIL scores. Although the results do not indicate that some schools work better with ICT than other schools in digital inclusion, this does not exclude that contextual and individual factors within the schools are important for equality and that all students have the opportunity to develop.

6.4   Discussion

The ICILS 2013 provides us with in-depth information on the factors related to ICT development at multiple levels along with international comparisons. The present contribution aims at highlighting the manner in which schools in the two Nordic countries are trying to bridge the achievement gaps within the frame of the respective ICT integration policies. We can draw several theoretically and practically important conclusions from our analyses using student achievement and teacher data (ICILS 2013) from Norway and Denmark. Concerning teachers' access to ICT, their use of ICT in instruction and their attitudes towards ICT at the school level, our study found no significant variation between the schools in Norway and Denmark. There was also no significant variation in teachers' use of ICT (application in teaching/for learning/for teaching at school). As one of our main findings, this lack of variation between schools seems to be an indicator of digital equality at the institutional level. The lack of variation between schools in these teacher variables, however, does not imply that no variation exists within the schools regarding teachers' access and use of ICT. Our subsequent findings suggest a particular structure of digital divide in Norwegian and Danish schools, and this inequality could be further highlighted by analysing the differences within schools and between individuals. In both Norway and Denmark, ICT is integrated as a learning dimension in all subjects, but it is up to the individual schools to implement the necessary practices for ICT integration. Irrespective of the local choices made, these practices are loyal to the national objectives.

Teachers' understanding of the initiatives taken by authorities, along with the concepts used to describe and assess student ability to use and succeed in using ICT (e.g. CIL, ICT literacy and digital competence), is multidimensional (Aesaert & van Braak, 2014). Thus, considering this multidimensionality, teachers might be

influenced while responding to the questionnaire items about the usefulness of ICT (Scherer, Siddiq, & Teo, 2015). In H1, we attempted to address the variation in teachers' views on the lack of ICT resources. Interestingly, variation was observed in teachers' views on the lack of ICT resources and teachers' collaboration between schools. ICT resources are presumed to be necessary for creating advantages in both student outcomes and staff attitudes (European Commission, 2013). Despite the high level of government ICT investments in education in both countries, some unequal distribution of these resources exists owing to geographical and other formal barriers (Volckmar, 2019). One explanation may be the local authority and responsibility for making the right choices and the priorities within the individual municipality and school. At the school level, the immediate responsibility for resource allocation and implementation of policies lies with the school staff. This implies that access to not only resources but also relevant knowledge is a prerequisite for schools attempting to achieve equity. Teachers who have reached a sufficient level of ICT self-efficacy are more likely to implement ICT into their teaching practices (Hatlevik, 2017). Teachers' personal ICT competence and attitudes (perceptions of their ICT skills) towards successful ICT implementation in instruction are strong predictors of their ICT use in teaching (Albion, Tondeur, Forkosh-Baruch, & Peeraer, 2015; Davis et al., 2013; Gerick, Eickelmann, & Bos, 2017; Ward & Parr, 2010).

In testing H2, we observed variation between schools in terms of teachers' collaborative practices. In both countries, there seemed to be less teacher collaboration with ICT use in schools where the teachers perceived a lack of ICT equipment and resources. When it comes to H2, the analyses revealed a more nuanced relationship. H2 did not hold when explaining the sufficient levels of variation in teachers' use of ICT for learning at school, teachers' use of ICT applications in teaching and teachers' use of ICT for teaching at school. However, the results for Denmark showed that teachers' backgrounds and their experience with ICT can explain variations in teachers' emphasis on developing ICT-based capabilities. Overall, the results showed that teachers' backgrounds and experience with ICT can explain the variations in their perceived self-efficacy and their views on collaboration between teachers. One way to interpret this is that there are no traces of inequality in teachers' use of ICT, but there are traces of digital inequality between teachers when it comes to their self-efficacy and views on collaboration. It certainly is important for teachers to gain experience with ICT to learn how to use ICT in general and to use ICT to teach and learn.

Gender was found to be a predictor of teachers' attitudes, which aligns with earlier research indicating that male teachers have higher ICT self-efficacy (Scherer et al., 2015; Wikan & Molster, 2011). Gender (being male; e.g. Broos, 2005), age (being younger) and more experience with ICT were also significant contributors to variations in teachers' emphasis on developing ICT-based capabilities in schools in Norway, thereby creating a slighter different profile from that of Denmark. A negative relation between teachers' age and perceptions of usefulness has also been noted in earlier studies (e.g. O'bannon & Thomas, 2014; Scherer et al., 2015; Vanderlinde et al., 2014). Our findings support existing research. It is only for

'perceived self-efficacy in using ICT at school' that gender, age and experience are significant in both countries. The main finding of this study is that gender, age and experience do not significantly explain the variation in teachers' attitudes and choices in these two Nordic countries. We can see diversity at the within-school level, and difference in treatment is required to create equal opportunities for all teachers in their ICT use in teaching. Equal distribution of ICT resources therefore might not be the best way to tackle the inequalities or diversity for creating equity in outcomes. Although equality could be achieved by sameness in treatment and the concept of justice, by overlooking individual factors and abilities, promoting equity is rather difficult. Typically, one would also expect that a lack of necessary ICT resources could help explain the variation in teachers' use of ICT, teachers' self-efficacy and their emphasis on developing ICT capabilities. Insufficient ICT equipment and a lack of technical and pedagogical support are pointed out as major hindrances in the effective use of ICT in teaching and learning (European Commission, 2013, p. 156). In the recently conducted ICILS 2018, although both school level and teacher data showed large differences in the availability of and appropriateness of ICT resources across countries, the teachers who were frequent ICT users in class were found to be more positive about teacher collaboration (Fraillon et al., 2019).

Concerning the explained variance in collaborative practices (H3), teachers' views on the lack of ICT resources and teachers' positive views on using ICT in teaching and learning were significant contributors in both countries. Overall, the results support this hypothesis, which indicates a lack of equity when it comes to experiencing collaboration. This means that some teachers experienced working in a supportive environment, whereas others experienced the opposite. Our assumption is that this variation provides teachers with different options and possibilities in terms of discussing ICT teaching and searching for support from colleagues. In-house training and adoption of ICT-related practices within schools contribute to the development of teachers' own ICT competence and support the improvement of a student-oriented pedagogical approach (Drent & Meelissen, 2008; Egeberg et al., 2012; Fraillon et al., 2014; Wang et al., 2014). However, notably, the situation is dependent on how much ICT is being used in terms of time and access and how well it is implemented in terms of teacher collaboration within schools (Fullan, 2007; Lindqvist, 2015).

When studying the contributors to variation in students' ICT literacy (CIL) scores, teachers' perceptions of a lack of ICT resources were found to be directly related to ICT literacy in both countries. This finding resonates with the fact that sufficient ICT resources along with technical support are key elements for ICT implementation in classrooms (European Commission, 2013). Nevertheless, it is pertinent that overall ICT investments also encompass areas such as teacher training and pedagogical support and do not only focus on material resources from higher levels of government.

As stated in H4, we expected to identify teacher collaboration as a significant contributor to student CIL scores. However, this was not revealed in our results. One explanation for this could be drawn from the ICILS study sampling design, in which

15 teachers were selected at random from all teachers teaching the target grade at each school (Fraillon et al., 2014). A second factor leading to the low collaboration finding could be that the sampled teachers were from different disciplines; therefore, they were not prone to collaboration in their teaching of the subject, possibly ignoring the need to use ICT (Wikan & Molster, 2011). The obvious benefit for teachers lies in making the best use of innovations in a collaborative environment and in developing their shared understanding. Vrasidas (2015) also reported that more than two-thirds of participating teachers who were provided with opportunities to learn from each other and collaborate with experts felt more prepared to integrate ICT in their classrooms. This highlights the potential importance of collaboration among teachers in terms of informal learning opportunities – for example, observing how other teachers use ICT in teaching as part of technology integration and teachers' professional development (Fraillon et al., 2014, 2019).

Among the attitude indicators, teachers' positive views about ICT use in instruction were significant predictors of collaboration in ICT use both in Norway and Denmark. Teachers with negative views towards ICT use in instructional practices or lower ICT self-efficacy may find collaborating with other advanced ICT users among their peers rather challenging. Furthermore, the absence of clear guidelines and school policies regarding ICT and teachers' characteristics and attitudes could play an important role in how their collaboration manifests in instructional practices.

Overall, variation was found between students concerning their CIL scores; however, when scrutinising the available variables from the survey, we did not identify any teacher variables that could explain sufficient levels of variance in the CIL scores. Our study cannot exclude the existence of the digital divide at the school or system level, but the most clear and comprehensive digital inequality was identified at the individual level. It seems, therefore, that the variance identified can be explained by the variance between students and not between schools. Krumsvik (2011) emphasised the importance of teachers using technology in instruction so that their students can achieve the digital competence aims set in the curriculum. The challenge is finding solutions that facilitate the equity of both access and use of ICT within schools by addressing the observed discrepancies in teachers' use of ICT in instruction.

6.4.1   Digital Inclusion/Equity

From a government perspective, ICT resources are intentionally distributed equally among schools, representing a step towards accomplishing digital equity. However, one could assume that factors of individual teacher, such as teacher competencies, teacher perceptions and their attitudes, might contribute to an extent towards inequality within schools. Haydn (2014) found that some teachers appear to be experts, whereas others have less expertise. In addition, as a guiding thought, providing teachers with support and appropriate pedagogical development is as important as ensuring ICT provision and support (European Commission, 2013, p. 156)

and should be prioritised. The ICILS 2018 reported that, across participating countries, teachers show higher usage levels of digital tools with general utility in classrooms than advanced digital learning tools (Fraillon et al., 2019). Without formal training courses in new digital technologies, much depends on the ability, compounded by the willingness, of the teachers to integrate ICT into instruction. At the individual level, teachers' personal and technology-related characteristics (e.g. prior experience with ICT and attitudes) play an important role in strengthening teachers' professional development involving ICT use in instruction (Gil-Flores, Rodríguez-Santero, & Torres-Gordillo, 2017). At the institutional level, aspects such as school policies concerning resource allocation, technology initiatives and revised strategies to support quality instruction and learning using ICT play a vital role in digital inclusion. For instance, the implementation of this institutional endeavour is reflected in Denmark, where a very high percentage of teachers report participating in professional development courses (ICILS 2018). In the case of scholarships for Norwegian teachers who pursue further education, the subjects mathematics, English, Norwegian, Sami and Norwegian sign language are given priority (The Norwegian Directorate for Education and Training, 2020). Among the 5775 teachers in 2020/2021, who are offered a scholarship or extra funding so that they use substitute teacher, only 419 teachers are given funds to study programming or professional digital competence.

Teachers' perceptions of the benefits of using ICT might be different from their actual perceptions of ICT use in instruction with respect to the problems and obstacles in the use of ICT in instruction (Carstens & Pelgrum, 2009). Therefore, it is essential for teachers to develop an updated teaching practice including optimal pedagogical use of ICT that supports not only students' learning processes but also their expertise in ICT literacy. Digital inclusion in schools would further be enhanced by constant efforts in meeting the ever-changing targets (e.g. resources) and by means of helping teachers become at ease and experienced in using ICT as part of their teaching.

Our analyses show that in most of the phenomena measured in the teacher survey in ICILS (related to both teachers' experience of using ICT, their views on ICT in teaching and learning, and their use of ICT in their teaching), there is little variation across schools in both Norway and Denmark. The small variation between schools is a challenge for the statistical analyses. However, from the equity perspective, less variation between schools supports the claim of high degrees of equality between schools. We consider this as an indicator of digital equality at the institutional level in both Norway and Denmark. Regarding students' CIL achievement, the main source of variance is not found at the school level but at the individual level, meaning that in these countries, observed equity is promoted more at the institutional level than at the individual level.

6.4.2   Limitations and Future Directions

Owing to the sampling design, the study did not provide a direct opportunity to connect either the students or teachers to a particular class (e.g. in Trends in International Mathematics and Science Study (TIMSS), an entire class is sampled, and one teacher per subject answers the teacher questionnaire; Martin, Mullis, Foy, & Hooper, 2016). This poses a clear limitation to our study and to understanding the relationship between teacher characteristics and students' ICT literacy. We attempted to aggregate the student scores at the school level and to distribute them to all teachers alike. Because ICT is integrated into all subjects and not treated as a specific subject, another limitation could be the ICILS test being too general and not directly related to student achievement in particular subject domains, although administering the self-report questionnaire to a large group of teachers gave us better knowledge of the teacher population. With the intention that teacher information should not be linked to individual students, a random sample of 15 teachers in schools with 21 or more teachers teaching the target grade regardless of the subject they taught was included in the ICILS (Fraillon et al., 2014, p. 34). This increased the complexity of situation because whether these teachers taught the students the years before remains unclear. We primarily relied upon the teachers' self-reports in our analyses and also did not test for measurement invariance to prove the equivalence of teacher views/beliefs between the two countries.

   The data used in this study originate from 2013, and there is a need for further research on the topic. The second round of ICILS was conducted in 2018, but only data for Denmark is available because Norway did not participate in the ICILS 2018 cycle. In looking at the trend data for Denmark, the use of ICT in teaching has increased from 2013 to 2018. For instance, in 2018, 72% of the teachers reported using ICT on a daily basis, whereas this number was 40% in 2013 (Bundsgaard et al., 2019). The Danish teachers also reported significant changes in the degree to which they emphasised teaching in CIL-relevant topics, and they were even more self-confident in using ICT in 2018 than in 2013. The forthcoming ICILS 2023 will provide opportunities to further examine what characterises digital diversity in Norwegian and Danish schools. In addition, the study will provide an opportunity to examine the developments from 2013 to 2023 in both countries.

6.5   Conclusion

This study aimed to examine teachers' access to ICT, use of ICT in instruction, perceptions of lack of resources, attitudes towards ICT and collaborative practices. Teachers' perceptions of a lack of ICT resources in schools hinder the effective implementation of ICT in Denmark and Norway. However, equipping schools with ICT resources alone without a more holistic approach is unlikely to be productive in the development of ICT skills and knowledge.

Although some variation between schools was visible in ICT-related teacher measures, the school systems and administrators play a significant role in transforming practices and policies designed for encouraging the use of ICT in instructional practices. According to Cox et al. (2003), teachers are critical concerning the use of ICT because it defines not only the type of resources incorporated but also how those resources are used within classroom activities and during class time. In addition, when appropriate technological resources for each discipline are used, positive effects on learning can be anticipated because the availability of ICT equipment allows for its more frequent use by teachers. Teachers need to work in supportive environments where, aside from warranting access to new technologies, ICT implementation is seen as integral and relevant to achieving educational goals.

Our results suggest that, first, education systems need to focus on direct resourcing (ICT) to schools with larger needs for ICT resources. Second, setting concrete targets for achieving more equity by promoting and facilitating the extensive and consistent use of ICT by teachers, particularly in their instructional practices, should be considered a priority. Finally, the importance of teachers' (and schools') roles in promoting equity should be highlighted by setting concrete targets for equipping teachers with better ICT skills and enhancing their competence in transferring these skills to both students and colleagues.

Appendix

The section 'ICT and teaching in your school' in the ICILS 2013 teacher questionnaire had the following items.

Teachers' Use of Specific ICT Applications (T_USEAPP)

Q. How often did you use the following tools in your teaching of the reference class this school year?

('Never', 'In some lessons', 'In most lessons' and 'In every or almost every lesson')

1. Tutorial software or [practice programs]
2. Digital learning games
3. Word processors or presentation software (e.g. [Microsoft Word®] and [Microsoft PowerPoint®])
4. Spreadsheets (e.g. [Microsoft Excel®])
5. Multimedia production tools (e.g. media capture and editing and web production)

6. Concept-mapping software (e.g. [Inspiration®] and [Webspiration®])
7. Data logging and monitoring tools
8. Simulations and modelling software
9. Social media (e.g. Facebook and Twitter)
10. Communication software (e.g. email and blogs)
11. Computer-based information resources (e.g. websites, wikis and encyclopaedias)
12. Interactive digital learning resources (e.g. learning objects)
13. Graphing or drawing software
14. E-portfolios

Teachers' Use of ICT for Learning (T_USELRN)

Q. How often does your reference class use ICT in the following activities?
  ('Never', 'Sometimes' and 'Often')

1. Working on extended projects (i.e. over several weeks)
2. Working on short assignments (i.e. within one week)
3. Explaining and discussing ideas with other students
4. Submitting completed work for assessment
5. Working individually on learning materials at their own pace
6. Undertaking open-ended investigations or field work
7. Reflecting on their learning experiences (e.g. using a learning log)
8. Communicating with students in other schools on projects
9. Seeking information from experts outside the school
10. Planning a sequence of learning activities for themselves
11. Processing and analysing data
12. Searching for information on a topic using outside resources
13. Evaluating information resulting from a search

Teachers' Use of ICT in Teaching Practices (T_USETCH)

Q. How often do you use ICT in the following practices when teaching your reference class?
  ('Never', 'Sometimes' and 'Often')

1. Presenting information through direct class instruction
2. Providing remedial or enrichment support to individual students or small groups of students
3. Enabling student-led whole-class discussions and presentations
4. Assessing students' learning through tests
5. Providing feedback to students

6. Reinforcing learning of skills through repetition of examples
7. Supporting collaboration among students
8. Mediating communication between students and experts or external mentors
9. Enabling students to collaborate with other students (within or outside school)
10. Collaborating with parents or guardians in supporting students' learning
11. Supporting inquiry learning

Teachers' ICT Self-Efficacy (T_EFF)

Q. How well can you do these tasks on a computer by yourself?
   ('I know how to do this', 'I could work out how to do this' and 'I do not think I could do this')

1. Producing a letter using a word processing program
2. Emailing a file as an attachment
3. Storing your digital photos on a computer
4. Filing digital documents in folders and subfolders
5. Monitoring students' progress
6. Using a spreadsheet program (e.g. [Lotus 1 2 3®, Microsoft Excel®]) for keeping records or analysing data
7. Contributing to a discussion forum/user group on the Internet (e.g. a wiki or blog)
8. Producing presentations (e.g. [PowerPoint® or a similar program]) with simple animation functions
9. Using the Internet for online purchases and payments
10. Preparing lessons that involve the use of ICT by students
11. Finding useful teaching resources on the Internet
12. Assessing student learning
13. Collaborating with others using shared resources such as [Google Docs®]
14. Installing software

Teachers' Emphasis on Teaching ICT Skills (T_EMPH)

Q. In your teaching of the reference class in this school year, how much emphasis have you given to developing the following ICT-based capabilities in your students?
   ('Strong emphasis', 'Some emphasis', 'Little emphasis' and 'No emphasis')

1. Accessing information efficiently
2. Evaluating the relevance of digital information
3. Displaying information for a given audience/purpose
4. Evaluating the credibility of digital information
5. Validating the accuracy of digital information

6. Sharing digital information with others
7. Using computer software to construct digital work products (e.g. presentations, documents, images and diagrams)
8. Evaluating students' approach to information searches
9. Providing digital feedback on the work of others (such as classmates)
10. Exploring a range of digital resources when searching for information
11. Providing references for digital information sources
12. Understanding the consequences of making information publically available online

Teachers' Positive Views on Using ICT in Teaching and Learning (T_VWPOS)

Q. To what extent do you agree or disagree with the following statements about using ICT in teaching and learning at school?
   ('Strongly agree', 'Agree', 'Disagree' and 'Strongly disagree')

1. Enables students to access better sources of information
2. Helps students to more effectively consolidate and process information
3. Helps students to learn to collaborate with other students
4. Enables students to more effectively communicate with others
5. Helps students to develop greater interest in learning
6. Helps students to work at a level appropriate to their learning needs
7. Helps students to develop skills in planning and self-regulation of their work
8. Improves the academic performance of students

Teachers' Negative Views on Using ICT in Teaching and Learning (T_VWNEG)

Q. To what extent do you agree or disagree with the following statements about using ICT in teaching and learning at school?
   ('Strongly agree', 'Agree', 'Disagree' and 'Strongly disagree')

1. Results in poorer writing skills among students
2. Only introduces organisational problems for schools
3. Impedes concept formation, which is better done with real objects than with computer images
4. Only encourages copying material from published Internet sources
5. Limits the amount of personal communication among students
6. Results in poorer calculation and estimation skills among students
7. Only distracts students from learning

Teachers' Lack of Computer Resources at School

(T_RESRC). Scale on six out of eight items.

Q. To what extent do you agree or disagree with the following statements about the use of ICT in teaching at your school?

('Strongly agree', 'Agree', 'Disagree' and 'Strongly disagree')

1. My school does not have sufficient ICT equipment (e.g. computers).
2. My school does not have access to digital learning resources.
3. My school has limited connectivity (e.g. slow or unstable speed) to the Internet.
4. The computer equipment in our school is out-of-date.
5. There is insufficient provision for me to develop expertise in ICT.
6. There is insufficient technical support to maintain ICT resources.

Teachers' Collaboration in Using ICT (T_COLICT)

Q. To what extent do you agree or disagree with the following practices and principles in relation to the use of ICT in teaching and learning?

('Strongly agree', 'Agree', 'Disagree' and 'Strongly disagree')

1. I work together with other teachers on improving the use of ICT in classroom teaching.
2. There is a common set of rules in the school about how ICT should be used in classrooms.
3. I systematically collaborate with colleagues to develop ICT-based lessons based on the curriculum.
4. I observe how other teachers use ICT in teaching.
5. There is a common set of expectations in the school about what students will learn about ICT.

References

Aesaert, K., & van Braak, J. (2014). Exploring factors related to primary school pupils' ICT self-efficacy: A multilevel approach. *Computers in Human Behavior, 41*, 327–341.

Ainscow, M., Dyson, A., Goldrick, S., & West, M. (2016). Using collaborative inquiry to foster equity within school systems: Opportunities and barriers. *School Effectiveness and School Improvement, 27*(1), 7–23.

Ainscow, M., & Miles, S. (2008). Making education for all inclusive: Where next? *Prospects, 38*(1), 15–34.

Ala-Mutka, K. (2011). *Mapping digital competence: Towards a conceptual understanding.* Luxembourg: European Union.

Albion, P. R., Tondeur, J., Forkosh-Baruch, A., & Peeraer, J. (2015). Teachers' professional development for ICT integration: Towards a reciprocal relationship between research and practice. *Education and Information Technologies, 20*(4), 655–673.

Bacigalupo, M., & Cachia, R. (2011). *Teacher Collaboration Networks in 2025. What is the role of teacher networks for professional development in Europe*. Notes from the Workshops held on June 6th and 7th at the Institute for Prospective Technological Studies of the European Commission Joint Research Centre.

Binkley, M., Erstad, E., Herman, J., Raizen, S., Ripley, M., & Miller-Ricci, M. (2012). Defining 21st century skills. In P. Griffin, B. McGaw, & E. Care (Eds.), *Assessment and teaching of 21st century skills* (p. 17e66). Dordrecht, The Netherlands: Springer. https://doi.org/10.1007/978-94-007-2324-5

Bremholm, J., & Bundsgaard, J. (2019). Danske elevers brug af it i skolen. In V. T. Christensen (Ed.), *PISA 2018: Danske Unge i en international sammenligning* (pp. 113–126). VIVE - Viden til velfærd.

Broos, A. (2005). Gender and information and communication technologies (ICT) anxiety: Male self-assurance and female hesitation. *Cyberpsychology & Behavior, 8*(1), 21–31.

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* (pp. 13–41*)*. Cham, Switzerland: Springer.

Bundsgaard, J., Bindslev, S. G., Caeli, E. N., Pettersson, M., & Rusmann, A. (2019). *Danske elevers teknologiforståelse: resultater fra ICILS-undersøgelsen 2018* (1st ed.). Aarhus, Denmark: Aarhus Universitetsforlag.

Bundsgaard, J., Pettersson, M., & Puck, M. R. (2014). *Digitale kompetencer. It i danske skoler i et internationalt perspektiv*. Aarhus, Denmark: Aarhus Universitetsforlag.

Caeli, E. N., & Bundsgaard, J. (2019). Datalogisk tænkning og teknologiforståelse i folkeskolen tur-retur. *Læring og Medier, 11*(19). https://doi.org/10.7146/lom.v11i19.110919

Calvani, A., Fini, A., Ranieri, M., & Picci, P. (2012). Are young generations in secondary school digitally competent? A study on Italian teenagers. *Computers & Education, 58*, 797–807.

Carstens, R., & Pelgrum, W. J. (2009). *Second information technology in education study*. SITES 2006 technical report.

Cha, S. E., Jun, S. J., Kwon, D. Y., Kim, H. S., Kim, S. B., Kim, J. M., et al. (2011). Measuring achievement of ICT competency for students in Korea. *Computers & Education, 56*(4), 990–1002. https://doi.org/10.1016/j.compedu.2010.11.003

Chapman, C., & Fullan, M. (2007). Collaboration and partnership for equitable improvement: Towards a networked learning system? *School Leadership & Management, 27*(3), 207–211. https://doi.org/10.1080/13632430701379354

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Abingdon, UK: Routledge Academic.

Corson, D. (2001). Ontario students as a means to a government's ends. *Our Schools/Our Selves, 10*(4), 55–77.

Cox, M., Webb, M., Abbott, C., Blakeley, B., Beauchamp, T., & Rhodes, V. (2003). *A review of the research literature relating to ICT and attainment*. https://dera.ioe.ac.uk/1600/1/becta_2003_attainmentreview_queensprinter.pdf

Cuban, L. (2013). *Inside the black box of classroom practice: Change without reform in American education*. Cambridge, MA: Harvard Education Press.

Davis, N., Eickelmann, B., & Zaka, P. (2013). Restructuring of educational systems in the digital age from a co-evolutionary perspective. *Journal of Computer Assisted Learning, 29*(5), 438–450.

Donnelly, D., McGarr, O., & O'Reilly, J. (2011). A framework for teachers' integration of ICT into their classroom practice. *Computers & Education, 57*(2), 1469–1483.

Drent, M., & Meelissen, M. (2008). Which factors obstruct or stimulate teacher educators to use ICT innovatively? *Computers & Education, 51*(1), 187–199.

Dybkjær, L., & Christensen, S. (1994). *Infosamfundet år 2000 – Rapport fra udvalget om "Informationssamfundet år 2000"*. Copenhagen, Denmark: Forskningsministeriet.

Egeberg, G., Gudmundsdottir, G. B., Hatlevik, O. E., Ottestad, G., Skaug, J. H., & Tømte, K. (2012). Monitor 2011. *Skolens digitale tilstand*. Senter for IKT i utdanningen.

Erstad, O. (2006). A new direction? Digital literacy, student participation and curriculum reform in Norway. *Education & Information Technologies, 11*, 415–429.

Erstad, O. (2015). Educating the digital generation-exploring media literacy for the 21st century. *Nordic Journal of Digital Literacy, 10*(Jubileumsnummer), 85–102.

Erstad, O., Kløvstad, V., Kristiansen, T., & Søby, M. (2005). *På vei mot digital kompetanse i grunnopplæringen*. Oslo, Norway: Universitetsforlaget.

Espinoza, O. (2007). Solving the equity – Equality conceptual dilemma: A new model for analysis of the educational process. *Educational Research, 49*(4), 343–363.

European Commission. (2013). *Survey of schools: ICT in education*. https://ec.europa.eu/digital-single-market/en/news/survey-schools-ict-education

Ferrari, A. (2012). Digital competence in practice: An analysis of frameworks. *JRC IPTS*. https://doi.org/10.2791/82116

Fogarty, R., & Pete, B. M. (2010). The Singapore vision: Teach less, learn more. In *21st century skills: Rethinking how students learn* (pp. 97–116). Bloomington, IN: Solution Tree Press.

Fraillon, J., Ainley, J., Schulz, W., Friedman, T., & Duckworth, D. (2019). *Preparing for life in a digital world: The IEA international computer and information literacy study 2018 international report*. Cham, The Netherlands: Springer Nature.

Fraillon, J., Ainley, J., Schulz, W., Friedman, T., & Gebhardt, E. (2014). *Preparing for life in a digital age: The IEA international computer and information literacy study international report*. Cham, The Netherlands: Springer Open.

Fraillon, J., Schulz, W., & Ainley, J. (2013). *International computer and information literacy study: Assessment framework*. International Association for the Evaluation of Educational Achievement. Herengracht 487, Amsterdam, 1017 BT, The Netherlands.

Fraillon, J., Schulz, W., Friedman, T., Ainley, J., & Gebhardt, E. (2015). *ICILS 2013 technical report*. International Association for the Evaluation of Educational Achievement (IEA). https://www.iea.nl/publications/technical-reports/icils-2013-technical-report

Fredriksson, U., Jedeskog, G., & Plomp, T. (2008). Innovative use of ICT in schools based on the findings in ELFE project. *Education and Information Technologies, 13*(2), 83.

Fullan, M. (2007). Change the terms for teacher learning. *The Learning Professional, 28*(3), 35.

Geiser, C. (2012). *Data analysis with Mplus*. New York: Guilford Press.

Gerick, J., Eickelmann, B., & Bos, W. (2017). School-level predictors for the use of ICT in schools and students' CIL in international comparison. *Large-scale Assessments in Education, 5*(1), *5*.

Gil-Flores, J., Rodríguez-Santero, J., & Torres-Gordillo, J. J. (2017). Factors that explain the use of ICT in secondary-education classrooms: The role of teacher characteristics and school infrastructure. *Computers in Human Behavior, 68*, 441–449.

Gill, L., Dalgarno, B., & Carlson, L. (2015). How does pre-service teacher preparedness to use ICTs for learning and teaching develop through their degree program? *Australian Journal of Teacher Education (Online), 40*(1), 36.

González-Sanmamed, M., Sangrà, A., & Muñoz-Carril, P. C. (2017). We can, we know how. But do we want to? Teaching attitudes towards ICT based on the level of technology integration in schools. *Technology, Pedagogy and Education, 26*(5), 633–647.

Gorski, P. C. (2002). How computers STILL cheat girls. *The Education Digest, 68*(3), 23–27.

Gorski, P. C. (2009). Insisting on digital equity: Reframing the dominant discourse on multicultural education and technology. *Urban Education, 44*(3), 348–364. https://doi.org/10.1177/0042085908318712

Hargreaves, A., & Fullan, M. (2012). *Professional capital: Transforming teaching in every school*. New York: Teachers College Press.

Hatlevik, O. E. (2017). Examining the relationship between teachers' self-efficacy, their digital competence, strategies to evaluate information, and use of ICT at school. *Scandinavian Journal of Educational Research, 61*(5), 555–567.

Hatlevik, O. E., & Christophersen, K. S. (2013). Digital competence in the beginning of upper secondary school: Digital inclusion. *Computers & Education*, 240–247. https://doi.org/10.1016/j.compedu.2012.11.015

Hatlevik, O. E., & Gudmundsdottir, G. B. (2013, April 1). An emerging digital divide in urban school children's information literacy: Challenging equity in the Norwegian school system. *First Monday., 18*(4).

Hatlevik, O. E., Guðmundsdóttir, G. B., & Loi, M. (2015). Digital diversity among upper secondary students: A multilevel analysis of the relationship between cultural capital, self-efficacy, strategic use of information and digital competence. *Computers & Education, 81*, 345–353.

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

Hawkins, B. L., & Oblinger, D. G. (2006). The myth about putting information online: No one cares what you say online. *Educause Review, 41*(5), 14–15.

Haydn, T. (2014). How do you get pre-service teachers to become 'good at ICT' in their subject teaching? The views of expert practitioners. *Technology, Pedagogy and Education, 23*(4), 455–469. https://doi.org/10.1080/1475939X.2014.892898

Hox, J. J. (2013). Multilevel regression and multilevel structural equation modelling. *The Oxford Handbook of Quantitative Methods, 2*(1), 281–294.

Institute of Museum and Library Services [IMLS]. (2011). *Building digitally inclusive communities: A guide to the proposed framework*. Washington, DC: Institute of Museum and Library Services.

Kozma, R. B. (2009). The knowledge ladder: Using ICT and education reform to advance economic and social development goals. In *Symposium on Education Technology in Schools: Converging Innovation and Creativity*.

Krumsvik, R. J. (2008). Situated learning and teachers' digital competence. *Education and Information Technologies, 13*(4), 279–290.

Krumsvik, R. J. (2011). Digital competence in the Norwegian teacher education and schools. *Högre utbildning, 1*(1), 39–51.

Lindqvist, M. J. H. (2015). Exploring activities regarding technology-enhanced learning in a one-to-one initiative. *Nordic Journal of Digital Literacy, 10*(4), 227–245.

Livingston, S., & Helsper, E. (2007). Gradations in digital inclusion: Children, young people and the digital divide. *New Media Society, 9*(4), 671–696.

Marsh, H. W., Hau, K. T., & Grayson, D. (2005). Goodness of fit in structural equation models. In A. Maydeu-Olivares & J. J. McArdle (Eds.), *Multivariate applications book series*. Mahwah, NJ: Erlbaum.

Martin, M. O., Mullis, I. V. S., Foy, P., & Hooper, M. (2016). *TIMSS 2015 international results in science*. Chestnut Hill, MA: Boston College, TIMSS & PIRLS International Study Center.

McCoach, D. B., & Adelson, J. L. (2010). Dealing with dependence (part I): Understanding the effects of clustered data. *Gifted Child Quarterly, 54*(2), 152–155.

McCormick, R. (2004). Collaboration: The challenge of ICT. *International Journal of Technology and Design Education, 14*(2), 159–176.

Meinck, S. (2015). *Sampling design and implementation*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Ministry of Education and Research. (1996). *IT i norsk utdanning. Plan for 1996–1999* [IT in Norwegian education. Plan for 1996–1999]. Ministry of Education and Research.

Ministry of Education and Research. (2000). *IT i norsk utdanning. Plan for 2000–2003. [IT in Norwegian education. Plan for 2000–2003]*. Ministry of Education and Research.

Ministry of Education and Research. (2004). *Program for digital kompetanse [Program for digital competence]*. Ministry of Education and Research.

Ministry of Education and Research. (2006). *Læreplanverket for kunnskapsløfte [curriculum for the knowledge promotion]*. Oslo, Norway: Ministry of Education and Research.

Mioduser, D., Nachmias, R., & Forkosh-Baruch, A. (2008). New literacies for the knowledge society. In J. Voogt & G. Knezek (Eds.), *International handbook of information technology in primary and secondary education, part one* (pp. 23–42). Boston, MA: Springer.

Muthén, B., & Muthén, L. (1998–2014). *Mplus Version 8.3* [statistical software package]. Muthén & Muthén.

O'bannon, B. W., & Thomas, K. (2014). Teacher perceptions of using mobile phones in the classroom: Age matters! *Computers & Education, 74*, 15–25.

OECD. (2015). *Students, computers and learning. Making the connection*. Paris, France: Organisation for Economic Co-operation and Development (OECD).

Pedró, F. (2007). The new millennium learners: Challenging our views on digital technologies and learning. *Nordic Journal of Digital Literacy, 2*(04), 244–264.

Pettersson, F. (2018). On the issues of digital competence in educational contexts–a review of literature. *Education and Information Technologies, 23*(3), 1005–1021.

Prestridge, S. (2017). Examining the shaping of teachers' pedagogical orientation for the use of technology. *Technology, Pedagogy and Education, 26*(4), 367–381.

Robinson, L., Cotten, S. R., Ono, H., Quan-Haase, A., Mesch, G., Chen, W., et al. (2015). Digital inequalities and why they matter. *Information, Communication & Society, 18*(5), 569–582.

Scheerder, A., van Deursen, A., & van Dijk, J. (2017). Determinants of internet skills, uses and outcomes. A systematic review of the second-and third-level digital divide. *Telematics and Informatics, 34*(8), 1607–1624.

Scherer, R., Siddiq, F., & Teo, T. (2015). Becoming more specific: Measuring and modeling teachers' perceived usefulness of ICT in the context of teaching and learning. *Computers & Education, 88*, 202–214.

Schulz, W., & Ainley, J. (2015). ICILS questionnaire development. In J. Fraillon, W. Schulz, T. Friedman, J. Ainley, & E. Gebhardt (Eds.), *ICILS 2013 technical report* (pp. 23–36). Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Schulz, W., & Friedman, T. (2015). Scaling procedures for ICILS questionnaire items. *ICILS Test Development, International Computer and Information Literacy Study: ICILS 2013: Technical Report*. https://www.iea.nl/publications/technical-reports/icils-2013-technical-report

The Norwegian Directorate for Education and Training. (2012). *Rammeverk for grunnleggende ferdigheter* [Framework for basic skills]. Norwegian Directorate for Education and Training, Oslo, Norway.

The Norwegian Directorate for Education and Training. (2020). *Resultat av søknadsbehandlingen – lærere [Result of application processing-teachers]*. https://www.udir.no/kvalitet-og-kompetanse/etter-og-videreutdanning/larere/resultat-av-soknadsbehandlingen%2D%2D-larere/

Tondeur, J., Aesaert, K., Pynoo, B., van Braak, J., Fraeyman, N., & Erstad, O. (2015). Developing a validated instrument to measure preservice teachers' ICT competencies: Meeting the demands of the 21st century. *British Journal of Educational Technology, 48*(2), 462–472. https://doi.org/10.1111/bjet.12380

Vanderlinde, R., Aesaert, K., & Van Braak, J. (2014). Institutionalised ICT use in primary education: A multilevel analysis. *Computers & Education, 72*, 1–10.

Volckmar, N. (2019). The enduring quest for equity in education: Comparing Norway and Australia. *Scandinavian Journal of Educational Research, 63*(4), 617–631.

Vrasidas, C. (2015). The rhetoric of reform and teachers' use of ICT. *British Journal of Educational Technology, 46*(2), 370–380.

Wang, S. K., Hsu, H. Y., Reeves, T. C., & Coster, D. C. (2014). Professional development to enhance teachers' practices in using information and communication technologies (ICTs) as cognitive tools: Lessons learned from a design-based research study. *Computers & Education, 79*, 101–115.

Ward, L., & Parr, J. M. (2010). Revisiting and reframing use: Implications for the integration of ICT. *Computers & Education, 54*(1), 113–122.

Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika, 54*(3), 427–450.

Warschauer, M. (2002, July 1). Reconceptualizing the digital divide. *First Monday, 7*(7).

Wastiau, P., Blamire, R., Kearney, C., Quittre, V., Van de Gaer, E., & Monseur, C. (2013). The use of ICT in education: A survey of schools in Europe. *European Journal of Education, 48*(1), 11–27.

Wikan, G., & Molster, T. (2011). Norwegian secondary school teachers and ICT. *European Journal of Teacher Education, 34*(2), 209–218.

Zhong, Z.-J. (2011). From access to usage: The divide of self-reported digital skills among adolescents. *Computers & Education, 56*, 736–746.

Chapter 7
Teachers' Role in Enhancing Equity—A Multilevel Structural Equation Modelling with Mediated Moderation

Trude Nilsen, Ronny Scherer ⓘ, Jan-Eric Gustafsson, Nani Teig, and Hege Kaarstein

Abstract Even though equity is an important aim for the Nordic countries, for many of these countries, the effect of a student's home background on their achievement seems to increase over time. If the aim is to reduce the effect of SES (socio-economic status) on student outcomes, there is a need to identify the factors that moderate this relation. One such factor could be teachers and their instruction because they have been found to be key to student outcomes. However, few have linked teachers and their instruction to equity, and fewer still have made this link in Nordic countries. The aim of the present study is to identify the aspects of teacher quality and their instruction that may reduce the relationship between SES and student achievement in the Nordic countries. Eighth-grade students from the only two Nordic countries participating in TIMSS 2015 (Norway and Sweden) were selected. Multigroup, multilevel (students and classes) structural equation models with random slopes were employed to investigate which aspects of teacher quality moderate the relation between SES and student science achievement via instructional quality. The findings show that teacher professional development and specialisation reduce the relation between SES and science achievement via instructional quality in Sweden, while there were no significant findings for Norway. This study contributes to the fields of equity and teacher effectiveness, demonstrating that teachers may make a difference in reducing inequity through their competence and instruction.

T. Nilsen (✉) · N. Teig · H. Kaarstein
Department of Teacher Education and School Research, University of Oslo, Oslo, Norway
e-mail: trude.nilsen@ils.uio.no

R. Scherer
Centre for Educational Measurement, University of Oslo, Oslo, Norway

J.-E. Gustafsson
Centre for Educational Measurement, University of Oslo, Oslo, Norway

University of Gothenburg, Gothenburg, Sweden

7.1 Background and Rationale

Educational systems around the world have long strived to increase educational equity, yet a large body of research has established a prevailing and substantial relation between socioeconomic status (SES) and student achievement (Kim, Cho, & Kim, 2019; OECD, 2016; Sirin, 2005), and this relation seems to have increased in the Nordic countries over time (Hansen, 2015; Nilsen, Bjørnsson, & Olsen, 2018; OECD, 2016). For Sweden, the level of equity is now below the OECD average, with a score point difference of 44 in science achievement associated with a one unit increase in the ESCS[1] (OECD, 2016). For Norway, the level of equity is not statistically different from the OECD average, with score point difference of 37 in science achievement associated with a one unit increase in the ESCS (OECD, 2016). This development is unfortunate, as it threatens the idea behind the Nordic model which is based on an ideal model of a "School for All" (see Chap. 2).

However, researchers have paid little attention to investigating the possible *mechanisms* through which SES is related to educational achievement (Berkowitz, Moore, Astor, & Benbenishty, 2017). Rather, SES is mostly utilized to control for selection bias when investigating effects of predictors on educational outcomes (Broer, Bai, & Fonseca, 2019). However, if educational systems aim to reduce the strength of the relationship between SES and student outcome, which is often used as an indicator of educational equity (see Chaps. 2 and 3), there is a need to identify factors that *moderate* this relation (Atlay, Tieben, Hillmert, & Fauth, 2019). In fact, knowledge about these factors could support educational systems with reducing educational gaps by manipulating factors, such as school climate, instructional quality, and teacher quality.

Although existing research has shown that teachers and their instruction are crucial for student outcomes, few studies have linked these aspects to equity (e.g., Darling-Hammond, 2015; Hwang, Choi, Bae, & Shin, 2018; Teig, Scherer, & Nilsen, 2018), and even fewer studies have been conducted in the Nordic countries. However, some studies from Germany and the United States found that high-quality teachers may enhance equity by reducing the gap between high- and low-SES students (Baumert et al., 2010; Darling-Hammond, 2015; Rjosk et al., 2014). While parents may support high-SES students may in their schoolwork (e.g., Tan, Lyu, & Peng, 2019), high-quality teachers may compensate for a lack of such support in low-SES students (Jeynes, 2005). Improving teacher quality, competence, and instruction may result in more students reaching their full potential (Atlay et al., 2019; Rivers & Sanders, 2002; Rjosk et al., 2014).

Researchers have linked formal teacher qualifications, including educational level, specialization, and professional development (PD), to high-quality teaching (Blömeke, Suhl, Kaiser, & Döhrmann, 2012). Despite mixed evidence on the effectiveness of teacher educational level, several studies have shown a substantial effect

---

[1] ESCS refers to the Program for International Student Assessment (PISA) index of Socio-Economic Status.

of teacher specialization on student outcomes and equity (e.g., Goe, 2007; Qin & Bowen, 2019). Research syntheses have also demonstrated a significant effect of teacher PD on student achievement (Goe, 2007; Kraft, Blazar, & Hogan, 2018; Timperley, Wilson, Barrar, & Fung, 2007). In the United States, some researchers have even suggested that increasing PD for teachers would contribute to closing the achievement gap between students (Darling-Hammond, 2015; Fischer et al., 2016).

To improve teacher quality, in 2013, Sweden implemented a massive teacher PD program in mathematics, which has since been extended to other subjects (Ringarp & Parding, 2018). Similarly, Norway has also made substantial investments in teacher PD (Regjeringen, 2014), albeit on a lesser scale than in Sweden. Given the decreasing levels of equity and the increasing emphasis on improving teacher quality in these Nordic countries (Hansen, 2015; OECD, 2016; Regjeringen, 2014; Ringarp & Parding, 2018), the question arises whether teacher quality may reduce the relation between SES and achievement.

However, teacher quality is rarely *directly* related to student outcomes; instead, it exerts an indirect effect via instructional quality (e.g. Baumert et al., 2010; Fauth et al., 2019). Hence, researchers that examine whether teacher quality may moderate the relation between SES and achievement should also consider indirect effects via instructional quality (i.e., a possible mediational path).

By taking into account these possible mechanisms of relationship between student SES and achievement, the overall aim of this study is twofold: (a) to identify the aspects of teacher qualifications and their instruction that may reduce the relation between SES and student achievement in Norway and Sweden (*moderation*) and (b) to examine whether the moderation effect of teacher qualifications is (partially) mediated via instructional quality (*mediated moderation*). More specifically, this study addressed these aims within the context of science education. Investigating educational equity in this context is of key significance as reforms in science education continue to promote scientific literacy as a fundamental goal of school science (Norris & Phillips, 2003). Despite considerable efforts in developing scientific literacy of all students, research has shown that those from underprivileged SES families are at a disadvantage when it comes to learning the language of science (Ryoo, 2009). Along this line, it is valuable to investigate the relationship between student SES and achievement, particularly by taking into account teacher qualifications and practices in science teaching.

7.2 Theoretical Framework

7.2.1 Educational Equity

One of the most important goals of most educational systems is to provide equitable opportunities and to enable all students to fully realize their academic potential, irrespective of their gender, ethnic belonging, or SES (Opheim, 2004). This has

been an especially important goal and the idea behind the Nordic model (see Chap. 2).

According to Espinoza (2007), equity and equality are interrelated and defined in a number of ways (see Chap. 2). One of Espinosa's definitions refers to "equality on average across social groups" and describes equal opportunities for all students to achieve high academic outcomes, no matter their social background (p. 353). To an extent, the present study belongs under this umbrella. However, the OECD and UNESCO have narrowed this broad concept (OECD, 2016; UNESCO, 2018). International large-scale assessments (ILSAs) and especially the OECD (2016) has had a significant influence on the conceptualization of equity. UNESCO has likewise had an impact on the conceptualization and measurement of equity and equality (2018). One of the most common indicators of educational equity is the strength of the relation between SES and student academic achievement. This indicator of equity is referred to as "impartiality" (UNESCO, 2018). While knowledge of this relation is important, it still does not answer whether and how schools may compensate for such inequity. "Redistribution" is a type of equity referring to compensating mechanisms for inequity (UNESCO, 2018). For example, low-SES schools may be allocated resources to compensate for students' disadvantage. In order for policy to enact compensatory approaches, it is vital to know what factors may reduce the impact of students' background (e.g., gender, ethnicity, etc.) on their academic outcome. This is exactly what the present study investigates by exploring how and if teacher quality and their instruction may reduce the impact of student SES on achievement.

7.2.2   Teacher Quality

Teacher quality is a broad concept and conceptualized somewhat differently across studies. Researchers have also used the concepts of teacher quality and teaching quality interchangeably. In this study, we separate the two concepts and refer to teacher quality as the skills, beliefs, and abilities the teachers bring into the classroom, whereas we define teaching quality, or instructional quality, as the teachers' behavior in the classroom and the quality of their instruction.

Goe (2007) suggested that the inputs of teacher quality include teacher *qualifications* (e.g., education, certification, experience) and teacher *characteristics* (e.g., self-efficacy, attitudes, beliefs). In a similar vein, Blömeke, Olsen, and Suhl (2016) proposed that teacher quality includes teacher qualifications (e.g., educational background, amount of experience in teaching, participation in PD) as well as personality characteristics, such as teachers' self-efficacy or beliefs. Focusing on ILSA studies, Klingebiel and Klieme (2016) applied a conceptual framework of teacher quality that consists of: (a) teacher qualifications including education and PD and (b) teacher competence involving teacher professional knowledge, beliefs, and non-cognitive or motivational factors. Despite using different labels to indicate some aspects of teacher quality, these studies have offered a similar conceptual

framework of teacher quality, which comprises both teacher qualifications and teacher competence/characteristics.

In the present study, we focus on teacher qualifications rather than their competence/characteristics for the following reasons. First, previous research has shown that teacher qualifications are related to educational equity (Darling-Hammond, 2015). For example, high-SES schools may have more qualified teachers than low-SES schools have (e.g. Darling-Hammond, 2006; Lankford, Loeb, & Wyckoff, 2002). In Norway, a study revealed a lack of certified teachers in schools with high proportions of minority students and students with special needs (Bonesrønning, Falch, & Strøm, 2005). Researchers identified a similar pattern in Sweden (Hansson & Gustafsson, 2017). Additionally, teacher qualifications (e.g., certification) may have larger effects on low-SES students than on high-SES students (e.g. Nye, Konstantopoulos, & Hedges, 2004), although studies including such moderation effects of teacher qualifications are rare.

Second, teacher qualifications—such as their specialization, educational level, and PD—are important factors that can be influenced through educational policy (e.g., through teacher education). Even though educational policy may influence teacher characteristics, such as increased self-efficacy through teacher education, this mechanism is difficult to establish or measure. Third, the present study emphasizes a Nordic perspective and comparison across the Nordic countries. Teacher competence measured by, for instance, a test within a certain domain has proven difficult to measure across countries (Blömeke & Delaney, 2014; Blömeke, Hsieh, Kaiser, & Schmidt, 2014). Due to the above-mentioned reasons, this study concentrates on the qualification aspect of teacher quality, more specifically on teacher education, specialization, and PD. The following sections discuss each of these aspect in detail.

Teacher Education Researchers conducting ILSA studies commonly measure teacher formal level of education using International Standard Classification of Education (ISCED) levels (e.g. Mullis, Martin, Foy, & Hooper, 2016). Of the Nordic countries, teachers in Finland have the highest education, where more than 90% of them have a master's degree or higher (Mullis, Martin, Foy, et al., 2016; OECD, 2016). While the effect of teachers' educational level has often been hard to establish and varies greatly from one country to another (Blömeke et al., 2016), some studies have demonstrated a significant effect of teacher's level of education on student achievement (Blömeke et al., 2016; Nilsen, Scherer, & Blömeke, 2018) and in enhancing equity (Heck, 2007).

Teacher Specialization Specialization in the content domain is an important part of teachers' qualifications as well as an indicator of their content knowledge and pedagogical content knowledge (Blömeke et al., 2014; Goe, 2007). Student learning depends to a large degree on teachers who have specialized in the subject they teach and whose content knowledge and pedagogical content knowledge are sound (e.g., Baumert et al., 2010; Blömeke et al., 2016; Goe, 2007; Nilsen, Scherer, et al., 2018).

Such teachers may also reduce the achievement gap between students (Baumert et al., 2010).

Teacher PD Research syntheses found that teacher PD may have significant effects on student achievement (Goe, 2007; Kraft et al., 2018; Timperley et al., 2007). However, for PD to have an effect on student learning, it needs to be of sufficient length and quality (Timperley et al., 2007). As such, sufficient teacher PD may be an important factor in reducing the achievement gap among different groups of students (e.g. Darling-Hammond, 2015).

7.2.3   Teacher Qualifications in Norway and Sweden

A natural point of departure for reviewing the teachers' formal qualifications in Norway and Sweden can be traced back to an important phenomenon known as the PISA shock in 2001 (Elstad, Nortvedt, & Turmo, 2009; Haugsbakk, 2013; Lundström, 2015; Tveit, 2013). Norwegian students produced results on the PISA 2001 that were so far below expectations that the Norwegian Minister of Education compared it with the failure to bring home any medals from the Winter Olympics (Elstad et al., 2009; Nortvedt, 2018; Tveit, 2013). Following the PISA shock in 2001, Norway implemented several policy changes, including reforming the National Curriculum for Grades 1–13 called the "Knowledge Promotion" and introducing a National Quality Assessment System that implemented national tests alongside participation in other ILSA studies, like TIMSS[2] (for more details, see Elstad et al., 2009). A similar line of events also took place in Sweden as the PISA shock had a profound impact on educational policy (Ringarp, 2016). The PISA shock may also have been an important factor that drove the implementation of national tests in both Norway and Sweden (Lundahl & Waldow, 2009; Lundström, 2015). In addition to these initiatives and actions that focused on improving student outcomes, the Norwegian and Swedish governments also reformed teacher education and made changes to employment regulations for teachers.

The teacher education practices and programs in Norway and Sweden are quite similar and are founded on the same principles, values, and traditions (Ringarp & Parding, 2018). Some large reforms in teacher education have had an impact on the current teacher education and qualifications in these countries. Norway implemented a large teacher education reform in 2010 that divided teacher education for Grades 1–10 into two types of programs: classroom teachers for Grades 1–7 and specialized teachers for Grades 5–10 that focus on one or two subjects (for more details, see e.g., Munthe, Malmo, & Rogne, 2011). Individuals interested in teaching Grades 8–13 have always had an alternative route to become a teacher in Norway by following a university program and specializing in one or two subjects. Since 2014, Norwegian teachers must have a minimum of 30 credit points (i.e., one full-time semester) in

---

[2] Trends in Mathematics and Science Study, see https://timssandpirls.bc.edu/

science in addition to the already required pedagogical education to be hired and teach the subject. This requirement was not included in the teacher hiring requirements before 2014. All science teachers in Norway now have until 2025 to fulfil the last extension of the requirements (Ministry of Education and Research, 2015).

Sweden implemented two large teacher education reforms in 1998 and 2011. Research became an integral part of education in the 1998 reform, as all teachers were required to attend the same educational program and educating students across different socioeconomic classes were specifically targeted to enhance equity (Ringarp & Parding, 2018). In the 2011 reform, the educational program was differentiated and split into four educational programs: preschool teacher, classroom teacher for primary school Grades 1–3 and 4–6, specialized teachers (e.g., science teachers for Grades 7–9 or upper secondary school), and teachers for vocational tracks (Ringarp & Parding, 2018). Since 2011, all Swedish teachers are required to obtain a teaching certificate, and science teachers who teach Grades 7–9 need at least 45 credit points[3] to teach the subject.

The investments in teachers' PD gained more widespread recognition since 2009 in Norway (Lagerstrøm, Moafi, & Revold, 2014) and since 2001 in Sweden. In Norway, school administrators (i.e., municipalities or counties) are responsible for meeting their teachers' need for PD. The main focus has been to ensure that teachers have the minimum required study credits (Ministry of Education and Research, 2008). In Sweden, teachers' PD is more centralized and described as one of the national steering devices for the government (Kirsten & Wermke, 2017). Substantial amounts of resources are invested for improving teacher quality, including granting teachers with 13 days per year to attend PD (Kirsten & Wermke, 2017). In 2013, Sweden implemented an extensive teacher PD program in mathematics, which has also been extended to other subjects (Boesen, Helenius, & Johansson, 2015; Ringarp & Parding, 2018).

In Norway, the politically prioritized subject has been mathematics (OECD, 2019), which could be why only a very small number of Norwegian students in TIMSS 2015 had teachers who participated in PD in science in the last 2 years (Martin, Mullis, Foy, & Hooper, 2016). Only 4% to 12% of Norwegian students were taught by teachers who participated in PD in other different topics. Conversely, this number ranged between 23% and 35% in Sweden (Martin et al., 2016).

In spite of all the reforms to improve teacher quality in Sweden and Norway, research examining teacher education and PD in these countries has been limited. Few studies have investigated whether these reforms have had an impact on student outcomes and educational equity or whether teacher qualifications relate to students' science learning outcomes, especially by comparing the results in both countries. However, some studies have found that PD implemented by the government was associated with high performance in Sweden (Gustafsson & Nilsen, 2017; Nilsen, Scherer, et al., 2018).

---

[3] One semester of full-time study is 30 credit points.

7.2.4 Instructional Quality

Similar to teacher quality, instructional quality is a broad concept operationalized differently across countries and studies (e.g. Blömeke et al., 2016; Ferguson & Danielson, 2014; Kuger, Klieme, Jude, & Kaplan, 2016; Kyriakides, Creemers, & Antoniou, 2009; Nilsen, Scherer, et al., 2018; Pianta & Hamre, 2009). Despite these differences, researchers in Europe (Blömeke et al., 2016; Kuger et al., 2016) have extensively used the framework of instructional quality from Klieme, Pauli, and Reusser (2009). According to this framework, instructional quality includes three main aspects: classroom management, cognitive activation, and teacher support.

Classroom management is often considered to be independent of the subject domain (Klieme et al., 2009). All subjects would require effective classroom management, including clear rules and procedures about the time spent on tasks and disciplinary situations. Since this study focuses on the context of science education, investigating a generic aspect like classroom management has become of less interest. In addition, classroom management has been frequently studied in research on instructional quality; hence, its relation to student outcome has been well established (Kyriakides et al., 2009; van Tartwijk & Hammerness, 2011). Thus, this particular aspect of instructional quality is not included in the present study.

In contrast with classroom management, *cognitive activation* is the aspect of instructional quality that is most dependent on the subject domain (Klieme et al., 2009; Kuger et al., 2016). In the domain of science, cognitive activation includes engaging students with cognitively challenging lessons through inquiry activities, such as interpreting data from scientific experiments (Minner, Levy, & Century, 2010; Teig, Scherer, & Nilsen, 2019). In general, cognitive activation comprises instructional activities that challenge students cognitively and engage them with high-level thinking, for example, through evaluating, integrating, and applying knowledge in the context of problem solving (Baumert et al., 2010; Hiebert & Grouws, 2007; Nilsen & Gustafsson, 2016).

Teacher support refers to practices related to the teacher's response to students' needs, including listening to and respecting students' ideas and questions and encouraging classroom discussions among students. A supportive teacher would show an interest in every student's learning, provide feedback, and adapt practices to the individual student's needs (Blömeke et al., 2016).

In addition to these three aspects, some studies have included a fourth aspect of instructional quality, known as *clarity of instruction* (Bellens, Van Damme, Van Den Noortgate, Wendt, & Nilsen, 2019; Bergem, Nilsen, & Scherer, 2016). Clarity of instruction relates to a clear and comprehensive teaching practice. To achieve clarity of instruction, the teacher must set clear learning goals, provide a summary at the end of the lesson, and link new and old topics (Bergem et al., 2016; Cohen & Grossman, 2016; Raudenbush, 2008). Although clarity of instruction could be integrated into the aspect of teacher support, the present study separates these two aspects of instructional quality.

Some studies have investigated the relation between teachers' instructional quality and educational equity; however, most of these studies were situated in Germany and the United States. Rjosk et al. (2014) investigated language instruction in German classrooms and found that cognitive activation mediated the relation between SES and achievement. Willms (2010) analyzed data from PISA 2006 and found that instructional quality mediated the relation between SES and achievement at the school level. Using data from TIMSS 2011, researchers have determined that instructional quality moderates the relation between SES and achievement (Gustafsson, Nilsen, & Hansen, 2018). Although the findings varied across the 50 countries who participated in TIMSS 2011, this study shows that instructional quality reduced the strength of the effect of SES on achievement in some countries (Gustafsson et al., 2018).

The body of extant literature on the moderating role of instructional quality is diverse, as the following two examples show. In a recently published study of a large German student sample, Atlay et al. (2019) examined the moderating role of instructional quality on the relation between SES and achievement in mathematics. Atlay et al. (2019) found that cognitively activating classrooms and good teacher support were beneficial especially for high-SES students; surprisingly, this study found support for a positive rather than negative moderation effect. In contrast, a study of the TIMSS 2015 national extensions in three countries (i.e., Germany, Belgium, and Norway) could not identify any moderation effect for any of the three core dimensions of instructional quality (Bellens et al., 2019). Considering these findings, the role of instructional quality as a possible moderator of the SES–achievement relation remains unsettled and warrants further empirical investigation.

7.3 Methodology

7.3.1 Data and Sample

We utilized large-scale data from TIMSS, the only study with representative samples at the national level that collects data from students and teachers in mathematics and science. Furthermore, TIMSS is the only ILSA that samples entire classes within schools, enabling investigations of factors explaining variance between classes. As the factor of teacher qualifications seems to be of more importance for student outcomes in lower secondary than in primary school (Goe, 2007; Nilsen, Scherer, et al., 2018), we selected Grade 8 students from the only two Nordic countries participating in the last cycle of TIMSS in 2015: Sweden and Norway.

7.3.2   Measures

Teacher Quality Teacher qualifications were used to measure teacher quality through the following indicators: (a) *educational level* from ISCED level 3 to 8; (b) *specialization* as determined by the major or main area of study in science education and in physics, biology, chemistry, or earth science; (c) *content of PD* or teachers' participation in various PD activities in the last 2 years, including science content, science pedagogy/instruction, curriculum, integration of information technology into science teaching, improving students' critical thinking or inquiry skills, and science assessment; and (d) *hours of PD* as determined by the number of hours teachers spent in PD in the last 2 years.

Instructional Quality We measured this construct using teachers' ratings of how often they would do certain practices (measured on a four-point scale from *never* to *every or almost every lesson*). In accordance with the framework of instructional quality (e.g. Klieme et al., 2009), we included five items pertaining to cognitive activation (e.g., "Ask students to complete challenging exercises that require them to go beyond the instruction"), teacher support (e.g., "Encourage classroom discussions among students"), and clarity of instruction (e.g., "Link new content to students' prior knowledge"). Note that TIMSS 2015 did not measure classroom management.

In addition, the measurement models of teacher qualifications and instructional quality demonstrated metric invariance across the Nordic countries (Nilsen & Gustafsson, 2016; Nilsen, Scherer, et al., 2018), which implies teachers from these countries interpreted both constructs similarly.

SES TIMSS 2015 measured students' SES by their responses on questions about parents' education, the number of books at home, and the educational resources at home. A composite score for SES[4] was estimated based on an item response theory model to represent students' individual socioeconomic background.

Science Achievement The TIMSS 2015 science assessment contained 250 items that covered topics in chemistry, physics, biology, and earth science. These items captured the breadth of the science domain as well as the range of cognitive dimensions (i.e., knowing, applying, and reasoning). Five plausible values were drawn from the achievement distribution to represent science achievement. The mean science achievement for both countries was slightly different; specifically, Swedish students had a mean of 522 with a standard deviation of 3.4, whereas Norwegian students had a mean of 509 with a standard deviation of 2.8.

---

[4] http://timssandpirls.bc.edu/timss2015/international-results/timss-2015/mathematics/home-environment-support/home-resources-for-learning/

Table 7.1 shows percentages of teacher characteristics and qualifications in Sweden and Norway. More detailed information on the questionnaires and descriptive statistics of the measures are available on the TIMSS 2015 website.[5]

Table 7.1 Percentages of teacher characteristics and qualifications in Norway and Sweden

Variables	Sweden ($n = 706$)	Norway ($n = 225$)
Gender		
Male	41.0	45.1
Female	59.0	54.9
Years of teaching experience		
<10 years	38.6	51.5
10–19 years	43.9	34.0
20–30 years	10.7	9.7
>30 years	6.8	4.9
Level of formal education		
Upper secondary	7.1	–
Short-cycle tertiary	2.4	2.4
Bachelor or equivalent	52.0	69.9
Master or equivalent	34.4	27.2
Doctor or equivalent	4.1	0.5
Major area of education		
Science and science education	47.7	17.2
Only in science	26.5	33.5
Only in science education	13.7	9.4
All other majors	5.0	39.9
No formal education beyond upper secondary school	7.1	–
The number of hours teachers attended PD in the past 2 years		
None	33.5	57.7
<6 h	27.3	18.9
8–15 h	22.2	10.7
16–35 h	7.1	4.6
>35 h	9.9	8.2
The content of PD teachers attended in the past 2 years		
Science content	32.0	19.4
Science pedagogy/instruction	27.2	15.8
Science curriculum	34.0	11.2
Integrating information technology into science	26.6	8.2
Improving student' critical thinking or inquiry skills	22.3	9.7
Science assessment	32.0	12.8
Addressing individual students' needs	30.3	8.2

[5] https://timssandpirls.bc.edu/timss2015/

7.3.3   Data Analysis

Two-group (i.e., Sweden, Norway) and multilevel (i.e., students nested in classes) structural equation modeling (SEM) with random slopes was employed. A random slope model allows each group (i.e., class) to have a different slope, which means that the explanatory variable (i.e., teacher qualifications and instructional quality) may have a different effect for each group.

SEM is a multivariate statistical analysis technique that includes confirmatory factor analysis (CFA). CFA generates factor loadings of indicators on an underlying latent factor. Along with the model fit indices, the factor loadings provide a measure for reliability and validity (Hox, Moerbeek, & Van de Schoot, 2017). SEM allows researchers to examine the relationships between multiple observed and unobserved variables, while providing explicit estimates of error variance parameters. It further enables complex modeling (e.g., multi-group and random slopes models) and complex patterns with intervening variables between the independent and dependent variables, and independent variables may also function as dependent variables (Preacher, Zyphur, & Zhang, 2010).

Furthermore, another great advantage of SEM is the possibility for multi-level approaches (MSEM) where it is possible to model at all levels simultaneously. MSEM with measurement models with multiple indicators is the most robust method for multi-level analyses with latent variables (Hox et al., 2017).

We specified cross-level interaction models with indirect effects to test which aspects of teacher qualifications moderate the relation between individual students' SES and science achievement via classroom instructional quality. All models were estimated using the software M*plus* 8.3 with the robust maximum likelihood estimation (Muthén & Muthén, 1998–2017). Prior to adding any structural models, multilevel confirmatory factor analyses were conducted to ensure reliable and valid measurement models of each construct at both the student and the classroom level. Indirect effects that may indicate (partial) mediation were estimated using the MODEL CONSTRAINTS option in M*plu*s, with Wald 95% confidence intervals. It should be noted that the coefficients provided in the results section were not standardized. All models follow the latent decomposition approach for variables that were measured at the student level but aggregated to the classroom level, following an approach presented by Marsh et al. (2009).

Figure 7.1 shows the conceptual model for the overall aim in the present study. The black dot reflects the random slope of the relation between SES and achievement at the student level. The arrows pointing to the dot represent the relation between the classroom level predictors, teacher qualifications and instructional quality, as well as the variation in the slope. In other words, the model shows how teacher qualifications and instructional quality may moderate the relation between SES and achievement at the student level. Furthermore, the model shows a mediation path where instructional quality mediates teacher qualifications' moderation of the relation between SES and achievement.
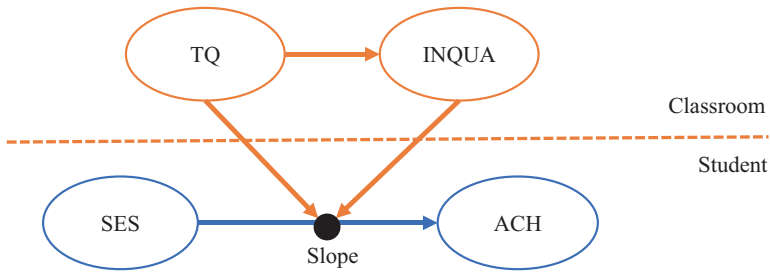
Fig. 7.1 The conceptual model of the overall aims in this study. TQ = teacher qualifications; INQUA = instructional quality; ACH = student achievement, SES = students' socioeconomic status
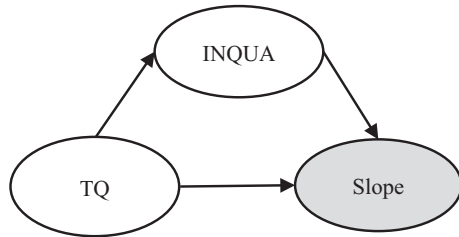
Fig. 7.2 The analytical model at the classroom level. TQ = teacher qualifications; INQUA = instructional quality; Slope = random slope of the relation between SES and achievement at the student level

This model creates a latent variable for the slope between SES and achievement at the classroom level. Hence, in addition to the relations shown at the classroom level, we investigate the relation between the predictors of teacher qualifications and instructional quality and the slope, as shown in Fig. 7.2. We further control for the relation between SES and achievement at the classroom level. Figure 7.2 reflects the analytical model at the classroom level created in M*plus*.

A direct moderation effect that enhances equity (by reducing the strength between SES and achievement) would require a negative, significant relation between teacher qualifications and the slope. A mediated moderation that enhances equity would require a negative, significant mediation effect from teacher qualifications to the slope via instructional quality.

7.4 Results

The purposes of the present study were to (a) identify various aspects of teacher qualifications (i.e., content of PD, hours of PD, teacher educational level, and teacher specialization) that may have contribute to reducing the relationship between student SES and achievement and (b) examine whether the moderation effect of teacher qualifications was (partially) mediated via instructional quality.

a) Content of PD: Sweden
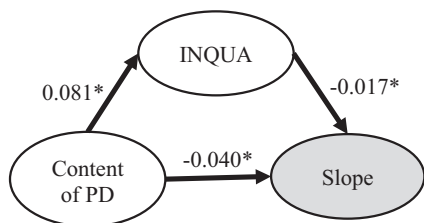
b) Content of PD: Norway
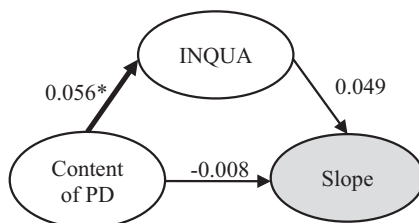
Fig. 7.3 Moderation model at the classroom level in (**a**) Sweden and (**b**) Norway. PD = profes-
sional development; INQUA = instructional quality; Slope = random slope of the relation between
SES and achievement at the student level. *$p < .05$

Figure 7.3 presents the main results at the classroom level for teachers' participa-
tion in various PD activities (content of PD). For the Swedish data (Fig. 7.3a), all
three relations between content of PD, instructional quality, and the slope were sig-
nificant, along with a significant mediation effect. The moderation coefficient was
negative and significant ($B = -0.040$), suggesting that content of PD reduced the
strength of the relation between SES and student achievement via instructional
quality. This may suggest good teaching quality—indicated by teachers who have
participated in various activities of PD—reduces the importance of student home
background for science achievement and hence, enhances equity among students.
For the Norwegian data, on the other hand, only the relation between content of PD
and instructional quality was significant, and neither mediation nor moderation
effects were evident (Fig. 7.3b).

We further controlled for the relation between classroom SES and achievement.
The results showed that the relation between SES and student achievement at the
classroom level was $B = 2.04$ ($SE = .158$, $p < .01$) for the Swedish data and $B = 1.22$
($SE = .141$, $p < .01$) for the Norwegian data. Hence, an increase of one unit in the
classroom-level SES scale was associated with a 204-point score increase in
Sweden. This change represents about twice the standard deviation of classroom-
level achievement. In Norway, a one-unit increase of the classroom-level SES scale
was associated with a 122-point score increase in classroom-level achievement.

With respect to the number of hours teachers spent on PD (hours of PD), we
identified a direct, significant, and negative moderation effect in Sweden (Fig. 7.4a).
The corresponding regression coefficient was smaller than for the model with the
content of PD, and there was no significant mediation effect. These results indicate
that the number of hours teachers spent on PD enhanced equity among students in
Sweden. For Norway, no evidence for moderation and mediation surfaced (Fig. 7.4b).

For the teachers' educational level, we found no significant moderation effects in
either country (Fig. 7.5).

With regard to the teacher specialization, we found a significant, direct, and neg-
ative moderation effect in the Swedish data (Fig. 7.6a), indicating that this aspect of
teacher qualification enhances equity. Once again, no significant moderation and
mediation effect surfaced for the Norwegian data (Fig. 7.6b).
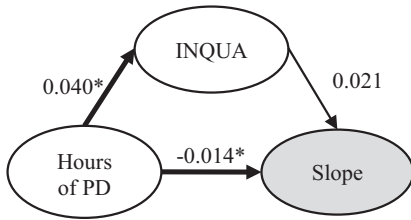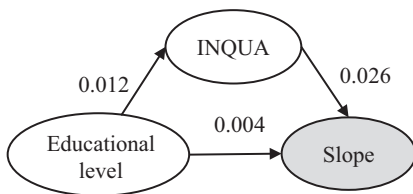
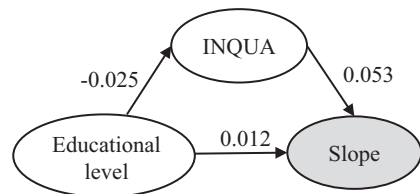a) **Hours of PD: Sweden**   b) **Hours of PD: Norway**
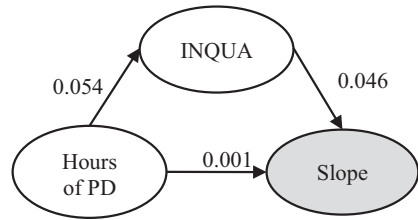
Fig. 7.4 Moderation model at the classroom level in (**a**) Sweden and (**b**) Norway. PD hours = hours of professional development; INQUA = instructional quality; Slope = random slope of the relation between SES and achievement at the student level. *p < .05

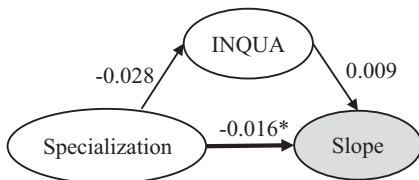a) **Teacher educational level: Sweden**   b) **Teacher educational level: Norway**

Fig. 7.5 Moderation model at the classroom level in (**a**) Sweden and (**b**) Norway. Educational level = educational level from ISCED level 3 to 8; INQUA = instructional quality; Slope = random slope of the relation between SES and achievement at the student level. *p < .05

a) **Teacher specialization: Sweden**   b) **Teacher specialization: Norway**
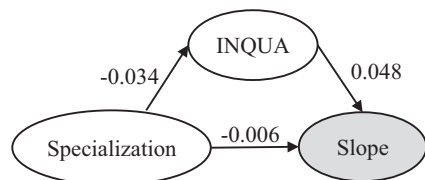
Fig. 7.6 Moderation model at the classroom level in (**a**) Sweden and (**b**) Norway. Specialization = teacher major or main area of study; INQUA = instructional quality; Slope = random slope of the relation between SES and achievement at the student level. *p < .05

7.5 Discussion

In this study, we investigated whether different aspects of teacher qualifications (i.e., content of PD, hours of PD, education level, and specialization) could reduce the strength of the relation between SES and achievement via their instructional quality. The results indicate that, in Sweden, teachers who participated in different PD activities helped enhance equity among students via their instruction. Moreover,

the length of these activities (i.e., hours of PD) also contributed to enhancing equity, although no mediation effect was detected. In Norway, we found no significant moderation or mediation effects for either teachers' participation in PD activities or the time they spent in these activities. With respect to teachers' educational level, we identified no significant effect for either the Norwegian or the Swedish data. For teacher specialization, conversely, there was a direct, significant, and negative moderation effect for Sweden. This finding indicates that teachers' area of specialization reduced the relation between SES and achievement and, thus, enhanced equity. In Norway, the moderation effect was insignificant for teacher specialization.

Both the content and number of hours teachers participated in PD contributed to enhance equity in Sweden. Given that Sweden has invested tremendous effort and resources into PD, these findings seem particularly promising. The findings were also in line with those from the United States (e.g. Darling-Hammond, 2015; Darling-Hammond, Hyler, & Gardner, 2017; Wilson, 2013), indicating that such efforts may reduce the performance gap between high- and low-SES students.

However, this study found no evidence that the number of hours teachers spent in PD contributed to enhancing equity in Norway, which could be due to several reasons. The number of teachers who participated in PD in Sweden was substantively larger than in Norway (Mullis, Martin, Foy, et al., 2016; Skolverket, 2016). In addition, relatively few Norwegian teachers participated in the TIMSS 2015 study; specifically, the study involved 225 teachers in Norway and 706 teachers in Sweden. The small sample in the Norwegian data could reduce the power of the statistical analyses, which might make it harder to detect findings that could in fact be significant. This explanation might be particularly true in the case of Norway, where fewer science teachers participated in the TIMSS study in comparison with Sweden. Among these participants, 57.7% of the Norwegian teachers stated they had never attended PD in the past 2 years in contrast to only 33.5% of the Swedish teachers (Table 7.1). Another possible explanation for the discrepancy might relate to how science teaching is delivered in both countries. In Norway, science is taught as an integrated subject whereas science is divided according to the subject domain (e.g., physics, chemistry, biology) in Sweden. Each subject domain would have a different teacher in Sweden. In other words, while only one science teacher is responsible for teaching a Grade 8 classroom in Norway, several subject-domain teachers are needed to accomplish similar tasks in Sweden. Taken together, the non-significant effects of PD in the Norwegian data might be attributable to the teachers' low participation in the PD activities and the small sample of teachers who participated in the TIMSS 2015 study.

Another plausible explanation could be due to differences in the quality of PD implemented in the two countries. For PD to have an impact on student learning, a certain level and type of quality are required (Boyle, Lamprianou, & Boyle, 2005; Timperley et al., 2007). Effective programs should be implemented for a considerable length of time and provide teachers with specific content focused on the curriculum and include active learning, collaborative activities, modeling of effective instruction, collegial collaboration, reflection, and continuous feedback (e.g. Darling-Hammond et al., 2017; Timperley et al., 2007). Sweden spent considerable resources and time

on PD for their teachers, and the courses were heavily based on research and structured planning (Gustafsson & Nilsen, 2017; Mullis, Martin, Goh, & Cotter, 2016; Ringarp & Parding, 2018). Conversely, the focus in Norway has been on mathematics teachers more so than science teachers, and even if the government in Norway has spent resources on teacher PD, it seems that few science teachers attended any programs (Martin et al., 2016).

While some studies have found a significant relation between PD and student outcome in Sweden (e.g. Gustafsson & Nilsen, 2017; Lindvall, 2017), no study has found a significant moderation of the relation between SES and achievement in Sweden. Although previous studies have investigated the moderation effects of PD on the relation between SES and achievement in Sweden and Norway, they did not include the *indirect* moderation effect via instructional quality (Nilsen & Bergem, 2020). This could be why this study found no significant moderation effect for Sweden or Norway. Including a mediated moderation model could then boost the power of the analyses as such a model to a large extent reflects the actual picture; in particular, teachers' qualifications in themselves are not valuable unless reflected in their teaching practices.

With regard to teacher specialization, the results showed that it contributed to reducing the importance of student home background in Sweden. Again, the findings for the Norwegian data were statistically insignificant. The aforementioned reasons for the lack of significant findings for Norway for content and hours of PD could also explain the insignificant effects of teacher specialization for Norway. Moreover, the Swedish data showed that 47.7% of the students had teachers who specialized in both science and science education, while only 17.2% of the students in Norway had access to such teachers (Table 7.1). Compared to the Norwegian teachers, Swedish teachers are required to take 50% more study points in science specialization to be formally qualified and allowed to teach this subject (Ministry of Education and Research, 2015; Skolverket, 2019). In addition, it is important to note that teacher specialization is only an indicator of teacher competence and not a direct assessment of teachers' knowledge and skills in science and science education. Such assessments are substantially more time-consuming for teachers and challenging to implement in ILSA studies. They require teachers to not only solve science tasks but also to answer the background questionnaire. Following this line of reasoning, the indirect assessment of teacher competence inherent in TIMSS could be one reason why the moderation and mediated moderation effects were not significant for teacher educational level in Norway or Sweden.

In summary, some possible explanations for why the content of PD, hours of PD, and teacher specialization reduced the strength of the relation between SES and achievement in Sweden and not in Norway may relate to the low statistical power (i.e., fewer teachers in the sample and fewer teachers who participated in PD activities in Norway) and the larger variations of students' SES in the Swedish data. Nevertheless, considering that previous studies have found that teachers' PD influenced student outcomes in Sweden but not in Norway (Gustafsson & Nilsen, 2017; Nilsen, Scherer, et al., 2018), it seems that the quality and length of the PD offered to the teachers in Sweden exceeded that of Norway. Perhaps improving the

quality and length of the training programs provided to science teachers in Norway could contribute to reducing the achievement gap between high- and low-SES students in Norway. Likewise, this suggestion could be applied to teacher specialization in Norway, as the average is substantially higher in Sweden than in Norway (Kaarstein, Nilsen, & Blömeke, 2016; Martin et al., 2016).

7.6   Limitations of the Study

As with all studies using ILSA data, no causal inferences can be drawn due to the cross-sectional design inherent in the studies. However, TIMSS has been repeated every 4 years since 1995, and the quality of this study has been enhanced for each cycle. TIMSS also implements a number of quality assurance procedures and pilots the survey in every cycle. In addition, this study's methodological approach of including multi-group and multilevel SEM is known to be the most robust and reliable analytical method for these types of research questions and offers higher levels of reliability and inferences.

Another limitation of this study relates to the low numbers of teachers who participated in TIMSS 2015, which may decrease the power in detecting significant findings. It may be argued that several of our findings would have been significant if more teachers had participated. This study could also suffer from construct underrepresentation when it comes to instructional quality. Although instructional quality is a multidimensional construct, TIMSS 2015 did not measure all aspects of instructional quality (e.g., classroom management). TIMSS 2019 has increased the emphasis on teacher practices by including all aspects of instructional quality. This change should consequently lead to higher validity and increase the power of the analyses in future studies.

7.7   Contributions and Implications

This study contributes to the knowledge base in the field of teacher quality and instructional quality. While it has been known for quite some time that instructional quality may mediate the relation between teacher quality and student outcome (Baumert et al., 2010; Blömeke et al., 2016), bringing together a mediation and moderation model represents a novel approach in this field. Although we found evidence only for the mediated moderation for teacher PD in Sweden, our findings indicate that researchers may want to examine such effects with teacher quality as the moderator. Teacher quality in itself (e.g., their specialization) is of little use unless it informs their classroom practices. For example, it is less likely that students achieve high learning outcomes from a teacher with a high educational level but with low instructional quality.

This study also contributes to the field of educational equity. While the number of studies examining equity is substantial, especially since the emergence of international large-scale studies, few have investigated the teacher's role in reducing inequity. Most of these studies have originated in the United States (e.g. Darling-Hammond, 2015), and very few have focused on the Nordic countries. It is especially interesting that professional development and teachers' specialization seemed to enhance equity in Sweden, given that Sweden has deviated from the Nordic model due to free school choice (Gustafsson & Yang Hansen, 2017, also see Chap. 2). In Chap. 3 there were strong indications that Sweden was an outlier compared to the other Nordic countries; regardless of how equity was measured and what methods were used, Sweden had a much lower level of equity. While Sweden's comparatively lower levels of equity is old news, our findings are uplifting, as Sweden's efforts to increase teacher competence may be a way back to the ideals behind the Nordic model.

One general implication of this study is that enhancing teachers' qualifications may increase the quality of their instruction and, ultimately, reduce the achievement gap between students. Providing PD for teachers and ensuring that teachers have sound qualifications may indeed reduce the effect of student home background on their achievement.

References

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*, 101211. https://doi.org/10.1016/j.learninstruc.2019.05.008

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., et al. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal, 47*(1), 133–180.

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).

Bergem, O. K., Nilsen, T., & Scherer, R. (2016). Undervisningskvalitet i Matematikk [Instructional quality in mathematics]. In O. K. Bergem, H. Kaarstein, & T. Nilsen (Eds.), *Vi kan lykkes i realfag* [We can succeed in mathematics and science]. Oslo, Norway: Universitetsforlaget

Berkowitz, R., Moore, H., Astor, R. A., & Benbenishty, R. (2017). A research synthesis of the associations between socioeconomic background, inequality, school climate, and academic achievement. *Review of Educational Research, 87*(2), 425–469. https://doi.org/10.3102/0034654316669821

Blömeke, S., & Delaney, S. (2014). Assessment of teacher knowledge across countries: A review of the state of research. In S. Blömeke, F.-J. Hsieh, G. Kaiser, & W. H. Schmidt (Eds.), *International perspectives on teacher knowledge, beliefs and opportunities to learn* (pp. 541–585). Dordrecht, The Netherlands: Springer.

Blömeke, S., Hsieh, F.-J., Kaiser, G., & Schmidt, W. (2014). *International perspectives on teacher knowledge, beliefs and opportunities to learn*. Retrieved from Dordrecht, The Netherlands.

Blömeke, S., Olsen, R. V., & Suhl, U. (2016). Relation of student achievement to the quality of their teachers and instructional quality. In *Teacher quality, instructional quality and student outcomes* (pp. 21–50). Cham, Switzerland: Springer.

Blömeke, S., Suhl, U., Kaiser, G., & Döhrmann, M. (2012). Family background, entry selectivity and opportunities to learn: What matters in primary teacher education? An international comparison of fifteen countries. *Teaching and Teacher Education, 28*(1), 44–55.

Boesen, J., Helenius, O., & Johansson, B. (2015). National-scale professional development in Sweden: Theory, policy, practice. *ZDM Mathematics Education, 47*(1), 129–141.

Bonesrønning, H., Falch, T., & Strøm, B. (2005). Teacher sorting, teacher quality, and student composition. *European Economic Review, 49*(2), 457–483.

Boyle, B., Lamprianou, I., & Boyle, T. (2005). A longitudinal study of teacher change: What makes professional development effective? Report of the second year of the study. *School Effectiveness and School Improvement, 16*(1), 1–27.

Broer, M., Bai, Y., & Fonseca, F. (2019). A review of the literature on socioeconomic status and educational achievement. In M. Broer, Y. Bai, & F. Fonseca (Eds.), *Socioeconomic inequality and educational outcomes: Evidence from twenty years of TIMSS* (pp. 7–17). Cham, Switzerland: Springer.

Cohen, J., & Grossman, P. (2016). Respecting complexity in measures of teaching: Keeping students and schools in focus. *Teaching and Teacher Education, 55*, 308–317.

Darling-Hammond, L. (2006). Securing the right to learn: Policy and practice for powerful teaching and learning. *Educational Researcher, 35*(7), 13–24.

Darling-Hammond, L. (2015). *The flat world and education: How America's commitment to equity will determine our future*. New York: Teachers College Press.

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

Elstad, E., Nortvedt, G., & Turmo, A. (2009). The Norwegian assessment system: An accountability perspective. *Cadmo, 15*, 89–103.

Espinoza, O. (2007). Solving the equity–equality conceptual dilemma: A new model for analysis of the educational process. *Educational Research, 49*(4), 343–363.

Fauth, B., Decristan, J., Decker, A.-T., Büttner, G., Hardy, I., Klieme, E., et al. (2019). The effects of teacher competence on student outcomes in elementary science education: The mediating role of teaching quality. *Teaching and Teacher Education, 86*, 102882.

Ferguson, R. F., & Danielson, C. (2014). How framework for teaching and tripod 7Cs evidence distinguish key components of effective teaching. In T. J. Kane, K. A. Kerr, & R. C. Pianta (Eds.), *Designing teacher evaluation systems* (pp. 98–143). New York: Wiley.

Fischer, C., Fishman, B., Levy, A. J., Eisenkraft, A., Dede, C., Lawrenz, F., et al. (2016). When do students in low-SES schools perform better-than-expected on a high-stakes test? Analyzing school, teacher, teaching, and professional development characteristics. *Urban Education*, 0042085916668953.

Goe, L. (2007). *The link between teacher quality and student outcomes: A research synthesis*. Washington, DC: National Comprehensive Center for Teacher Quality.

Gustafsson, J.-E., & Nilsen, T. (2017). *Changes in Mathematics performance from 2011 to 2015 in Norway and Sweden as a function of teacher professional development*. Paper presented at the ECER 2017, Copenhagen, Denmark.

Gustafsson, J.-E., Nilsen, T., & Hansen, K. Y. (2018). School characteristics moderating the relation between student socio-economic status and mathematics achievement in grade 8. Evidence from 50 countries in TIMSS 2011. *Studies in Educational Evaluation, 57*, 16–30.

Gustafsson, J.-E., & Yang Hansen, K. (2017). Changes in the impact of family education on student educational achievement in Sweden 1988–2014. *Scandinavian Journal of Educational Research, 62*(5), 719–736.

Hansen, K. Y. (2015). *Trend in educational inequality in the Nordic countries between 2000 and 2009: Evidence from OECD PISA studies*. Paper presented at the ECER, Budapest, Hungary.

Hansson, Å., & Gustafsson, J.-E. (2017). Pedagogisk segregation: Lärarkompetens i den svenska grundskolan ur ett likvärdighetsperspektiv [Pedagogic segregation. Teacher competence in the Swedish elementary school from an equity perspective]. *Pedagogisk Forskning i Sverige* [Pedagogic Research in Sweden], *21*(1–2), 56–78.

Haugsbakk, G. (2013). From Sputnik to PISA shock – New technology and educational reform in Norway and Sweden. *Education Inquiry, 4*(4), 23–22. https://doi.org/10.3402/edui.v4i4.23222

Heck, R. H. (2007). Examining the relationship between teacher quality as an organizational property of schools and students' achievement and growth rates. *Educational Administration Quarterly, 43*(4), 399–432.

Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. *Second Handbook of Research on Mathematics Teaching and Learning, 1*, 371–404.

Hox, J. J., Moerbeek, M., & Van de Schoot, R. (2017). *Multilevel analysis: Techniques and applications*. New York: Routledge.

Hwang, J., Choi, K. M., Bae, Y., & Shin, D. H. (2018). Do teachers' instructional practices moderate equity in mathematical and scientific literacy?: An investigation of the PISA 2012 and 2015. *International Journal of Science and Mathematics Education, 16*(1), 25–45.

Jeynes, W. H. (2005). A meta-analysis of the relation of parental involvement to urban elementary school student academic achievement. *Urban Education, 40*(3), 237–269.

Kaarstein, H., Nilsen, T., & Blömeke, S. (2016). Lærerkompetanse [Teacher competence]. In O. K. Bergem, H. Kaarstein, & T. Nilsen (Eds.), *Vi kan lykkes i realfag* [We can succeed in mathematics and science]. Oslo, Norway: Universitetsforlaget.

Kim, S. W., Cho, H., & Kim, L. Y. (2019). Socioeconomic status and academic outcomes in developing countries: A meta-analysis. *Review of Educational Research, 89*(6), 875–916.

Kirsten, N., & Wermke, W. (2017). Governing teachers by professional development: State programmes for continuing professional development in Sweden since 1991. *Journal of Curriculum Studies, 49*(3), 391–411.

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

Klingebiel, F., & Klieme, E. (2016). Teacher qualifications and professional knowledge. In E. Klieme, S. Kuger, N. Jude, & D. Kaplan (Eds.), *Assessing contexts of learning* (pp. 447–468). Cham, Switzerland: Springer.

Kraft, M. A., Blazar, D., & Hogan, D. (2018). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. *Review of Educational Research, 88*(4), 547–588.

Kuger, S., Klieme, E., Jude, N., & Kaplan, D. (2016). *Assessing contexts of learning: An international perspective*. Cham, Switzerland: Springer.

Kyriakides, L., Creemers, B., & Antoniou, P. (2009). Teacher behaviour and student outcomes: Suggestions for research on teacher training and professional development. *Teaching and Teacher Education, 25*(1), 12–23.

Lagerstrøm, B. O., Moafi, H., & Revold, M. K. (2014). *Kompetanseprofil i grunnskolen Hovedresultater 2013/2014* [Competence profile in elementary school. Main results 2013/2014]. Retrieved from Oslo-Kongsviner. https://www.ssb.no/utdanning/artikler-og-publikasjoner/_attachment/197751?_ts=148a1618d30

Lankford, H., Loeb, S., & Wyckoff, J. (2002). Teacher sorting and the plight of urban schools: A descriptive analysis. *Educational Evaluation and Policy Analysis, 24*(1), 37–62.

Lindvall, J. (2017). Two large-scale professional development programs for mathematics teachers and their impact on student achievement. *International Journal of Science and Mathematics Education, 15*(7), 1281–1301.

Lundahl, C., & Waldow, F. (2009). Standardisation and 'quick languages': The shape-shifting of standardised measurement of pupil achievement in Sweden and Germany. *Comparative Education, 45*(3), 365–385.

Lundström, U. (2015). Systematic quality work in Swedish schools: Intentions and dilemmas. *Scandinavian Journal of Public Administration, 19*(1), 23–44.

Marsh, H. W., Lüdtke, O., Robitzsch, A., Trautwein, U., Asparouhov, T., Muthén, B., et al. (2009). Doubly-latent models of school contextual effects: Integrating multilevel and structural equation approaches to control measurement and sampling error. *Multivariate Behavioral Research, 44*(6), 764–802.

Martin, M. O., Mullis, I. V. S., Foy, P., & Hooper, M. (2016). *TIMSS 2015 international results in science*. Chestnut Hill, MA: Boston Collage, TIMSS & PIRLS International Study Centre.

Ministry of Education and Research. (2008). *Quality in school*. Oslo, Norway: Ministry of Education and Research.

Ministry of Education and Research. (2015). *Competence for quality. Strategy for professional development for teachers and school leaders towards 2025*. Oslo, Norway: Ministry of Education and Research.

Minner, D. D., Levy, A. J., & Century, J. (2010). Inquiry-based science instruction – What is it and does it matter? Results from a research synthesis years 1984 to 2002. *Journal of Research in Science Teaching, 47*(4), 474–496.

Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2016). *TIMSS 2015 international results in mathematics*. Boston: Boston College, TIMSS & PIRLS International Study Center.

Mullis, I. V. S., Martin, M. O., Goh, S., & Cotter, K. (2016). *TIMSS 2015 encyclopedia: Education policy and curriculum in mathematics and science*. Boston: Boston College, TIMSS & PIRLS International Study Center.

Munthe, E., Malmo, K.-A. S., & Rogne, M. (2011). Teacher education reform and challenges in Norway. *Journal of Education for Teaching, 37*(4), 441–450. https://doi.org/10.1080/0260747 6.2011.611012

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide* (8th ed.). Los Angeles: Muthén & Muthén.

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), 26-pages.

Nilsen, T., Bjørnsson, J. K., & Olsen, R. V. (2018). 7. Hvordan har likeverd i norsk skole endret seg de siste 20årene? [How has equity changed over the last 20 years?]. *Tjue år med TIMSS og PISA i Norge* [Twenty years of TIMSS and PISA in Norway] (pp. 150–172).

Nilsen, T., & Gustafsson, J.-E. (2016). *Teacher quality, instructional quality and student outcomes: Relationships across countries, cohorts and time. IEA research for education. Volume 2*. International Association for the Evaluation of Educational Achievement ….

Nilsen, T., Scherer, R., & Blömeke, S. (2018). 3. The relation of science teachers' quality and instruction to student motivation and achievement in the 4th and 8th grade: A Nordic. *Northern Lights on TIMSS and PISA 2018, 61*.

Norris, S. P., & Phillips, L. M. (2003). How literacy in its fundamental sense is central to scientific literacy. *Science Education, 87*(2), 224–240.

Nortvedt, G. A. (2018). Policy impact of PISA on mathematics education: The case of Norway. *European Journal of Psychology of Education, 33*(3), 427–444.

Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis, 26*(3), 237–257.

OECD. (2016). *PISA 2015 results (Volume I). Excellence and equity in education*. Paris: OECD Publishing.

OECD. (2019). *Improving school quality in Norway: The new competence development model. Implementing education policies*. Paris: OECD Publishing.

Opheim, V. (2004). *Equity in education: Country analytical report Norway*. Oslo, Norway: NIFU STEP.

Pianta, R., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher, 38*(2), 109–119.

Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods, 15*(3), 209.

Qin, L., & Bowen, D. H. (2019). The distributions of teacher qualification: A cross-national study. *International Journal of Educational Development, 70*, 102084.

Raudenbush, S. W. (2008). Advancing educational policy by advancing research on instruction. *American Educational Research Journal, 45*(1), 206–230.

Regjeringen. (2014). *Over 3000 lærere får tilbud om videreutdanning* [Over 3000 teachers are offered professional development]. Regjeringen. Retrieved from https://www.regjeringen.no/no/aktuelt/Over-3000-larere-far-tilbud-om-videreutdanning/id757511/

Ringarp, J. (2016). PISA lends legitimacy: A study of education policy changes in Germany and Sweden after 2000. *European Educational Research Journal, 15*(4), 447–461.

Ringarp, J., & Parding, K. (2018). I otakt med tiden?: Lärarprofessionens ställning sett via lärarut-bildningens utveckling i Sverige, 1962–2015 [At odds with time? The position of the teaching profession seen through the development of teacher education in Sweden]. In M. Buchardt, J. E. Larsen, & K. B. Staffensen (Eds.), *Professionerne og deres uddannelser* [The professionals and their education] (pp. 49–68). Köpenhamn, Denmark: Selskabet for skole- og uddannelseshistorie.

Rivers, J. C., & Sanders, W. L. (2002). Teacher quality and equity in educational opportunity: Findings and policy implications. *Teacher Quality*, 13–23.

Rjosk, C., Richter, D., Hochweber, J., Lüdtke, O., Klieme, E., & Stanat, P. (2014). Socioeconomic and language minority classroom composition and individual reading achievement: The mediating role of instructional quality. *Learning and Instruction, 32*, 63–72.

Ryoo, K. (2009). *Learning science, talking science: The impact of a technology-enhanced curriculum on students' science learning in linguistically diverse mainstream classrooms.* Unpublished doctoral dissertation, Stanford University, Stanford, CA, USA.

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

Skolverket. (2016). *TIMSS 2015. Svenska grundskoleelevers kunskaper i matematik och natur-vetenskap i ett internationellt perspektiv* [TIMSS 2015. Swedish elementary school students' knowledge of mathematics and science in an international perspective]. Retrieved from Stockholm. https://www.skolverket.se/skolutveckling/forskning-och-utvarderingar/internationella-jamforande-studier-pa-utbildningsomradet/timss-internationell-studie-om-kunskaper-i-matematik-och-naturvetenskap-hos-elever-i-arskurs-4-och-8#h-Rapporter

Skolverket. (2019). *Lärarlegitimation och förskollärarlegitimation* [Teacher certificate and Preschool Teacher certificate]. Retrieved from https://www.skolverket.se/regler-och-ansvar/lararlegitimation-och-forskollararlegitimation

Tan, C. Y., Lyu, M., & Peng, B. (2019). Academic benefits from parental involvement are stratified by parental socioeconomic status: A meta-analysis. *Parenting*, 1–47. https://doi.org/10.1080/15295192.2019.1694836

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

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

Timperley, H., Wilson, A., Barrar, H., & Fung, I. (2007). *Teacher professional development and learning: Best evidence synthesis iteration (BES)*. Wellington, New Zealand: Ministry of Education.

Tveit, S. (2013). Educational assessment in Norway. *Assessment in Education: Principles, Policy & Practice*, 1–17. https://doi.org/10.1080/0969594X.2013.830079

UNESCO. (2018). *Handbook on measuring equity in education*. UNESCO Institute for Statistics FHI 360 Education Policy Data Centre, Oxford Policy Management, University of Cambridge: UNESCO Institute for Statistics, Montreal, Quebec.

van Tartwijk, J., & Hammerness, K. (2011). The neglected role of classroom management in teacher education. *Teaching Education, 22*(2), 109–112.

Willms, J. D. (2010). School composition and contextual effects on student outcomes. *The Teachers College Record, 112*(4), 3–4.

Wilson, S. M. (2013). Professional development for science teachers. *Science, 340*(6130), 310–313.

Chapter 8
The Case for Good Discipline? Evidence on the Interplay Between Disciplinary Climate, Socioeconomic Status, and Science Achievement from PISA 2015

Ronny Scherer (ID)

Abstract In both educational and psychological research, the relation between socioeconomic status (SES) and academic achievement is the most widely examined contextual effect. While several research syntheses have reported evidence of positive and significant SES–achievement relations (i.e., higher SES is associated with better academic achievement in several domains), they also reported substantial variation across educational contexts, such as classrooms, schools, and educational systems, and proposed mechanisms underlying these relations. This chapter addressed this variation and tested three hypotheses on the interplay between socioeconomic status, the disciplinary climate in science lessons, and science achievement—the *compensation hypothesis*, the *mediation hypothesis*, and the *moderation hypothesis*. Utilizing the Programme for International Student Assessment (PISA) 2015 data from the Nordic countries (Denmark, Finland, Iceland, Norway, and Sweden), multilevel structural equation modeling provided evidence to test the contextual, indirect, and cross-level interaction effects. While evidence for the compensation hypothesis existed in most Nordic countries, evidence supporting the mediating and moderating roles of the disciplinary climate for the SES–achievement relation was sparse.

Keywords Disciplinary climate · Multilevel structural equation modeling · Programme for International Student Assessment (PISA) · Science achievement

R. Scherer (✉)
Centre for Educational Measurement, University of Oslo, Oslo, Norway
e-mail: ronny.scherer@cemo.uio.no

Good classroom discipline, an orderly learning environment, and few disruptions of instruction are considered prerequisites for a good school climate and instructional quality. While most of the extant research has been concerned with establishing that a disciplinary climate—a climate that requires the definition of desirable student behaviors and the prevention of undesirable student behaviors (Hochweber, Hosenfeld, & Klieme, 2014)—is significantly related to academic achievement (Berkowitz, Moore, Astor, & Benbenishty, 2017), less effort has been made to establish this relation in the context of equity or equality (Atlay, Tieben, Hillmert, & Fauth, 2019). Specifically, moving beyond merely describing the socioeconomic status (SES)–achievement relation as an indicator of (in-)equality, researchers and policy makers have become more and more interested in studying the following: (a) the extent to which a disciplinary climate may compensate for the effect of SES on academic achievement, (b) the mechanisms behind the relations among SES, achievement, and disciplinary climate, and (c) the extent to which the disciplinary climate may decrease possible achievement gaps between students of different SES (Berkowitz et al., 2017; Ning, Van Damme, Van Den Noortgate, Yang, & Gielen, 2015). However, the body of evidence clarifying the role disciplinary climate plays for SES, academic achievement, and the SES–achievement relation is diverse. For instance, while some evidence suggests that a disciplinary climate is directly related to achievement above and beyond SES (Bellens, Van Damme, Van Den Noortgate, Wendt, & Nilsen, 2019), some evidence suggests that it may mediate the relation between SES and achievement (Liu, Van Damme, Gielen, & Van Den Noortgate, 2015). Some further evidence suggests that a good disciplinary climate moderates the SES–achievement relation (Ning et al., 2015). This diversity in the nature of the relations among SES, disciplinary climate, and academic achievement ultimately results in different interpretations of the role disciplinary climate plays: While some researchers may conclude that a good disciplinary climate is related to better achievement independent of students' or schools' SES, others may conclude that a good disciplinary climate is more likely to occur in high-SES schools, resulting in better achievement. Finally, other researchers may conclude that a good disciplinary climate is associated with smaller achievement gaps—in other words, in schools with a good disciplinary climate, the achievement differences are hardly retraceable to SES differences. In the extant literature, these three perspectives have been summarized in three hypotheses—namely, the *compensation*, *mediation*, and *moderation hypotheses* (Berkowitz et al., 2017). Through the lenses of these hypotheses, the Programme for International Student Assessment (PISA) 2015 data of the five Nordic countries (Denmark, Finland, Iceland, Norway, and Sweden) were analyzed, and the evidence base for or against a compensation, mediation, or moderation mechanism describing the relations among SES, disciplinary climate, and science achievement was examined. Ultimately, the resultant evidence could clarify the role of disciplinary climate for SES, achievement, and the SES–achievement relation for the PISA 2015 Nordic country data and highlight plausible conclusions that could be drawn in the context of equity and equality. Following the framework proposed by Willms and Tramonte (2019), this study considers the relation between SES and disciplinary climate an indicator of *equity* (i.e., representing possible differences in

the opportunities to access a good disciplinary climate in school science lessons), while the relation between SES and science achievement is seen as an indicator of *equality* (i.e., representing possible differences in educational outcomes). These conceptualizations resonate with the "equality–equity model" proposed by Espinoza (2007), which can be characterized as follows: (a) possible SES differences in disciplinary climate may represent differences in access to education, or more precisely, access to the same quality of education to address basic educational needs; and (b) possible SES differences in science achievement (i.e., educational achievement based on test performance in the dimension of "output") represent inequalities for students across social groups.

8.1 Theoretical Framework

8.1.1 Disciplinary Climate and Academic Achievement

The disciplinary climate represents one of the most extensively studied aspects of schooling and instruction (Atlay et al., 2019; Seidel & Shavelson, 2007). Although a plethora of conceptualizations exist, the extant body of literature seems to converge in that the disciplinary climate represents a climate in schools and/or classrooms that requires the identification of desirable and the prevention of undesirable student behaviors (Hochweber et al., 2014). This conceptualization clearly goes beyond strategies to handle disruptive behavior in educational settings (Atlay et al., 2019) and comprises instructional approaches, such as setting and communicating classroom rules, establishing routines, providing an orderly and functional classroom or school setting, monitoring school and/or classroom activities, and intervening if necessary (e.g., Hochweber et al., 2014; Seidel & Shavelson, 2007). To add to the complexity, teachers must adapt these approaches to the specific classroom or school contexts, especially in socially diverse settings with substantial variation in SES or minority status (Emmer & Stough, 2001; Rjosk et al., 2014). In this sense, establishing a good disciplinary climate is considered part of teacher competence, and the instructional approaches taken to accomplish it is part of instructional quality (Lipowsky et al., 2009; Seidel & Shavelson, 2007). Despite this anchoring in the instructional and professional teacher competence frameworks, the disciplinary climate concept has also found its way into the frameworks of school climate. In these frameworks, a good disciplinary climate is a subdimension of school safety and comprises conflict resolution; clarity, fairness, and consistency of rules; and the belief in school rules (M.-T. Wang & Degol, 2015). Bringing together the conceptualizations of disciplinary climate as part of instructional quality and school climate, Scherer and Nilsen (2017) found that a safe and orderly school environment is also characterized by good classroom management, which can result in better school achievement. Moreover, a good disciplinary climate forms the prerequisite for engaging in other instructional activities, such as cognitive activation and teacher

support (Klieme, Pauli, & Reusser, 2009). In this sense, the disciplinary climate helps teachers create learning environments that support students' learning.

A large body of research testifies to the consistently positive and significant relation between a good disciplinary climate and academic achievement across educational contexts, subject areas, and countries (e.g., Bellens et al., 2019; Berkowitz et al., 2017; Seidel & Shavelson, 2007; M. C. Wang, Haertel, & Walberg, 1993). However, this relation may vary in individual-level (student) data in which perceptions of disciplinary climate are assessed and classroom- or school-level data in which aggregated perceptions of disciplinary climate are evaluated with a certain reliability. For instance, Fauth, Decristan, Rieser, Klieme, and Büttner (2014) found a significant correlation between disciplinary climate and academic achievement for the classroom level but not the student level. In their study of eighth-graders, Blank and Shavit (2016) found significant relations at the student and classroom level but not at the school level. Considering this variation, the specification of the appropriate level of analysis is critical to interpreting the relation between disciplinary climate and academic achievement (Marsh et al., 2012).

8.1.2   Socioeconomic Status and Academic Achievement

SES represents the social standing or class of an individual or group and comprises measures of parental education, income, and occupation (APA, 2006; Willms & Tramonte, 2019). The concept serves as a proxy for possible inequalities with respect to students' background, and it has been studied extensively in relation to educationally relevant outcome variables, especially academic achievement (Thomson, 2018). This perspective focuses on achievement as the output of education and quantifies the possible influence of unequal conditions (SES) on it (i.e., inequalities on average across social groups; Espinoza, 2007). Given the popularity of this perspective, a plethora of studies examining the SES–achievement relation exists across academic domains and school subjects. While reviewing this large body of research is beyond the scope of this chapter, the chapter brings to attention some knowns and unknowns.

Several research syntheses have agreed that a statistically significant and positive relation between SES and academic achievement exists across domains, SES measures, and measures of academic achievement (e.g., Broer, Bai, & Fonseca, 2019; Harwell, Maeda, Bishop, & Xie, 2016; Kim, Cho, & Kim, 2019; Scherer & Siddiq, 2019; Sirin, 2005; van Ewijk & Sleegers, 2010; White, 1982). Despite this consistent finding, the corresponding effect sizes ranged from small ($\bar{r} = 0.12$) to moderate ($\bar{r} = 0.32$) coefficients and varied across study, sample, and measurement characteristics (e.g., gender and grade-level composition in the sample, country of origin, types of achievement measures). Moreover, the statistical approaches most data analysts have taken to describe SES–achievement relations have been limited to correlational analyses of student-level data (Willms & Tramonte, 2019). This observation brings to light one key issue, that is, the appropriate level of analysis at

which the SES–achievement relation is described. Clearly, students' SES has a substantive meaning for individual students and is considered a powerful variable explaining achievement differences between students. At the same time, SES has a substantive meaning for classrooms and schools, representing the classroom or school SES composition (Thomson, 2018). Recognizing that SES and academic achievement can also be related at some level of clustering requires a multilevel approach to describing achievement gaps and composition effects (Marsh et al., 2009).

8.1.3   Three Hypotheses on the Interplay Between Disciplinary Climate, SES, and Academic Achievement

Bringing together the two lines of research describing the relation between SES and academic achievement and the relation between disciplinary climate and academic achievement, the core question this chapter assesses is how these three concepts play together. More specifically, while both lines of research have established significant links between the two pairs of concepts, the role of the disciplinary climate—as an aspect of both school climate and instructional quality—in academic achievement after controlling for SES, as well as the relation between academic achievement and SES, remains unclear.

Berkowitz et al. (2017) argued that the "scientific evidence establishing directional links and mechanisms between SES, school climate, and academic performance is inconclusive" (p. 425), especially due to the different perspectives educational researchers have taken to describe these links and mechanisms. Synthesizing these perspectives in 78 empirical studies, the authors identified three core hypotheses that describe the interplay between aspects of school climate, SES, and academic achievement; these are the compensation, mediation, and moderation hypotheses (see Fig. 8.1).

The *compensation hypothesis* assumes that the disciplinary climate explains variation in academic achievement at the student and school levels above and beyond SES (Fig. 8.1a). It further assumes that the disciplinary climate contributes to "academic achievement beyond the expected outcomes based on SES background" (Berkowitz et al., 2017, p. 426). In this sense, support for this hypothesis could be interpreted as evidence for a compensating effect of disciplinary climate. Notably, this hypothesis does not make any assumptions on the link between SES and disciplinary climate—it only considers these two concepts as explanatory variables of academic achievement side-by-side, and therefore, it is commonly tested using contextual or single-level regression models. In their systematic review, Berkowitz et al. (2017) noticed that the compensation hypothesis is the dominating perspective researchers take to describe the interplay between SES, achievement, and climate variables. In the context of large-scale international assessments, indeed, many studies tested this hypothesis and obtained evidence that climate
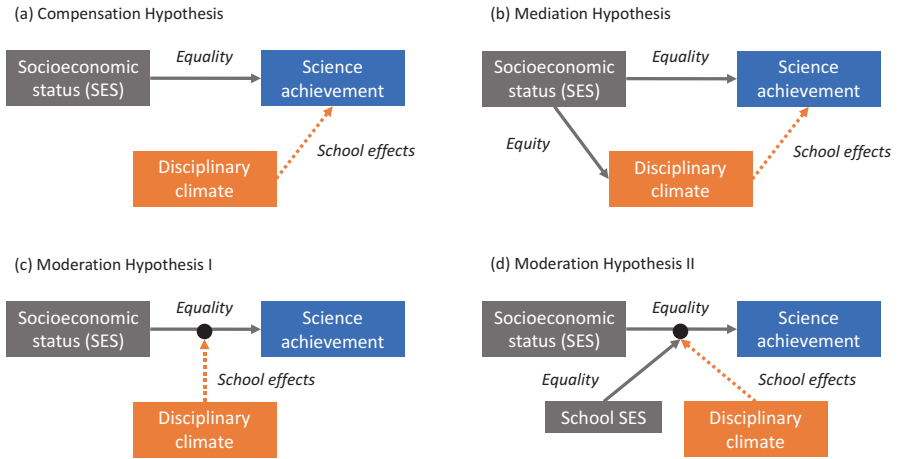
Fig. 8.1 Conceptual models framing of the relations among the three constructs. (Adopted from Willms & Tramonte, 2019)

variables (represented as instructional quality or school climate) were significantly (and positively) related to academic achievement beyond SES at the student level and some level of clustering (e.g., Bellens et al., 2019; Ning et al., 2015; Rjosk et al., 2014; Shin, Lee, & Kim, 2009). This hypothesis takes the perspective of equality as it describes the relation between SES and educational outcomes—however, it only considers the additional variance explanation in educational outcomes through instructional variables (i.e., schooling) without a link between differences in SES and differences in disciplinary climate.

The *mediation hypothesis* assumes a mechanism underlying the relation between SES and academic achievement via disciplinary climate (Fig. 8.1b). Researchers testing this hypothesis argue that "a school's SES influences its social climate, which in turn influences academic achievement" (Berkowitz et al., 2017, p. 426). In this sense, schools with a low average SES may struggle with establishing safe and orderly learning environments, and thus, be more likely to show low achievement (G. Chen & Weikart, 2008). Despite the causal claims behind this hypothesis, it is worth noting that the mediation mechanism is considered a school- or classroom-level mechanism rather than a student-level one (Liu et al., 2015). However, classroom or school climate variables are often assessed via student ratings, which are aggregated to the classroom or school level (Marsh et al., 2012); this allows researchers to test this hypothesis for individual students' perceptions. In a slightly different context, Schmidt, Burroughs, Zoido, and Houang (2015) tested for student-level mediation and found support for significant indirect effects of individual SES on academic achievement via perceptions of opportunities to learn. In contrast to the moderation hypothesis, the mediation hypothesis adds the link between SES and disciplinary climate and an equity perspective to the compensation hypothesis by considering possible gaps in encountering or having access to a positive disciplinary climate. It also assumes a sequence of relations among variables, that is, SES →

Disciplinary Climate → Achievement. Such a sequence entails that variation in disciplinary climate may be due to variation in SES, while variation in achievement may be due to variation in the disciplinary climate. Typically, multilevel mediation models are used to test this hypothesis (Preacher, Zyphur, & Zhang, 2010). This hypothesis takes the perspective of equity, as it describes the relation between SES and instructional variables (i.e., opportunities to experience instructional quality even with different needs resulting from varying socioeconomic background) and the perspective of equality, as it describes the SES–achievement relation. However, the SES–achievement relation is established only in the case of partial mediation and does not exist in the case of full mediation.

The *moderation hypothesis* assumes that the disciplinary climate may explain variation in the relation between students' SES and their individual achievement across classrooms or schools (Fig. 8.1c). In other words, classrooms or schools of different disciplinary climate may show different SES–achievement relations (Berkowitz et al., 2017). In the case of negative moderation effects, a positive disciplinary climate is associated with smaller achievement gaps in classrooms or schools (Nilsen, Bloemeke, Yang Hansen, & Gustafsson, 2016). However, some empirical studies found positive moderation effects that pointed to a widening of the achievement gaps with better disciplinary climate (Gustafsson, Nilsen, & Hansen, 2018), while others could not identify any significant moderation (Bellens et al., 2019). Typically, researchers use cross-level interaction models to test the moderation hypothesis and address the extent to which differences in classroom or school conditions are associated with smaller achievement gaps (Jehangir, Glas, & van den Berg, 2015). Put differently, school conditions may facilitate the reduction of inequalities among students and/or improve their educational outputs irrespective of their background. A variation of this hypothesis includes classroom or school SES as another predictor of SES–achievement next to disciplinary climate (Fig. 8.1d). This variation allows researchers to examine the moderation effects of disciplinary climate above and beyond those of SES. Although the moderation effects are interpreted in a way that establishes disciplinary climate as the moderator, the empirical models testing these effects also allow for an alternative interpretation, in which SES is considered the moderator. Such an interpretation would entail that the relation between disciplinary climate and science achievement is smaller in high-SES schools than it is in low-SES schools. The moderation hypothesis takes the perspective of equality, describing the relation between SES and educational outcomes and considering possible moderation effects to be effects of schooling (Willms & Tramonte, 2019). In this sense, disciplinary climate may decrease possible inequalities in educational achievement across social groups (Espinoza, 2007).

The three hypotheses represent three lenses through which the interplay between disciplinary climate as an aspect of school climate and instructional quality, SES, and achievement can be examined.

8.1.4   The Present Study

This study focuses on the relations between disciplinary climate, socioeconomic status, and achievement in the context of science. The reasons for focusing on the context of science education are manifold: First, science is considered a core subject across many educational systems, including those of the Nordic countries, and it is a core domain of the existing large-scale assessments, such as PISA and TIMSS, which inform educational policy making (Kavli, 2018). Second, many educational systems struggle to provide equal opportunities for students to learn science; such inequalities may result in less frequent career choices in science, and they may ultimately pose a threat to national economic and technological competitiveness and equity (OECD, 2017a). Third, career choices in science are not determined only by students' attitudes toward and motivation to learn science; rather, a remarkable body of research has shown that this aspiration is also determined by students' home background, the distribution of capital, and parents' social status (Archer et al., 2012). Fourth, many countries around the world are promoting science education to provide students with equal opportunities to learn the subject (Bianchini, 2017). Fifth, inequalities in science education and achievement may create inequalities in science capital and vice versa; such inequalities affect students' participation in society as scientifically literate citizens (Archer, Dawson, DeWitt, Seakins, & Wong, 2015).

Utilizing the PISA 2015 data of the Nordic countries (Denmark, Finland, Iceland, Norway, and Sweden), the secondary analyses were aimed at examining the evidence for the three dominating hypotheses on the role of disciplinary climate: the compensation, mediation, and moderation hypotheses (see Fig. 8.1). In light of these three hypotheses, this chapter addresses the following three research questions (RQs):

RQ 1   *To what extent does disciplinary climate explain variations in science achievement above and beyond socioeconomic status?*
RQ 2   *To what extent does disciplinary climate mediate the relation between socioeconomic status and science achievement?*
RQ 3   *To what extent does the disciplinary climate explain between-school variation in the relation between socioeconomic status and science achievement?*

Given that indicators of disciplinary climate are commonly assessed via students', parents', teachers', or principals' reports (M.-T. Wang & Degol, 2015), these assumptions may hold not only at the individual (within) level, where *perceptions* of the disciplinary climate are in the focus, but also at the aggregated (between) level, where shared perceptions about the school are in focus (Marsh et al., 2012). In other words, the three hypotheses may be tested for different levels of analysis— in PISA 2015, these levels refer to the student and the school level, with disciplinary climate assessed via student reports. Accounting for the multilevel nature of the data, this study considers several types of specificity via the following approaches: (a) This study compares the evidence for the three hypotheses across the five

participating Nordic countries, taking a comparative perspective, and at the same time, allowing for *country specificity*; (b) as noted above, this study tests the three hypotheses for the student *and* the school level, accounting for *level specificity*; and (c) this study explores the role of disciplinary climate for the relation between SES and science achievement across the three core dimensions of SES (i.e., education, income, and occupation; APA, 2006), allowing for *SES measurement specificity*. The information about the extent to which the three hypotheses can or cannot be supported across these specific conditions adds to the evidence base on the interplay between disciplinary climate, socioeconomic status, and academic achievement. To summarize, the present study examines disciplinary climate in science lessons in terms of the following issues: (a) whether it explains variation in science achievement above and beyond SES, (b) whether it mediates the relation between SES and science achievement, and (c) whether it moderates the relation between SES and science achievement. In this respect, the relation between SES and disciplinary climate (i.e., students' reported disciplinary climate in the schools they were placed in) was considered to be an indicator of equity and interpreted as the degree to which students were given opportunities to access a good disciplinary climate in science lessons. The relation between SES and science achievement was considered an indicator of (in-)equality that provides information about the degree to which SES differences in achievement exist (Espinoza, 2007; Willms & Tramonte, 2019).

Although the approach taken in this study was guided by three hypotheses in the context of equity and equality, the country comparisons were mainly exploratory, especially with respect to the evidence for or against the existence of a "Nordic model." Despite the lack of a clear definition and a measurable framework of a Nordic model of education (Lundahl, 2016), the main goals of the Nordic school systems converge in that they strive for equity, participation, and welfare (Antikainen, 2006). However, these commonalities do not ensure that equal opportunities to learn, or in the context of this study, equal access to a good and positive disciplinary climate, exist across the Nordic countries. In fact, there is some evidence of substantial differences between them (OECD, 2017b; Sortkær & Reimer, 2018). Moreover, the existing international large-scale assessments suggest that the Nordic countries are far from scoring equally in the core domains of reading, science, and mathematics, and although relatively small, differences in measures of SES have arisen (OECD, 2016, 2019). Hence, exploring the differences and similarities in the information the three hypothesized models provide about the interplay of SES, science achievement, and disciplinary climate addresses whether evidence for a Nordic model exists in relation to the present models. For instance, possible differences in the contextual effects of schools' disciplinary climate on students' science achievement after controlling for SES may point to the fact that the possibilities to contribute to a better science achievement above and beyond the SES differences may not be equally exploited or provided across the Nordic countries. At the same time, such cross-country differences should not be overinterpreted as evidence against a Nordic model of education, especially because of the lack of a clear-cut framework that defines the dimensions and indicators of the model and because common efforts to create equity in the Nordic countries may not necessarily lead to the same results in

education systems (Blossing, Imsen, & Moos, 2014; Lundahl, 2016). In this sense, the present study explores rather than hypothesizing on cross-country differences and similarities in the proposed models and does not argue that similarities have been caused by a "Nordic model."

8.2   Data and Methodological Approaches

8.2.1   PISA 2015 Science Data of the Nordic Countries

The sample underlying the secondary analyses of the PISA 2015 data comprised the student samples of five Nordic countries, namely, Denmark, Finland, Iceland, Norway, and Sweden. Table 8.1 provides a brief summary of the corresponding sample sizes and the intraclass correlations (ICC_1) of the relevant variables. Each variable exhibited substantial between-school variation, and thus, allowed for decomposing their variances into the corresponding within and between parts (Snijders & Bosker, 2012). Notably, the smallest intraclass correlation for science achievement was apparent for the Icelandic data, while the Swedish data exhibited the largest coefficient. The disciplinary climate scale score varied the most between schools for the Norwegian data and the least for the Finnish data. Finally, between-school variation in the SES measures varied across the measures; nonetheless, consistently across the Nordic countries, the least variation occurred for the home possessions (HOMEPOS) measure.

8.2.1.1   Science Achievement

In PISA 2015, the concept of scientific literacy comprised the three following core competencies: explaining phenomena scientifically, evaluating and designing scientific enquiry, and interpreting data and evidence scientifically (OECD, 2017a).

Table 8.1 Description of the Nordic country samples included in the secondary data analyses

Country	Students N	Schools n	Intraclass correlation ICC_1				
			Science achievement	HISEI	HOMEPOS	PARED	Disciplinary climate
Denmark	7161	333	0.173	0.164	0.098	0.102	0.159
Finland	5882	168	0.159	0.132	0.098	0.131	0.120
Iceland	3371	124	0.042	0.111	0.037	0.072	0.165
Norway	5456	229	0.093	0.062	0.044	0.052	0.212
Sweden	5458	202	0.211	0.164	0.072	0.074	0.184
Overall	27328	1056	0.179	0.153	0.106	0.191	0.180

Note. Cases with completely missing data on all relevant variables were excluded. The ICC_1 of the WLE score for disciplinary climate (DISCLISCI) is reported here

Through a series of tasks requiring these competencies in the content domains labeled "Physical," "Living," and "Earth and Space," students' science achievement was measured and represented as a set of plausible values (OECD, 2017b). The secondary analyses included the plausible values PV1SCIE-PV10SCIE as indicators of the overall scientific literacy, yet not the plausible values specific to the three competencies or the content domains due to their high intercorrelations. Readers are kindly referred to the PISA 2015 Technical Report for more details about the psychometric properties and the design of the scientific literacy assessment (OECD, 2017b).

8.2.1.2 Socioeconomic Status

Students' socioeconomic status was measured by several indicators in PISA 2015. These indicators were summarized in three subscale scores by means of item response theory modeling as follows: highest parental education (HISEI), parental education (PARED), and HOMEPOS. Performing principal component analysis, these three scores were then combined with composite SES indicators, namely, the Index of Economic, Social, and Cultural Status (ESCS). Given the psychometric issues associated with this composite SES score (Cronbach's α values ranged between 0.53 and 0.65 for the Nordic countries; see OECD, 2017b), this chapter presents the results of the separate analyses for each of the three subscale scores. Moreover, due to the considerable heterogeneity of factor loadings within and between countries, SES was not represented as a latent variable measured by the three subscales scores to avoid biased estimates of structural parameters in structural equation models (Rhemtulla, van Bork, & Borsboom, 2019).

8.2.1.3 Disciplinary Climate in School Science Lessons

The disciplinary climate in school science lessons was assessed by students' ratings of five statements on a four-point scale ranging from 0 (*Never or hardly ever*) to 3 (*Every lesson*) (ST097; see OECD, 2017b). Some of these statements addressed the same aspect of disciplinary climate (e.g., "Students don't listen to what the teacher says" [Q01] and "The teacher has to wait a long time for students to quiet down" [Q03]), and a two-level confirmatory factor analysis suggested that residual covariances among two pairs of items existed (i.e., Q04 − Q05, Q01–Q03) beyond a within and a between latent variable representing disciplinary climate. To circumvent these redundancies and avoid construct-irrelevant multidimensionality, the three items—Q01, Q02, and Q04—served as manifest indicators. The correlation between the within- and between-level latent variables with the scale score DISCLISCI and the perfect correlations found provide some evidence for the validity of this approach. The within-level reliabilities ranged between $\omega_W = 0.79$ and $\omega_W = 0.84$, and the between-level reliabilities ranged between $\omega_B = 0.98$ and $\omega_B = 0.99$ across countries for the disciplinary climate scale comprising the three items.

8.2.2 Multilevel Structural Equation Modeling of the PISA 2015 Science Data

8.2.2.1 Analytic Setup

To test the models representing the three hypotheses (see Table 8.1), multilevel structural equation modeling (MSEM) described the measurement and structural models at the student (within) and school (between) levels in the statistical software package M*plus* 8.2 (Muthén & Muthén, 1998-2017). The representation of the corresponding statistical models is provided in Fig. 8.2; a more detailed description within the MSEM framework can be found in the Supplementary Material A1 (see Figs. A1 and A2). Extending multilevel regression modeling, MSEM allows researchers to account not only for sampling error but also measurement error using latent variables (Marsh et al., 2009). The observed variables were decomposed into their latent within and between parts and specified the corresponding measurement and structural models to test the three hypotheses. Specifically, disciplinary climate was represented as a latent variable at both the student and school level measured by three observed indicators; science achievement and the SES measures were represented by one observed variable each. All models were estimated by means of robust maximum likelihood estimation, and possible missing values were handled through the full-information maximum likelihood procedure. Moreover, the student and school weights were employed to adjust for possible selection bias and differences in the sampling probabilities. Student weights were scaled to the cluster and school weights to the sample. For the models involving the science achievement scores (i.e., the set of 10 plausible values), the analyses were performed for each plausible analysis separately and combined the resultant model parameters using
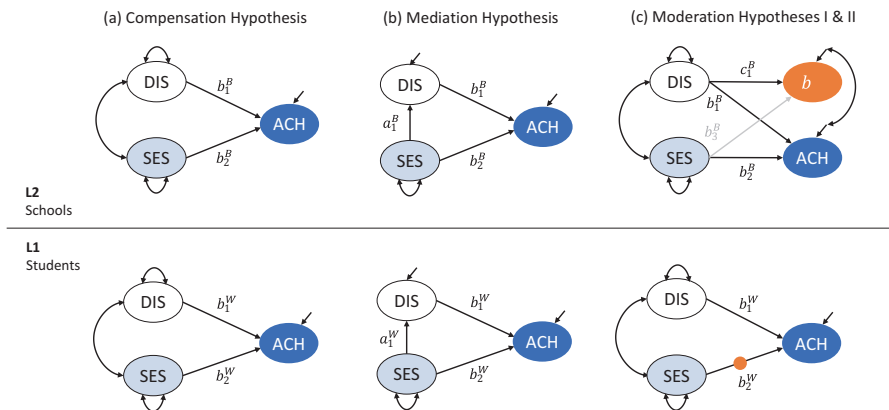
Fig. 8.2 Representation of the three hypotheses as multilevel structural equation models. *Note. ACH* Science achievement, *DIS* Disciplinary climate perceptions, *SES* Socioeconomic status, *B* Between, *W* Within. Random slopes are indicated in orange. The path coefficient b_3^B is only estimated for the testing of the moderation hypothesis II, yet not moderation hypothesis I

Rubin's combination rules. The M*plus* software facilitates this procedure via the TYPE = IMPUTATION option.

For the cross-level interaction models (moderation hypotheses I and II), the information criteria (AIC and BIC) were used to compare competing models; models exhibiting lower AIC and BIC values were preferred. To back these comparisons, likelihood-ratio tests were performed to examine the differences between different cross-level interaction models. For the contextual and mediation models, model fit was evaluated with the help of several fit indices, including the Satorra-Bentler corrected chi-square statistic (SB-χ^2), the comparative fit index (CFI), the root mean square error of approximation (RMSEA), and the level-specific standardized root mean square residual (SRMR$_W$, SRMR$_B$), and the partial saturation approach was performed to identify possible sources of misfit (Ryu, 2014). The common guidelines for evaluating the goodness-of-fit (CFI ≥ 0.95, RMSEA ≤0.06, and SRMR ≤0.08) served as additional sources of information (Kline, 2015). All models were estimated as single-group two-level models first and multigroup two-level models second; the latter allowed for the country-specific reporting of the relevant model parameters.

8.2.2.2 Evaluating the Disciplinary Climate Measurement Model

Students' perceptions of the disciplinary climate were represented as a latent variable at both the student and the school levels. To ensure the cross-level measurement invariance of these two latent variables and establish the same meaning of the respective constructs, factor loadings were constrained to being equal across levels (Stapleton, Yang, & Hancock, 2016). To support this constraint, multilevel confirmatory factor analysis models with and without these equality constraints were compared using fit indices and chi-square difference testing. After establishing that cross-level invariance held for data of each of the five Nordic countries, multilevel CFA models were extended to multigroup multilevel CFA models and tested for cross-country measurement invariance of the latent variables at both the student and school levels. This testing procedure was needed to establish that a sufficient degree of comparability across countries and levels was given to meaningfully compare the relations among variables. All model comparisons were based on the differences in CFI, RMSEA, SRMR-within, SRMR-between, and chi-square difference testing following the commonly applied guidelines for invariance testing (i.e., ΔCFI ≤ −0.010, ΔRMSEA ≤0.015, ΔSRMR ≤0.030; (Chen, 2007).

For the construct of the disciplinary climate, the results provided evidence that cross-country metric invariance at both the student and school levels and cross-level metric invariance held; the changes in the model fit statistics after adding invariance constraints did not deteriorate the model fit substantially. The final multigroup multilevel CFA model imposing these invariance constraints showed a very good fit to the data, SB- χ^2 (18) = 51.6, $p < .001$, CFI = 0.998, RMSEA = 0.019, SRMR$_W$ = 0.014, SRMR$_B$ = 0.014. Furthermore, the factor loadings of all three items were high across countries ($\lambda_W = 0.70$–0.85, $\lambda_B = 0.91$–1.00).

8.2.2.3 Evaluating the Structural Models

After examining the measurement models of disciplinary climate, the structural models were estimated. The models testing the compensation hypothesis were *contextual models* with latent-variable centered predictors of science achievement (*ACH*), and the contextual effect (cont_{DIS}) was represented as the difference between the between-level (b_1^B) and within-level (b_1^W) direct effects of disciplinary climate (*DIS*), $\mathrm{cont}_{DIS} = b_1^B - b_1^W$. The standardized contextual effect with the corresponding effect size ES_2 were obtained (Marsh et al., 2009; see Supplementary Material S1). To test the mediation hypothesis, *multilevel mediation models* with indirect effects of the SES measures on science achievement via disciplinary climate at both levels were estimated. Given that all these variables were measured at the student level and aggregated to the school level, these mediation models can be classified as 1–1-1 multilevel mediation models (Preacher et al., 2010), with a contextual indirect effect represented as the difference between the between-level (ind_B) and within-level (ind_W) indirect effects, $\mathrm{cont}_{ind} = \mathrm{ind}_B - \mathrm{ind}_W$ (Nagengast & Marsh, 2012). The standardized squared indirect effect served as the corresponding effect size (Lachowicz, Preacher, & Kelley, 2018). Finally, the moderation hypotheses were tested with the help of *cross-level interaction models* (Aguinis, Gottfredson, & Culpepper, 2013).

8.3 Results

8.3.1 Compensation Hypothesis (RQ 1)

As noted above, the compensation hypothesis accounted for the level and SES measure specificity in the PISA 2015 data. Along these lines, the subsequent reporting contains the corresponding regression coefficients for the student and the school level and each of the three SES measures in Table 8.2. The regression coefficients describe the relation between disciplinary climate and science achievement after controlling for SES at the student level (b_1^W) and the school level (b_1^B). Next to the variance explanations, they served as the criteria used to determine whether the compensation hypothesis could be supported (see Fig. 8.2). A representation of the results is provided in Fig. 8.3, and a more detailed description including the model fit indices is given in Supplementary Material S2.

8.3.1.1 Compensation Hypothesis at the Student Level

Consistent across countries and SES measures, students' perceptions of disciplinary climate predicted their science achievement above and beyond SES, with standardized regression coefficients ranging between $b_1^W = 0.037$ and $b_1^W = 0.102$ and overall variance explanations between $R_W^2 = 1.5\%$ and $R_W^2 = 6.3\%$. These

Table 8.2 Standardized coefficients of the student- and school-level regression models (Compensation hypothesis; see Fig. 8.2)

	L1: Student Level		L2: School Level		Contextual effect	
Country	b_1^W (SE)	b_2^W (SE)	b_1^B (SE)	b_2^B (SE)	stdcont$_{DIS}$(SE)	ES$_2$(SE)
SES measure: HISEI						
Denmark	0.077 (0.022)*	0.224 (0.016)*	0.188 (0.088)*	0.684 (0.072)*	0.053 (0.041)	0.111 (0.086)
Finland	0.097 (0.021)*	0.230 (0.018)*	0.139 (0.169)	0.542 (0.096)*	0.021 (0.072)	0.044 (0.153)
Iceland	0.063 (0.024)*	0.153 (0.021)*	0.278 (0.150)#	0.463 (0.157)*	0.029 (0.034)	0.058 (0.070)
Norway	0.089 (0.022)*	0.222 (0.017)*	0.507 (0.131)*	0.484 (0.121)*	0.110 (0.046)*	0.228 (0.097)*
Sweden	0.037 (0.022)#	0.233 (0.019)*	0.266 (0.083)*	0.759 (0.059)*	0.121 (0.044)*	0.249 (0.092)*
Overall	0.070 (0.011)*	0.228 (0.009)*	0.492 (0.052)*	0.216 (0.060)*	0.065 (0.028)*	0.140 (0.061)*
SES measure: HOMEPOS						
Denmark	0.081 (0.022)*	0.213 (0.017)*	0.159 (0.091)	0.765 (0.051)*	0.039 (0.044)	0.081 (0.092)
Finland	0.102 (0.021)*	0.174 (0.019)*	0.119 (0.106)	0.809 (0.102)*	0.014 (0.046)	0.029 (0.094)
Iceland	0.062 (0.024)*	0.106 (0.022)*	0.340 (0.158)*	0.410 (0.188)*	0.042 (0.036)	0.086 (0.074)
Norway	0.081 (0.023)*	0.217 (0.019)*	0.398 (0.119)*	0.487 (0.116)*	0.081 (0.042)#	0.169 (0.088)#
Sweden	0.041 (0.023)#	0.174 (0.019)*	0.543 (0.102)*	0.431 (0.093)*	0.246 (0.060)*	0.532 (0.141)*
Overall	0.073 (0.011)*	0.188 (0.010)*	0.320 (0.068)*	0.284 (0.063)*	0.108 (0.033)*	0.236 (0.073)*
SES measure: PARED						
Denmark	0.082 (0.022)*	0.121 (0.018)*	0.213 (0.108)*	0.648 (0.107)*	0.061 (0.051)	0.128 (0.108)
Finland	0.096 (0.021)*	0.139 (0.017)*	0.098 (0.096)	0.837 (0.085)*	0.008 (0.042)	0.016 (0.086)
Iceland	0.063 (0.024)*	0.162 (0.019)*	0.348 (0.137)*	0.430 (0.165)*	0.045 (0.033)	0.091 (0.068)
Norway	0.088 (0.022)*	0.098 (0.020)*	0.504 (0.119)*	0.337 (0.120)*	0.107 (0.044)*	0.224 (0.094)*
Sweden	0.041 (0.023)#	0.168 (0.019)*	0.302 (0.134)*	0.635 (0.120)*	0.134 (0.066)*	0.281 (0.141)*
Overall	0.072 (0.011)*	0.137 (0.009)*	0.388 (0.056)*	0.458 (0.046)*	0.137 (0.029)*	0.297 (0.064)*

Note. *W* Within (student) level, *B* School (between) level, *stdcont$_{DIS}$* standardized contextual effect, *ES$_2$* Effect size of the contextual effect (Marsh et al., 2009). * $p < .05$, # $p < .10$
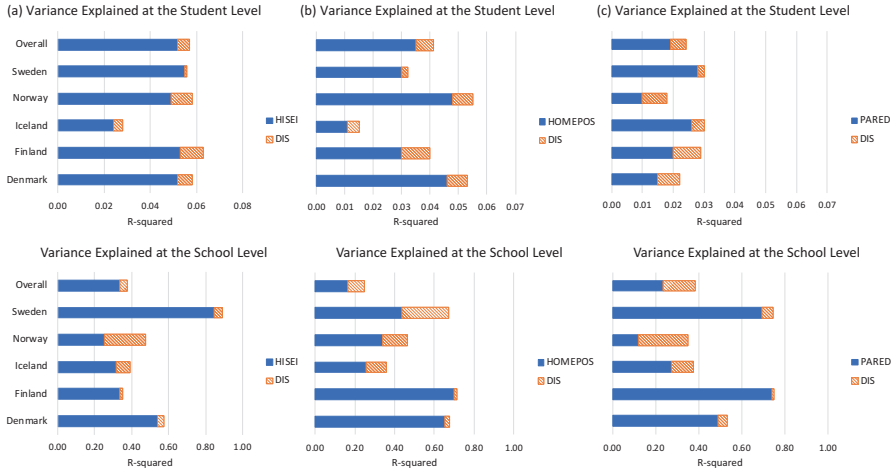
Fig. 8.3 Variance explanations of science achievement at the student and the school level by SES and disciplinary climate (DIS).
Note. The variance explanation of DIS is based on models in which both SES and DIS were included—these values indicate the additional contribution of DIS to the variance explanation by the SES measure

coefficients varied slightly between countries, and the Swedish data exhibited the smallest compensation effects. SES was a consistently strong predictor of individual science achievement, and the measure of disciplinary climate perceptions added only between $R^2_W = 0.1\%$ and 1.0% to the variance explanation by SES (see Fig. 8.3).

8.3.1.2 Compensation Hypothesis at the School Level

The school-level regression coefficients of the disciplinary climate measure ranged between $b^B_1 = 0.098$ and $b^B_1 = 0.543$ across countries and SES measures. Notably, the Icelandic, Norwegian, and Swedish data showed the largest effects across all SES measures and supported the compensation hypothesis at the school level. Except for the HOMEPOS measure, the Danish data also provided evidence backing the compensation hypothesis; however, there was no support for the Finnish data. The overall variance explanations at the school level ranged between $R^2_B = 35.2\%$ and $R^2_B = 74.8\%$. As for the student-level data, SES was a consistently strong predictor of school science achievement—the measure of disciplinary climate added between $R^2_B = 0.9\%$ and $R^2_B = 23.1\%$ to this variance explanation. The largest added values occurred for the Norwegian data (SES measures HISEI and PARED) and the Swedish data (SES measure HOMEPOS; see Fig. 8.3).

8.3.1.3 Contextual Direct Effects

The contextual effects—that is, the effects of school-level disciplinary climate on individual science achievement after controlling for school SES, individual SES, and perceptions of disciplinary climate—were statistically significant only for the Norwegian and the Swedish data. These effects were positive and ranged between $ES_2 = 0.14$ and $ES_2 = 0.53$. Notably, these effect sizes varied between the SES measures. Specifically, while they were of similar size for the SES measures HISEI and PARED for both Norway and Sweden, they differed to a larger extent between the countries for the HOMEPOS measure, with a larger effect for the Swedish data. Moreover, the effect was the largest among all effects for Sweden.

8.3.2 Mediation Hypothesis (RQ 2)

To test the mediation hypothesis, the indirect effects, along with the squared standardized indirect effects as effect sizes for both the student and the school level, were examined. Figure 8.4 shows the resultant direct and indirect effects for all SES measures, countries, and levels, and the Supplementary Material S2 contains all relevant model parameters.
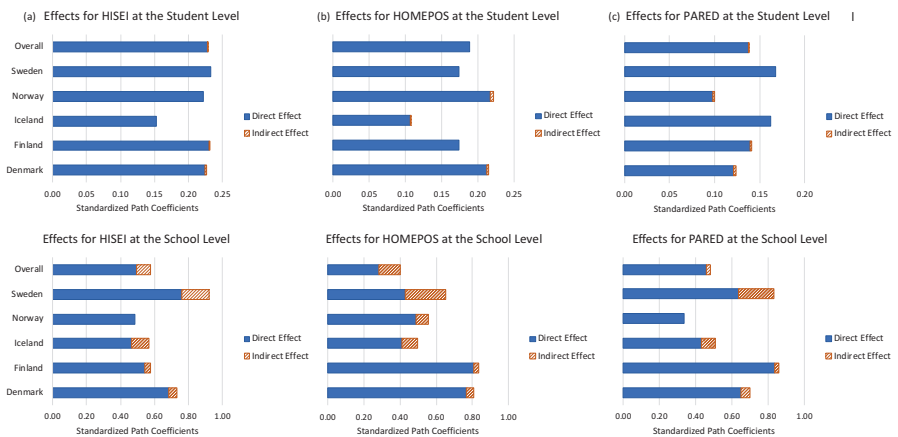
Fig. 8.4 Direct, indirect, and total effects of the SES measures on science achievement via disciplinary climate.
Note. Standardized path coefficients are shown

8.3.2.1 Mediation Hypothesis at the Student Level

Across all analytic conditions, there was no evidence supporting that the indirect within-level effects were different from zero. All effects were small, and the corresponding effect sizes were zero. Overall, the mediation hypothesis could not be supported for the student-level data.

8.3.2.2 Mediation Hypothesis at the School Level

In contrast to the student level, the mediation models at the school level exhibited significant and positive indirect effects for the Swedish data across all SES measures ($ind_B = 0.160$–0.220), with the highest value for the HOMEPOS measure. The corresponding effect sizes ranged between $\nu = 0.026$ and $\nu = 0.048$, and these can be considered small (Lachowicz et al., 2018).

8.3.2.3 Contextual Indirect Effects

Only in the case of the Swedish data did a positive and statistically significant difference between the school- and the student-level indirect effects occur across all SES measures. Nevertheless, this contextual effect surfaced because the indirect effect did not exist in the student-level model, whereas it was present in the school-level model.

8.3.3   Moderation Hypotheses (RQ 3)

Concerning the first moderation hypothesis (see Fig. 8.2c), there was evidence for a positive moderation of the relation between SES and science achievement only for the Swedish data and only for the SES measures HISEI, $c_1^B = 1.031$, $SE = 0.453$, $p = .023$, and HOMEPOS, $c_1^B = 31.663$, $SE = 10.403$, $p = .002$. These moderation effects suggested an increase in the SES–achievement relation with a better disciplinary climate. However, given the large standard errors, these effects must be interpreted with caution. No further cross-level interaction effects in the other countries and across the other analytic conditions could be found.

Concerning the second moderation hypothesis (see Fig. 8.2c), there was no evidence for the role of disciplinary climate in science lessons as a moderator of the SES-achievement relation. After introducing school SES as a possible moderator, the moderating effects of disciplinary climate for the Swedish data disappeared (see Supplementary Material S2). In fact, there was evidence for a significant cross-level interaction effect of school SES under the following conditions: (a) HISEI: Iceland, Norway, and Sweden, $b_3^B = 0.039$–0.056, $ps < .05$; (b) HOMEPOS: Finland, $b_3^B = -64.106$, $SE = 12.322$, $p = .004$; and (c) PARED: Denmark, $b_3^B = 1.242$, $SE = 0.566$,

$p = .028$, and Iceland, $b_3^B = 2.985$, $SE = 1.329$, $p = .025$. While the SES-achievement relation was stronger for higher values of HISEI or PARED in countries with significantly positive moderation effects, the relation was smaller for higher values of HOMEPOS in the Swedish data. Once again, the latter effect must be interpreted with caution due to the large standard error. Nevertheless, the moderation by disciplinary climate was not supported, and the moderation by SES differed across countries and SES measures.

8.3.4   Summary of the Main Findings

Table 8.3 visualizes the main findings; overall, the testing of the three hypotheses revealed the following results:

- *Compensation hypothesis:* Consistent evidence for the relation between disciplinary climate perceptions and science achievement after controlling for SES at

Table 8.3 Summary of the main findings

	Compensation hypothesis			Mediation hypothesis			Moderation hypothesis
Country	L1 (students)	L2 (schools)	Contextual effect	L1 (students)	L2 (schools)	Contextual effect	Cross-level interaction effect
SES measure: HISEI							
Denmark	Yes	Yes	No	No	No	No	No
Finland	Yes	No	No	No	No	No	No
Iceland	Yes	Yes	No	No	No	No	No
Norway	Yes	Yes	Yes	No	No	No	No
Sweden	Yes	Yes	Yes	No	Yes	Yes	Yes
Overall	Yes	Yes	Yes	No	Yes	Yes	No
SES measure: HOMEPOS							
Denmark	Yes	No	No	No	No	No	No
Finland	Yes	No	No	No	No	No	No
Iceland	Yes	Yes	No	No	No	No	No
Norway	Yes	Yes	Yes	No	No	No	No
Sweden	Yes	Yes	Yes	No	Yes	Yes	Yes
Overall	Yes	Yes	Yes	No	Yes	Yes	No
SES measure: PARED							
Denmark	Yes	Yes	No	No	No	No	No
Finland	Yes	No	No	No	No	No	No
Iceland	Yes	Yes	No	No	No	No	No
Norway	Yes	Yes	Yes	No	No	No	No
Sweden	Yes	Yes	Yes	No	Yes	Yes	No
Overall	Yes	Yes	Yes	No	No	No	No

the student level across all countries and measures of SES was found. At the same time, these relations varied between countries, with Finland and Norway showing the largest and Sweden the smallest effects. The variance explanations over and above SES were consistently small. Moreover, consistent evidence was found supporting the compensation hypothesis for the school level, except for the Finnish sample across all SES measures and the Danish sample for the home possession measure. The effects varied across SES measures even within countries; the Norwegian and Swedish data indicated consistently strong relations between disciplinary climate and science achievement and substantial variance explanations over and above SES at the school level. Contextual effects—that is, the effects of school-level disciplinary climate on individual science achievement across countries—existed only for the Norwegian and Swedish sample, with larger effect sizes for the latter.

- *Mediation hypothesis:* There was no evidence supporting the mediation hypothesis for the student level; only the Swedish data supported the existence of an indirect effect at the school level, which was consistent across all SES measures.
- *Moderation hypothesis:* The moderation hypothesis was supported only for the Swedish data and the SES measures representing HOMEPOS and HISEI status.

8.4   Discussion

8.4.1   *The Three Hypotheses in the Context of Equity and Equality*

As educational inequalities exist in academic achievement due to differences in students' SES, and ultimately, the classroom and school SES composition, identifying possible classroom and school factors that may compensate, mediate, or moderate these inequalities is key to educational research and policy making (Cresswell, Schwantner, & Waters, 2015). In this sense, the three hypotheses proposed by Berkowitz et al. (2017) provide different lenses through which the role of such factors can be investigated. Using this framework, this study focused on disciplinary climate in science lessons as a school factor and obtained evidence for or against the three hypotheses.

Specifically, in all three hypotheses, a link between SES and science achievement was assumed, which represented inequalities in educational outcomes (Willms & Tramonte, 2019). This link existed across the five Nordic countries and across the two levels of analysis, and indeed, indicated the presence of outcome inequalities due to differences in SES between students within schools and between schools. The consistent and moderate association between SES and achievement is well in line with the existing body of research and testifies to the strong explanatory power of SES (e.g., Kim et al., 2019; Sirin, 2005; Thomson, 2018). While striving for reducing the SES–achievement relation is a key goal for educational effectiveness

and school improvement (Scherer & Nilsen, 2019), explaining the possible mechanisms through which it operates is almost equally important (Berkowitz et al., 2017). In fact, knowledge about these mechanisms can provide insights into the roles of classroom or school factors from different perspectives—the mechanisms examined through the three hypotheses in this chapter were based on different assumptions about the role of the disciplinary climate, and ultimately, provided different interpretations.

The evidence supporting the *compensation hypothesis* suggests that a good disciplinary climate is indeed related to better science achievement after controlling for SES. In other words, disciplinary climate may compensate for educational inequalities due to SES. Notably, this finding was consistent across the five Nordic countries for both students and schools. At the individual (student) level, the compensation mechanism indicates that more positive perceptions of disciplinary climate in science lessons are associated with better science achievement after controlling for possible SES differences between students within a school. At the school level, the same interpretation holds for shared perceptions of disciplinary climate, school-average SES, and science achievement (Ning et al., 2015). One may argue that schools in the sample of Nordic countries succeed in achieving high due to establishing a good disciplinary climate in lessons, independent of their SES composition (e.g., Bellens et al., 2019).

The limited evidence for the *mediation* and *moderation hypotheses* for the PISA 2015 Nordic data may have several explanations and interpretations, which are as follows:

- Mediation only occurred for some countries at the school level but not the student level. Consequently, for the present data, this assumption represents school-level mechanisms implying that a good disciplinary climate is more likely to be found in high-SES schools, and ultimately, contributes to better school achievement (Liu et al., 2015). However, the student-level assumption that high-SES students are more likely to perceive the disciplinary climate more positively, and in turn, achieve better, still needs to be backed conceptually and empirically.
- For the Norwegian data, mediation could not be detected due to the missing link between-school SES and disciplinary climate across all three SES measures. The Finnish and Icelandic data showed the same pattern for the HOMEPOS and PARED measures. Consistently across all SES measures, the Finnish data further showed insignificant relations between disciplinary climate and achievement. The missing SES–climate link may be interpreted as an indicator of equal opportunities for students in schools to experience a good disciplinary climate (Willms & Tramonte, 2019). The missing climate–achievement link may be interpreted as a lack of school effects that could contribute to better achievement—for the Finnish PISA 2015 data, disciplinary climate was not a predictor of science achievement at the school level after controlling for both variables for school SES.
- Moderation by disciplinary climate is based on the assumption that a positive disciplinary climate may be associated with smaller achievement gaps (Ning et al., 2015). As there was no evidence supporting this assumption, the hopes

associated with disciplinary climate as a possible factor reducing inequalities in educational outcomes could not be fulfilled for the present data. However, this observation is in line with previous studies that could not identify moderation effects (e.g., Bellens et al., 2019). Notably, tracing such effects with sufficient power depends on several factors, including the complexity of the cross-level interaction models and the decomposition of the moderator variable into its within and between parts (Aguinis et al., 2013). Possible methodological issues may prohibit the substantive interpretation of the effects.

Concerning whether a uniform "Nordic model" regarding the three hypotheses exists, the findings indicated cross-country differences not only in the sizes of the relations among SES, disciplinary climate, and science achievement but also in the conclusions following them. These differences emerged for the compensation hypothesis in the Danish and Finnish data and for the mediation and moderation hypotheses for the Norwegian and Swedish data, yet without consistent effects across SES measures. This observation brings forward the question of possible explanations for these differences. Although desirable, the present data do not provide opportunities to explore direct causal explanations, and any explanation at the level of educational systems (e.g., considering educational reforms and policy making) would need to be substantiated by external data sources and (quasi-)experimental research designs (Rutkowski & Delandshere, 2016). In this sense, researchers are encouraged to explore and investigate possible explanatory variables for the differences identified in the study; such variables could offer further insights into what may characterize a "Nordic model."

8.4.2 Limitations and Future Directions

The secondary data analyses and possible inferences drawn from their results have at least two limitations worth noting: First, the disciplinary climate was assessed by student ratings as part of the PISA 2015 background questionnaire, and the corresponding items referred to the "disciplinary climate in school science lessons." This reference to the school level rather than the classroom level hinders classroom-level inferences (Scherer, Nilsen, & Jansen, 2016). Instead, given the level of analysis, the interpretation of the construct is more in line with that of school climate rather than instructional quality.

Second, some methodological approaches taken in the secondary data analyses have not yet been fully developed. For instance, little is known about the importance of cross-level measurement invariance in cross-level interaction models with moderating school-level variables that are aggregated student-level variables (Jak, 2019), especially when detecting the cross-level interaction effect. Moreover, some relations in the analytic models may be curvilinear rather than linear (Teig, Scherer, & Nilsen, 2018). In this sense, methodological research on these issues will help

readers fully understand the models that are used to describe the relations among the three constructs.

8.5   Conclusions and Implications

The secondary data analyses of the PISA 2015 data from five Nordic countries resulted in consistent and robust evidence supporting the compensation hypothesis, that is, the disciplinary climate's contribution to science achievement above and beyond SES at both the student and school levels. At the same time, only limited evidence supporting the mediation hypothesis—with some exceptions for school-level data—and the moderation hypothesis surfaced. These observations point to the following conclusions: (a) although educational inequalities may exist, a good disciplinary climate is associated with better science achievement; and (b) inequalities in the opportunities to experience a good disciplinary climate (due to differences in SES) may not translate into inequalities in science achievement. Considering these conclusions, this study has several implications: From a substantive perspective, the three hypotheses may indeed represent educationally relevant lenses through which the role of disciplinary climate for SES, academic achievement, and the SES–achievement relation could be examined. This chapter has shown that these hypotheses can be converted into testable statistical models. From a methodological perspective, any study investigating the interplay between disciplinary climate, SES, and achievement should consider several levels of analysis and examine the meaning of the construct at these levels (e.g., student perceptions vs. shared perceptions of students within a school). In addition, the study highlighted the importance of measurement invariance to facilitate similar construct meaning across countries and levels.

This chapter further reveals some implications for the understanding of equity and equality in school contexts: First, the hope that disciplinary climate—a core school condition and indicator of instructional quality—can compensate efficiently and directly for possible achievement gaps in the domain of science could not be substantiated with the present data and selection of countries. This calls into question possible compensatory mechanisms and effects of the disciplinary climate as a malleable contextual variable. Second, the mechanisms describing the role of school conditions for addressing possible achievement gaps are far from clear cut; in fact, the PISA 2015 data did not provide clear support for any of them. This implies that the researchers' theoretical perspectives on equity and equality will mainly determine the evaluation of the specific mechanism. Third, the three mechanisms tested in the secondary analyses shed light on different aspects of equity and equality; while the moderation hypothesis is based on the suggestion that equality in education can be increased by school conditions, the mediation hypothesis considers the dependencies between equality and equity via school conditions.

Concerning the elements describing a Nordic model, the study revealed some homogeneity in the findings across these countries—and some heterogeneity as

well. Consistently, a compensation mechanism describing the interplay between equality and school conditions arose at different levels of analysis; however, the other mechanisms could hardly be traced. In this sense, achievement differences in science can partly be compensated for by a positive disciplinary school climate—the school condition studied in this chapter. Therefore, it seems that this compensation mechanism represents an element of the Nordic model. However, these findings do not imply a possible reduction of achievement gaps in science through a better disciplinary climate.

References

Aguinis, H., Gottfredson, R. K., & Culpepper, S. A. (2013). Best-practice recommendations for estimating cross-level interaction effects using multilevel modeling. *Journal of Management, 39*(6), 1490–1528. https://doi.org/10.1177/0149206313478188

Antikainen, A. (2006). In search of the Nordic model in education. *Scandinavian Journal of Educational Research, 50*(3), 229–243. https://doi.org/10.1080/00313830600743258

APA. (2006). *Report of the APA task force on socioeconomic status*. Washington, DC: APA.

Archer, L., Dawson, E., DeWitt, J., Seakins, A., & Wong, B. (2015). "Science capital": A conceptual, methodological, and empirical argument for extending bourdieusian notions of capital beyond the arts. *Journal of Research in Science Teaching, 52*(7), 922–948. https://doi.org/10.1002/tea.21227

Archer, L., DeWitt, J., Osborne, J., Dillon, J., Willis, B., & Wong, B. (2012). Science aspirations, capital, and family habitus: How families shape Children's engagement and identification with science. *American Educational Research Journal, 49*(5), 881–908. https://doi.org/10.3102/0002831211433290

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*. https://doi.org/10.1016/j.learninstruc.2019.05.008

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 Assessment in Education, 7*, 1. https://doi.org/10.1186/s40536-019-0069-2

Berkowitz, R., Moore, H., Astor, R. A., & Benbenishty, R. (2017). A research synthesis of the associations between socioeconomic background, inequality, school climate, and academic achievement. *Review of Educational Research, 87*(2), 425–469. https://doi.org/10.3102/0034654316669821

Bianchini, J. A. (2017). Equity in science education. In K. S. Taber & B. Akpan (Eds.), *Science education: An international course companion* (pp. 455–464). Rotterdam, The Netherlands: Sense Publishers.

Blank, C., & Shavit, Y. (2016). The association between student reports of classmates' disruptive behavior and student achievement. *AERA Open, 2*(3). https://doi.org/10.1177/2332858416653921

Blossing, U., Imsen, G., & Moos, L. (2014). Schools for all: A Nordic model. In U. Blossing, G. Imsen, & L. Moos (Eds.), *The Nordic education model: 'A School for All' encounters Neo-Liberal policy* (pp. 231–239). Dordrecht, The Netherlands: Springer Netherlands.

Broer, M., Bai, Y., & Fonseca, F. (2019). *Socioeconomic inequality and educational outcomes*. Cham, Switzerland: Springer.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Chen, G., & Weikart, L. A. (2008). Student background, school climate, school disorder, and student achievement: An empirical study of new York City's middle schools. *Journal of School Violence, 7*(4), 3–20. https://doi.org/10.1080/15388220801973813

Cresswell, J., Schwantner, U., & Waters, C. (2015). *A review of international large-scale assessments in education: Assessing component skills and collecting contextual data*. Paris, France: OECD Publishing.

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

Espinoza, O. (2007). Solving the equity–equality conceptual dilemma: A new model for analysis of the educational process. *Educational Research, 49*(4), 343–363. https://doi.org/10.1080/00131880701717198

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

Gustafsson, J.-E., Nilsen, T., & Hansen, K. Y. (2018). School characteristics moderating the relation between student socio-economic status and mathematics achievement in grade 8. Evidence from 50 countries in TIMSS 2011. *Studies in Educational Evaluation, 57*, 16–30. https://doi.org/10.1016/j.stueduc.2016.09.004

Harwell, M., Maeda, Y., Bishop, K., & Xie, A. (2016). The surprisingly modest relationship between SES and educational achievement. *The Journal of Experimental Education, 85*(2), 197–214. https://doi.org/10.1080/00220973.2015.1123668

Hochweber, J., Hosenfeld, I., & Klieme, E. (2014). Classroom composition, classroom management, and the relationship between student attributes and grades. *Journal of Educational Psychology, 106*(1), 289–300. https://doi.org/10.1037/a0033829

Jak, S. (2019). Cross-level invariance in multilevel factor models. *Structural Equation Modeling: A Multidisciplinary Journal, 26*(4), 607–622. https://doi.org/10.1080/10705511.2018.1534205

Jehangir, K., Glas, C. A. W., & van den Berg, S. (2015). Exploring the relation between socio-economic status and reading achievement in PISA 2009 through an intercepts-and-slopes-as-outcomes paradigm. *International Journal of Educational Research, 71*, 1–15. https://doi.org/10.1016/j.ijer.2015.02.002

Kavli, A.-B. (2018). TIMSS and PISA in the Nordic countries. In N. C. O. Ministers (Ed.), *Northern lights on TIMSS and PISA* (pp. 11–30). Copenhagen, Denmark: Nordic Council of Ministers.

Kim, S. w., Cho, H., & Kim, L. Y. (2019). Socioeconomic status and academic outcomes in developing countries: A meta-analysis. *Review of Educational Research, 89*(6), 875–916. https://doi.org/10.3102/0034654319877155

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

Kline, R. B. (2015). *Principles and practice of structural equation modeling* (4th ed.). New York: Guilford Press.

Lachowicz, M. J., Preacher, K. J., & Kelley, K. (2018). A novel measure of effect size for mediation analysis. *Psychological Methods, 23*(2), 244–261. https://doi.org/10.1037/met0000165

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

Liu, H., Van Damme, J., Gielen, S., & Van Den Noortgate, W. (2015). School processes mediate school compositional effects: Model specification and estimation. *British Educational Research Journal, 41*(3), 423–447. https://doi.org/10.1002/berj.3147

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

Marsh, H. W., Lüdtke, O., Nagengast, B., Trautwein, U., Morin, A. J. S., Abduljabbar, A. S., et al. (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., Lüdtke, O., Robitzsch, A., Trautwein, U., Asparouhov, T., Muthén, B., et al. (2009). Doubly-latent models of school contextual effects: Integrating multilevel and structural equation approaches to control measurement and sampling error. *Multivariate Behavioral Research, 44*(6), 764–802. https://doi.org/10.1080/00273170903333665

Muthén, L. K., & Muthén, B. O. (1998-2017). *Mplus user's guide* (8th ed.). Los Angeles, CA: Muthén & Muthén.

Nagengast, B., & Marsh, H. W. (2012). Big fish in little ponds aspire more: Mediation and cross-cultural generalizability of school-average ability effects on self-concept and career aspirations in science. *Journal of Educational Psychology, 104*(4), 1033–1053. https://doi.org/10.1037/a0027697

Nilsen, T., Bloemeke, S., Yang Hansen, K., & Gustafsson, J.-E. (2016). Are school characteristics related to equity? The answer may depend on a country's developmental level. *IEA Policy Briefs, 10*, 1–7.

Ning, B., Van Damme, J., Van Den Noortgate, W., Yang, X., & Gielen, S. (2015). The influence of classroom disciplinary climate of schools on reading achievement: A cross-country comparative study. *School Effectiveness and School Improvement, 26*(4), 586–611. https://doi.org/10.1080/09243453.2015.1025796

OECD. (2016). *PISA 2015 results (volume I): Excellence and equity in education*. Paris, France: OECD Publishing.

OECD. (2017a). *PISA 2015 assessment and analytical framework: Science, reading, mathematic, financial literacy and collaborative problem solving*. Paris, France: OECD Publishing.

OECD. (2017b). *PISA 2015 technical report*. Paris, France: OECD Publishing.

OECD. (2019). *PISA 2018 results (Volume I)*. Paris, France: OECD Publishing.

Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods, 15*(3), 209–233. https://doi.org/10.1037/a0020141

Rhemtulla, M., van Bork, R., & Borsboom, D. (2019). Worse than measurement error: Consequences of inappropriate latent variable measurement models. *Psychological Methods*, No Pagination Specified-No Pagination Specified. https://doi.org/10.1037/met0000220

Rjosk, C., Richter, D., Hochweber, J., Lüdtke, O., Klieme, E., & Stanat, P. (2014). Socioeconomic and language minority classroom composition and individual reading achievement: The mediating role of instructional quality. *Learning and Instruction, 32*, 63–72. https://doi.org/10.1016/j.learninstruc.2014.01.007

Rutkowski, D., & Delandshere, G. (2016). Causal inferences with large scale assessment data: Using a validity framework. *Large-Scale Assessments in Education, 4*(1), 6. https://doi.org/10.1186/s40536-016-0019-1

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

Scherer, R., & Nilsen, T. (2017). *Relations between school climate, instructional quality, and educational outcomes*. Paper presented at the 7th IEA International Research Conference (IRC), Praque, Czech Republic.

Scherer, R., & Nilsen, T. (2019). Closing the gaps? Differential effectiveness and accountability as a road to school improvement. *School Effectiveness and School Improvement, 30*(3), 255–260. https://doi.org/10.1080/09243453.2019.1623450

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*(110). https://doi.org/10.3389/fpsyg.2016.00110

Scherer, R., & Siddiq, F. (2019). The relation between students' socioeconomic status and ICT literacy: Findings from a meta-analysis. *Computers & Education, 138*, 13–32. https://doi.org/10.1016/j.compedu.2019.04.011

Schmidt, W. H., Burroughs, N. A., Zoido, P., & Houang, R. T. (2015). The role of schooling in perpetuating educational inequality. *Educational Researcher, 44*(7), 371–386. https://doi.org/10.3102/0013189x15603982

Seidel, T., & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research Design in Disentangling Meta-Analysis Results. *Review of Educational Research, 77*(4), 454–499. https://doi.org/10.3102/0034654307310317

Shin, J., Lee, H., & Kim, Y. (2009). Student and school factors affecting mathematics achievement. *School Psychology International, 30*(5), 520–537. https://doi.org/10.1177/0143034309107070

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

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). London: SAGE Publications.

Sortkær, B., & Reimer, D. (2018). Classroom disciplinary climate of schools and gender – Evidence from the Nordic countries. *School Effectiveness and School Improvement, 29*(4), 511–528. https://doi.org/10.1080/09243453.2018.1460382

Stapleton, L. M., Yang, J. S., & Hancock, G. R. (2016). Construct meaning in multilevel settings. *Journal of Educational and Behavioral Statistics, 41*(5), 481–520. https://doi.org/10.3102/1076998616646200

Teig, N., Scherer, R., & Nilsen, T. (2018). More isn't always better: The curvilinear relationship between inquiry-based teaching and student achievement in science. *Learning and Instruction, 56*, 20–29. https://doi.org/10.1016/j.learninstruc.2018.02.006

Thomson, S. (2018). Achievement at school and socioeconomic background-an educational perspective. *NPJ Science of Learning, 3*, 5. https://doi.org/10.1038/s41539-018-0022-0

van Ewijk, R., & Sleegers, P. (2010). The effect of peer socioeconomic status on student achievement: A meta-analysis. *Educational Research Review, 5*(2), 134–150. https://doi.org/10.1016/j.edurev.2010.02.001

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/10.3102/2F00346543063003249

Wang, M.-T., & Degol, J. L. (2015). School climate: A review of the construct, measurement, and impact on student outcomes. *Educational Psychology Review, 28*(2), 315–352. https://doi.org/10.1007/s10648-015-9319-1

White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin, 91*(3), 461–481. https://doi.org/10.1037/0033-2909.91.3.461

Willms, J., & Tramonte, L. (2019). The measurement and use of socioeconomic status in educational research. In L. Suter, E. Smith, & B. Denman (Eds.), *The SAGE handbook of comparative studies in education* (pp. 289–304). London: SAGE Publications Ltd.

Chapter 9
Improving Equity Through National-Level Assessment Initiatives

G. A. Nortvedt, K. B. Bratting, O. Kovpanets, A. Pettersen, and A. Rohatgi

Abstract This chapter investigates how a national-level assessment initiative may improve equity in early years numeracy education, taking the Norwegian mapping tests for primary grades 1–3 as an example. Three assessments, one test for each grade level, were launched in the 2013–2014 school year and have been used every year since. In accordance with Nordic model principles, the test content is available to teachers to ensure familiarity with the test content and the formative use of the assessment outcomes to improve teaching and learning for students identified as at risk of lagging behind. Analysis of student data reveals that, 6 years after the first implementation, no inflation can be seen in test scores. Thus, an exposed assessment may remain robust within an educational system that aspires to transparency, such as the Norwegian one. However, analyses of interview data and achievement data reveal that teachers often struggle to use the assessment outcomes to improve teaching. These results suggest that the initiative to improve equity in primary school numeracy education depends on teachers' assessment literacy. In accordance with Nordic model principles, schools have significant autonomy and are responsible for identifying professional development needs for their teachers. This research confirms the dilemmas in the Nordic model between national-level and local initiatives and responsibilities.

Keywords At-risk students · Numeracy · Mapping tests · Assessment for learning · Teachers' assessment literacy · Equity in education

G. A. Nortvedt (✉) · K. B. Bratting · O. Kovpanets · A. Pettersen · A. Rohatgi
Department of Teacher Education and School Research, University of Oslo, Oslo, Norway
e-mail: g.a.nortvedt@ils.uio.no

225

9.1   Introduction

The Nordic Education Model is grounded in a social democratic ideology and an egalitarian philosophy. Its core values are equity and equal opportunities, inclusion and social justice, embedded in national school laws and curriculum documents to define 'A School for All' that ensures all students are given opportunities to reach their maximum potential (Imsen, Blossing, & Moos, 2017; Telhaug, Mediås, & Aasen, 2006). As such, the educational authorities in the Nordic countries implement policies and tools that not only describe educational equity in a Nordic context but also aims to assist schools in striving for equity. In Norway, for instance, national mapping tests in numeracy are available at the primary school level as part of the Norwegian quality assessment system (NQAS). This is not a unique situation; national governments often implement assessment strategies or policies to enhance students' opportunities to learn (Nortvedt & Buchholtz, 2018). The three mapping tests, one for each of the grade levels 1–3, are designed to identify students at risk of lagging behind who would benefit from more targeted teaching. Therefore, the tests are conducted with the aim of offering all students the opportunity to be successful in learning, and as such, improving equity in learning opportunities. Each test is accompanied by support material[1] for the teachers and schools to enhance the schools' efforts as they strive to improve mathematics education for all. The mapping tests differ from many other national-level assessments in some important aspects. For instance, the test data are owned by the local school and not reported in national league tables, and test results should be used formatively (Blömeke & Olsen, 2018).

After a period of test development, piloting and standardisation, the same tests remain in use for at least 5 years consecutively. Moreover, the tests have a high ceiling effect by design, ensuring that targeted students can solve many of the tasks in the test. This means that, unlike typical screening tests, the Norwegian mapping tests provide teachers with information about what identified students know and can do (Nortvedt, 2018). Over time, test content is expected to become highly familiar to teachers. Such transparency connected to national-level initiatives is within the Nordic model principles (Telhaug et al., 2006) and supposed to foster equity. Moreover, transparency enables teachers to further develop their assessment literacy due to opportunities to work with the test content and results.

In this chapter, we relate equity to the policy level and policy-level initiatives. In particular, we address whether national policy initiatives and assessment tools can contribute to equity in schools. As it is the teachers who administer the tests and interpret and use the test outcomes to inform their teaching, their work with the

---

[1] Teacher guides and a national website hosted by the Directorate for Education and Training. While the Grade 2 test is compulsory, grade 1 and 3 tests are voluntary. Still, almost all schools use all three assessments. Taken together, in this chapter we refer to the tests and supplementary material as the assessment

mapping tests is an important part of the national-level initiative. Indeed, trust in teachers to take on this responsibility is embedded in the Norwegian initiative.

We are aware that, with the implementation of such national assessment tools as mapping tests, there is a question about the extent to which they contribute to equity. Both the quality of the assessment and their use may be an issue (Stobart, 2008). Moreover, previous research has shown that teachers' assessment literacy is a critical aspect of their use of assessment data (Popham, 2009). As such, our aim with this chapter is to discuss how, through such national-level initiatives as the Norwegian mapping tests, an education system can enhance equity regarding student learning opportunities. For this purpose, we draw on analysis of student assessment data and teacher interviews.

9.2 Theoretical Framework

This section presents an overview of previous research that serves as a framework for our study. Key aspects of equity, assessment for learning and assessment literacy are discussed before presenting previous research on national-level assessment initiatives in the Norwegian context.

9.2.1 Equity, Equality and Inclusion in Education

The term *equity* is frequently used in both educational research in general and in mathematics education in particular, but often, no clear definition is provided, and the term is used in relation to different issues (Buchholtz et al., 2020; Espinoza, 2007; Roos, 2018). Moreover, equity is often used interchangeably with *equality*, causing confusion and ambiguity in the research literature (e.g. Espinoza, 2007; Zhu, 2018). We follow Rousseau and Tate (2003), who state that equity is associated with fairness or justice in terms of provision of education, while equality is related to sameness, non-discrimination or the state of being equal. Samoff (1996) highlights how equitable education necessitates structural inequalities, for example, to offer adapted education and differentiation.

Some teachers may consider equity in terms of inclusion (Nortvedt & Wiese, 2020). In mathematics education research, the concept of inclusion can refer to both inclusion in society (taking part in the classroom) and inclusion in the form of adapted teaching (Roos, 2018). This is in line with Espinoza's (2007) argument that a set of definitions and conceptualisations should be used that address different dimensions and stages of the educational process rather than striving for a unique understanding of equity and equality. Further, the National Council for Teachers of Mathematics research team argues that equity includes components related to both conditions of learning and outcomes. Their main concern is 'how mathematics education research can contribute to understanding the causes and effects of inequity, as

well as strategies that effectively reduce undesirable inequities of experience and achievement in mathematics education' (Gutstein et al., 2005, p. 94). According to Zhu (2018), individualised approaches are necessary to achieve equity in mathematics education, taking into account differences in students' individual needs and providing differentiated treatments rather than regarding and treating all students equally.

9.2.2   Assessment for Learning

Assessment for learning (AfL) is an important tool to adapt teaching and learning activities to the needs of the individual student. As defined, AfL constitutes 'all those activities undertaken by teachers and/or by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged' (Black & Wiliam, 1998, pp. 7–8). Further, Wiliam (2011) argue that the most important purpose of educational assessment is to serve and support learning. Previous studies have shown that good assessment practices can lead to improved learning (Hattie, 2009; Hattie & Timperley, 2007), including improved achievement and understanding in mathematics (Wiliam, 2007). As such, many educational systems have attempted to implement such assessment practices as AfL, but research shows that learning how to practice AfL is challenging for teachers (Hopfenbeck et al., 2017; Nortvedt, Santos, & Pinto, 2016). AfL is strongly connected to ideas of equity in education. Formative use of assessment data should result in targeted interventions and ensure that all students are engaged in challenging mathematics learning (Heritage & Wylie, 2018).

The term *mapping tests* traditionally denotes assessments that are used to identify (map) what students can do (Ginsburg, 2016), with mathematics mapping tests often focussing on student misconceptions (Burkhardt & Schoenfeld, 2018), for instance, in relation to understanding numbers and number operations (e.g. Wiliam, 2007). As such, mapping tests have traditionally been used in the Nordic countries to provide tools for teachers that can be used to inform teaching (Räsänen et al., 2019). However, Gersten et al. (2009) claim that mapping tests only have an effect when followed up with targeted interventions. In other words, implementing national-level assessments alone is not sufficient to improve equity.

In addition to mapping tests, screening tests have been used in special needs education to identify students at risk of learning difficulties or lagging behind (Gersten et al., 2009). The main aim of screening tests is to divide students into groups, not provide information about students that can inform teaching. This aim influences the assessment design and screening tests are usually designed to provide information mainly around the cut-off score to avoid erroneous classification of individual students. As such, it is challenging to use screening tests formatively.

The focus on AfL could be disrupted if teachers and schools perceive national-level assessment initiatives, such as mapping tests, as high-stakes tests. Internationally, researchers have raised a concern that, when test content is known

to teachers, it could lead to increases in test scores rather than increased student achievement (e.g. Harlen, 2007). Moreover, increases in scores could represent test inflation due to teachers practicing the test content with their students (e.g. Stobart, 2008). Prior research has repeatedly found that teachers who administer what they perceive as high-stakes tests focus on the content of the tests, administer repeated practice tests, train students how to respond to specific types of questions and adopt transmission styles of teaching (Stobart, 2008). Such behaviours stand in the way of using assessment outcomes formatively to support the learning process (Brookhart, 2011; Burkhardt & Schoenfeld, 2018; Popham, 2009; Reay & Wiliam, 1999). Therefore, teachers' assessment literacy is fundamentally important for their understanding of the purpose of the assessment and their ability to use the assessment outcomes formatively (Popham, 2009).

9.2.3 Teachers' Assessment Literacy

Teacher assessment literacy can be defined as their understanding of the principles of sound assessment (Popham, 2009; Stiggins, 2005). This includes knowledge about tests, interpretation of test results, and most importantly, understanding how to apply these results to improve student learning. These elements are key aspects of assessment literacy because adjusting instruction and knowing what to teach next are critical components of AfL from an equity perspective (Heritage, Kim, Vendlinski, & Herman, 2009). According to Brookhart (2011), teachers need to be able to analyse tests to determine what knowledge and thinking skills are required for students to solve the test items. Such analytical skills can assist teachers in using assessment results to plan their future instruction and adapt it to all students. As part of this, teachers should be able to administer external assessments and interpret their results to form decisions regarding students, classrooms and schools (Brookhart, 2011; Campbell & Collins, 2007).

A positive attitude towards the use of assessment data to assist any student lagging behind is an important aspect of teachers' assessment literacy. Importantly, teachers need to be able to cooperate with school leaders and teaching colleagues in interpreting and using assessment data to the best advantage of their students. This is an important contribution to equity because it fulfils the fundamental principle of adapted education that is a core value in the Nordic educational systems (Telhaug et al., 2006). Assessment literacy is closely related to understanding diversity and adaption of instruction. For instance, research has shown that teachers often believe classroom tests provide more cognitive diagnostic information than national-level tests do regarding students' learning processes, consequences for meaningful learning and use of learning strategies (Leighton, Gokiert, Cor, & Heffernan, 2010). Such beliefs could indicate a gap in the teacher's assessment literacy that might influence the extent to which the teacher will be able to use assessment data from external tests to enhance student learning.

9.2.4   National-Level Assessments from an Equity Perspective

The Nordic model emphasises education for all, and early intervention and AfL are implemented through national policies to ensure equitable education (Imsen et al., 2017; Telhaug et al., 2006). Within the Nordic countries, children have the right not only to go to their neighbourhood school but also to receive education that will help them fulfil their potential (Buchholtz et al., 2020). This is implemented in the Norwegian Educational Act, for instance, by means of the principles of inclusion, AfL and adapted teaching for all students (Forskrift til opplæringslova, 2006). Policy-level initiatives to steer and strengthen learning in schools through national-level efforts focus mainly on the curriculum; however, they also consider assessment practices. In an international context, research has shown that national-level efforts often prioritise the use of summative assessment for accountability and monitoring purposes, rather than formative-oriented assessment formats (Stobart, 2008). This is somewhat different in Norway, where only formative assessment is implemented in primary education (Forskrift til opplæringslova, 2006).

The NQAS differs from many other systems in that it includes national-level assessments to be used formatively (Andreasen & Hjörne, 2014; Blömeke & Olsen, 2018; Elstad, Nortvedt, & Turmo, 2009). Regarding primary school, Sweden has national tests in mathematics in grade 3 and Denmark in grades 2–6. In both countries, teachers should use the national tests to determine the extent to which students have reached curriculum goals (Skolverket, n.d.; Børne- og undervisninsministeriet, n.d.).[2] According to Andreasen and Hjörne (2014), these assessments function primarily as external summative assessments in contrast to the formatively oriented Norwegian mapping tests. However, both Denmark and Sweden have national policies highlighting that teachers should use test outcomes as part of their on-going assessment of their students. In this respect, the Swedish and Danish primary schools could function formatively.

For a national-level effort to contribute to equity, it should be used to adapt teaching and assessment to the needs of individual students. According to Nordenbo et al. (2009), it is crucial for teachers to find that they can use the national-level assessments outcomes in their work and feel ownership over the assessment data, as well as to perceive that they can influence matters regarding implementation of the assessment; these factors all influence teachers' intentions to use the assessments.

In our opinion, it is not sufficient that assessments are formatively oriented; test outcomes also need to be used formatively to improve instruction. If schools and teachers simply use the test score for comparison, this will lead to a mere summative use of the test, which will stand in the way of School for All (Andreasen

---

[2] Finland do not have external assessments at the primary level aimed at individual students at the primary level (https://www.infofinland.fi/sv/livet-i-finland/utbildning/det-finlandska-utbildnings systemet).

& Hjörne, 2014). This may indicate that the formative use of the test is necessary for the assessment to contribute to equity.

9.2.5   The Norwegian Context

In 2006, the Norwegian Ministry of Education and Research released the white paper titled 'Early Intervention for Lifelong Learning', presenting a national policy for how the education system may contribute to social equalisation (Kunnskapsdepartementet, 2006). This white paper refers to the Organisation for Economic Co-ordination and Development (OECD) evaluation of assessment practices in Norway, which pointed to Norwegian schools having weak strategies for following up students lagging behind due to a lack of information on student progression. Unclear descriptions of expected learning outcomes and a lack of mapping tools for identifying students in need of extra teaching were also highlighted. National-level research also demonstrated that Norwegian teachers tended to 'wait and see' when students demonstrated difficulties (Nordahl & Hausstätter, 2009; Solli, 2005).

Following advice given in the policy, the first primary school mapping tests were introduced in 2008. The second generation of mapping tests was implemented in 2014 and is still in use (Utdanningsdirektoratet, 2018). The mapping tests were intended as a tool that could support teachers and schools in identifying students at risk at an early stage and help teachers adapt their teaching to these students' needs (Nortvedt, 2018). In other words, although the tests are taken by all students, they mainly provide information about the identified students.

9.3   The Present Study

The aim of this paper is to investigate how national-level assessments might contribute to equity in schools using the Norwegian mapping tests introduced in 2014 as our case. As such, we aim to answer three research questions (RQs):

- RQ1: What happens to the test quality when an assessment is exposed over time in terms of its psychometric properties?
- RQ2: To what extent are students identified as at risk at one grade level still at risk at the next? In other words, to what extent does the assessment contribute to improved learning for the students identified as at risk?
- RQ3: To what extent do teachers understand and use the outcomes of the mapping tests to improve learning for students identified as part of their assessment literacy?

9.4   Method

To answer the three research questions, this chapter draws on quantitative and qualitative data related to different aspects of the implementation and use of the mapping tests in Norwegian mathematics classrooms as follows: quantitative data at the student level from the mapping test implementations in 2014–2019 and qualitative data at the teacher level from semi-structured interviews conducted in 2016. By combining the strengths of both quantitative and qualitative aspects of data analysis and large datasets, we aim to provide complementary and deeper knowledge that can contribute to educational research on equity as it is understood in a Nordic context.

9.4.1   Design

Addressing RQ1, student-level data from the mapping test implementation in 2014–2017 were used to investigate the test quality of each of the three mapping tests. The main aim was to investigate whether the assessments retain their psychometric properties over the period of 4 years. Data from the test implementation in 2015–2017 were linked to data from the first implementation in 2014, applying a concurrent calibration using Xcalibre to investigate whether students of a given ability level had the same probability of getting a certain total score on a test across implementations.

   To address RQ2, we drew on data from 11 schools that were invited to participate in a three-year project, providing item-level data for their students for each year (2018–2020). Data from 2018 and 2019 were used to investigate what happened over time with students who were identified as at-risk students in grades 1 or 2.

   Finally, addressing RQ3, data on the teacher level ($N = 7$) from semi-structured interviews were used to investigate how teachers conceive, implement and follow up on the mapping tests. Teachers' engagement with the mapping tests provides insights into how the mapping tests are used and the extent to which they might contribute to enhancing equity.

9.4.2   Samples and Recruitment

Sample 1 comprises data on the item and student levels (grades 1–3) for each mapping test implementation from 2014 to 2017. A new sample was selected each year, meaning that sample 1 is suitable for investigating test quality (Table 9.1).

   Sample 2 is a convenience sample consisting of item-level data from grade 1–3 students in 11 schools. The total sample is presented in Table 9.2. It should be noted

Table 9.1   Sample for each test implementation for grades 1, 2 and 3 in 2014–2017

	2014	2015	2016	2017	Total
Grade 1	2281	989	969	906	5145
Grade 2	2483	1102	1073	886	5544
Grade 3	2286	967	947	888	5088

Table 9.2   Sample for each test implementation for grades 1, 2 and 3 in 2018–2019

	2018	2019	Total
Grade 1	425	268	693
Grade 2	354	363	717
Grade 3	316	213	529

that, due to students changing schools or a lack of parental consent[3] to participate in the study, this sample is limited to only a part of the total sample for 2018 and 2019. This means that the combined sample participating in both grades 1 (2018) and 2 (2019) includes 259 students, while the combined sample participating in grades 2 (2018) and 3 (2019) includes 150 students. As these samples are small, both quantitative and qualitative analyses are necessary to analyse the data.

For both samples 1 and 2, the school principal was first approached and asked if the school could participate in the data collection. For sample 1, one school class at each grade level was invited to participate. For sample 2, all classes/students in grades 1–3 were invited to participate.

Sample 3 consists of seven teachers from four schools across two school districts (see Table 9.3). Six of the teachers were recruited through the school principal to participate in the study. The seventh teacher (David) was purposefully selected for the study due to his previous interest in the mapping tests and the lack of male teachers in the sample. All seven teachers provided informed consent to participate in the study.

9.4.3   Data Collection

To collect data on the item level for all students, the schools were asked to provide student booklets for each student. Data were coded and registered for later analysis, and one database was constructed for each assessment for each year. In addition, a combined database for each grade level comprising data from 2014–2017 was made, and two linked databases were constructed from sample 2 students who had participated in two consecutive years.

---

[3] Sample 2 data are collected with student identities and require parental consent. As such, the project has been reported to the Norwegian Centre for Research Data (NSD; project number 58107). Sample 1 data are collected without school and student identity and therefor do not require parental consent.

Table 9.3 Background information for participating grade 1, 2 and 3 teachers

Teacher	School	Grade level	Experience with implementing the mapping test
Anna	A	2	Experienced
Anita	A	1	Experienced in special needs but first year as a regular teacher
Bente	B	1	Experienced
Brita	B	1	Experienced
Camilla	C	1	Administered the mapping tests for the first time in 2016
Carina	C	3	Much experience at higher levels with other assessments; first time administering mapping tests
David	D	2	Experienced

The first author of this chapter conducted semi-structured interviews with seven teachers after the mapping test implementation in 2016. Each interview took place in a secluded room in the participant's school and lasted 60 min on average. All interviews were audiotaped and later transcribed. Two grade 1 teachers working closely together (Bente and Brita) were interviewed together. All other interviews were individual. The teachers were asked how they prepared for and implemented the mapping test with their students, analysed the test outcomes and followed up the mapping test results with identified students. The tests were taken in late March or early April, and the interviews were conducted in late June.

9.4.4 Data Analysis

Regarding RQ1, item response theory (IRT)-based test-equating procedures in the form of concurrent calibration were performed to investigate the extent to which item characteristics were maintained over time or whether test inflation occurred. Concurrent calibration was the preferred test-equating procedure because it allows pairwise comparison of test characteristics across two timepoints. The assumption here is that the test measures the same construct at both administrations.

As the tests were not changed between 2014 and 2019, and because the same tests were implemented at each timepoint, all test items have been treated as anchor items. Thus, the ability estimates (θ) from the different test administrations (at the same grade level) resulting from such calibration will be on the same scale as one another, making the scores from two tests comparable because both the a and b parameters are invariant across the population.

To investigate how the mapping tests affect students over time (RQ2), a small subsample comprised data on two timepoints for students moving from grade 1 to grade 2 and for students moving from grade 2 to grade 3. These data were used to investigate how student results typically develops across the two timepoints. In this

analysis, we primarily used descriptive statistics, such as averages, cross-tables, the analysis of variance (ANOVA) test and chi-square analysis.

Regarding RQ3, the interviews were analysed using meaning condensation following Kvale and Brinkmann (2009) to identify the teachers' conceptions of both the mapping tests and the identified students. This analysis aimed to uncover teachers' experiences and their reflections on test administration and data analysis in addition to following up on students.

In the first stage, three of the authors analysed the data separately (Kvale & Brinkmann, 2009). In the next stage, the authors alternated between working individually and collaboratively to enable meaning condensation and interpretation of the interview data.

Table 9.4 illustrates meaning condensation of natural units of teacher statements. During the interviews, teachers provided rich descriptions of their work and reflections, enabling their talk to be broken down into natural 'meaning units' that were analysed using meaning condensation. Finally, derived meanings were interpreted. All quotes used in the results section have been translated from Norwegian to English by the authors. Rather than translating them word by word, the translations focus on representing the core ideas and understandings expressed by the teachers to better align with the applied analytical process.

Table 9.4 Illustration of meaning condensation and interpretation

Natural unit (teacher statement)	Central themes	Interpretation
Student follow-up—We have sometimes made course groups,	Establish groups to organise follow-up	Making groups organise the follow-up
We score, I do think that mapping is for finding out where on the map students are and what they need to practice.	Mapping test will help teachers identify what students can do and need to work on	Mapping test is a tool for teachers
So, we have made course groups, for instance, if we see that many share the same challenges in relation to a topic, so, we have made them across classrooms, because we have three classes, so I think the courses are OK.	Following up groups of students with similar challenges across classes makes courses work well	Following up in groups is appropriate
[Oppfølgingen av elevene; vi har kjørt noen ganger litt kursgrupper, vi retter, jeg tenker jo at kartlegging er for å se hvor elevene er i terrenget og hva man trenger å øve på. Så har man kjørt litt kursgruppe for eksempel hvis man ser så, så mange som har utfordringer hvis det gjelder ett emne så har man kjørt, smelt dem sammen. og det er på tvers av trinnene da (hmmh) for det er tre klasser hos oss. så synes jeg går helt kurant.]		

9.5   Results

In this section, we present the results following the order of the research questions that guided our investigations. The insights gained from the data analysis related to the three RQs, as well as the relationship between the three outcomes, are further elaborated on in the discussion section.

9.5.1   What Happened to the Mapping Test Quality After Five Test Administrations?

RQ1 focussed on what happens to the quality of the tests when the assessments are exposed over time. Specifically, do the assessments retain their psychometric properties even after four test administrations? Figure 9.1 shows the test response function (TRF) for the grade 2 test for the 2014–2017 test administrations. The curves more or less overlap, revealing that a student with a certain ability level in 2015–2017 had more or less the same probability of providing the same proportion of correct responses as a student with the same ability level in 2014. This means that the expected test performance is the same across years, and the examinees show the same expected distribution of performance in the four test administrations. The cut-off score calculated in 2014 is 41 points ($\theta = -1.366$). This is close to where the test has the maximum information and the measurement error is very small (0.20).
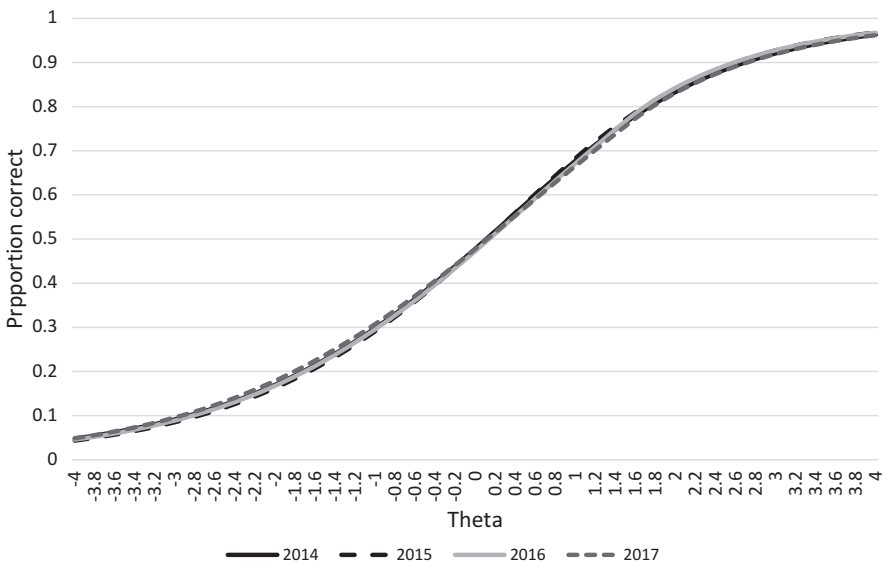
Fig. 9.1   Test response function for the grade 2 test for the 2014–2017 test administrations

Previous research in other countries has often found test inflation in exposed assessments. Test inflation typically happens for two different reasons, which are as follows: (1) teachers practise with students so they know how to respond to the test questions in advance and (2) teachers use their familiarity with the test and the test outcomes to improve their teaching. In Fig. 9.1, this would have been the case if the TRF graphs representing 2015–2017 test administrations rose above the line representing the 2014 administration. However, as shown in this figure, no test inflation was observed for the grade 2 mapping test.

Similar outcomes were obtained for the grades 1 and 3 mapping tests. Taken together, these outcomes lead to two likely interpretations, which are as follows: (1) there is no inflation in test scores due to test robustness, and (2) schools seemingly do not succeed in utilising the assessment data to improve mathematics instruction in primary grades 1–3. While the first interpretation points to test quality, the second points to potentially low assessment literacy or interest in using the assessment data. Neither interpretation can be excluded based on the current analysis.

9.5.2   What Happens over Time to Students Identified as 'At Risk' in Grade 1 or 2?

Data from the linked database, comprising data from the 2018 and 2019 samples, were used to investigate our second research question on what happened to students identified as being at risk in grades 1 or 2 in 2018: Were these students still below the cut-off score in 2019?

Table 9.5 shows the outcome patterns for the students ($N = 259$) who attended grade 1 in 2018 and grade 2 in 2019, while Table 9.6 shows the outcome patterns for the students ($N = 150$) who attended grade 2 in 2018 and grade 3 in 2019. Table 9.5 reveals that approximately 20% of the 259 students going from grade 1 to grade 2 were below the cut-off score in grade 1, grade 2 or both years. In this sample, nearly 1 in 10 students was identified as at risk in the two consecutive school years. While 5% of the grade 1 students were no longer identified as at risk in the following year, there was also a relatively large group of students (7%) who were not identified in grade 1 but fell below the cut-off score in grade 2 and were identified as at risk.

Similar patterns were observed for the transitions from grade 2 to grade 3 (Table 9.6). Nearly 20% of the students were identified as at risk in one or both school years. Fewer students (5%) were below the cut-off in both years. The same

Table 9.5 Achievement levels of students in grade 1 (2018) and grade 2 (2019)

		Grade 2, 2019		
		Identified as at risk	Not identified as at risk	Total
Grade 1, 2018	Identified as at risk	9%	5%	14%
	Not identified as at risk	7%	79%	86%
Total		16%	84%	100%

Table 9.6 Achievement levels of students in grade 2 (2018) and grade 3 (2019)

		Grade 3, 2019		
		Identified at risk	Not identified as at risk	Total
Grade 2, 2018	Identified as at risk	5%	6%	11%
	Not identified as at risk	10%	79%	89%
Total		16%	84%	100%

Table 9.7 Average scores in grades 1 and 2 for groups of students identified as at risk in both years, increasing from at-risk status, falling to at-risk status or scoring above the cut-off score in both years

	%	Average score in grade 1 (SD)	Average score in grade 2 (SD)
At risk both years	9%	32.6 (4.562)	28.5 (8.523)
Increasing from at risk	5%	36.1 (2.397)	46.1 (2.290)
Falling to at risk	7%	43.1 (2.198)	37.4 (3.789)
Above both years	79%	47 (2.648)	50.9 (3.231)

Note: Maximum score (cut-off) for grade 1 is 50 (39) and for grade 2 is 55 (41)

number of students moved from below to above the cut-off score. In this sample, 1 in 10 students was not identified as at risk in grade 2 but was identified as at risk in grade 3.

The outcomes may indicate that some teachers succeed in using the test results to improve student learning for their students. At the same time, they also show that some students identified as at risk in 2018 were still identified as at risk in 2019. This may indicate that the second school year did not provide students with sufficient opportunities for learning numeracy.

Table 9.7 shows the average scores for the groups of students that scored below the cut-off score in grade 1, grade 2 or both years. The tests have a ceiling effect, which affects the average scores for the group scoring above the cut-off score in both years. However, for the other three groups, average scores can be calculated, and an ANOVA test demonstrates significant differences between the four groups in both years [$F(3,255) = 221.615$, $p < .001$ and $F(3,255) = 264.286$, $p < .001$], with one exception: The students who have improved their results from below to above the cut-of score, in grade 3, does not score significantly lower than the group of student who scored about both years.

The group average scores presented in Table 9.7 indicate that the students who were identified as at risk in both years scored significantly below the cut-off in grade 1, with an average of 32.6 points (cut-off 39 points), but they scored even further from the cut-off score in grade 2, when the average scores was 28.5 points (cut-off 41 points). This indicates that the teachers did not succeed in increasing the at-risk students' attainment, and the increased standard deviation supports this interpretation. The students who transitioned from below to above the cut-off score, on

Table 9.8 Average scores in grades 2 and 3 for groups of students identified as at risk in both years, increasing from at-risk status, falling to at-risk status or scoring above the cut-off score in both years

	%	Average score in grade 2 (SD)	Average score in grade 3 (SD)
At risk both years	6%	32.1 (8.880)	46 (12.052)
Increasing from at risk	5%	40 (1.414)	63.88 (3.137)
Falling to at risk	10%	47.7 (3.826)	53.27 (4.605)
Above both years	79%	51.9 (2.6435)	68.4 (3.123)

Note: Maximum score (cut-off) for grade 2 is 55 (41) and for grade 3 is 72 (59)

average, were closer to the cut-off values, and at the same time, they scored well above the cut-off score in grade 2. In addition, the standard deviation was smaller for the second year, indicating that the students were more similar regarding achievement levels in 2019 compared to 2018. There was also a group of students who, on average, scored well above the cut-off in grade 1 but below the cut-off in grade 2. Judging by the increased standard deviation, more variation is visible in student achievement in grade 2 for the latter group.

Table 9.8 shows the group average scores for students in grades 2 and 3, showing similar patterns to those revealed for the grade 1–grade 2 transition. Table 9.8 indicates that, at this level, the students who were identified as at risk in both years also scored significantly below the cut-off in both years. This indicates that the teachers did not succeed in increasing the at-risk students' attainment, and again, the larger standard deviation supports this interpretation. Judging by the larger standard deviation for the students identified as at risk both years, more variation is visible in student achievement for this group, something that may make it more challenging for teachers to interpret the test outcomes and response patterns of these students.

9.5.3   To What Extent Does the Mapping Test Function as a Tool for Teachers to Support Student Learning?

Teachers' assessment literacy is a determinant of their work with mapping tests. For this reason, the teachers were asked about how they prepare for, administer and follow up the mapping test with their students. During the interviews, the teachers also shared their views and experiences about the mapping test and their work with the students identified as being at risk.

In the responses, four of the interviewed teachers express that in their view the mapping test could work as a tool for teachers and help them identify topics to address with their students. David's metaphor about placing students on a map is related to AfL:

… Mapping students is done to see where on the map the students are and what we need to practice more. [… kartlegging er for å se hvor elevene er i terrenget og hva man trenger å øve på.]

At the same time, our analysis revealed that the purpose of the mapping test may be somewhat unclear to some of the interviewed teachers. Anna's reflection below illustrates this and shows that, although she also highlights the formative aspect of the assessment, she is uncertain whether this is an external assessment or a tool for teachers. This exemplifies how teachers might struggle with understanding what distinguishes one test from another:

Anna: But I do not really know what the purpose [of the assessment] is. Is it like a national test where you should give feedback immediately? Or is it more like a tool for us teachers, you know? [Men jeg vet egentlig ikke helt hva som er målet. Er det som en nasjonal prøve som man skal gi tilbakemelding med en gang? Eller er det et verktøy for oss lærere ikke sant?]

Although emphasising the formative aspect, David also indicates that the mapping test provided insight into his teaching, suggesting that teachers may see alternative uses for mapping test data.

Bente and Brita, the two grade 1 teachers, report that they administered the test according to set guidelines, and they devote considerable time to analysing the test results. Even so, they express scepticism towards the test, partly because they believe the students are too young, and there is a risk of the testing being an uncomfortable experience for some students. They clearly express that conducting the mapping test is something they are obligated to do, and they are somehow sceptical of the test results. However, they view the assessment as a tool they can use to improve their instruction.

To prepare for the test, teachers need to go through tutorial materials that include instructions for how to administer the test. All the interviewed teachers state that it is important to create standardised conditions for all students in the testing. Internal school guidelines, in addition to the national guidelines, help the teachers create equal conditions when adapting the test situation to individual students as well. At the same time, however, the teachers sometimes feel the guidelines contribute to inequity. The test is timed so that students with naïve or rigid strategies will not have time to finish calculation tasks using these strategies. In particular, the time restraints are viewed as frustrating by the teachers, who find them unfair for low-achieving students:

Anita: We got a little frustrated with the time restriction because some first-graders would have done much better if the test wasn't timed. Because then I think everyone could have shown what they knew, not how much they could accomplish in a certain amount of time. [Vi ble litt frustrerte av det med tida fordi noen førsteklassinger hadde gjort det bra hvis det ikke var på tid. Fordi da tenker jeg da at alle hadde fått vist hva de kunne, ikke hvor mye de kunne prestere på et visst tidsrom.]

As Anita's statement exemplifies, the teachers feel their students would be able to show more of their competence if they had more time to respond to the test items. Thus, the teachers sometimes express that the test results do not reflect the

perceived level of competence of their students. The teachers also mention other factors, for example, the scoring procedures or unfamiliar item formats, which they feel affect student results.

The interviewed teachers show an awareness of other factors influencing test outcomes related to the student or the student's background, including learning difficulties, difficult situations at home, lower attention levels, misconceptions, linguistic challenges or careless mistakes. This is expressed by David in the following quotation, in which he points to factors outside school that influenced the test taking of two of his students, and consequently, wrongfully identified them as at risk:

> David: I think a lot of things in her life aren't so easy for her in general (…) and if then in a way her life outside of school has taken hold at an unlucky time, it might explain, right (…) So two of the students that are at risk, it is not mathematics interventions but other interventions that are needed. [Her for den ene sin del så tenker jeg at hun ikke har det så lett generelt {…} Og hvis da på en måte livet hennes utenfor skolen har gjort seg gjeldende på et uheldig tidspunkt så kan det forklare, ikke sant {…} Så to av de elevene de har under bekymringsgrensa så er det ikke matematikkfaglig tiltak, men andre tiltak som er nødvendig.]

All seven teachers indicate that they spend considerable time preparing for, administrating and attempting to understand the outcomes of the mapping test. None of the interviewed teachers report any difficulties scoring the tests, but analysing the data is challenging for many of them. Judging by his statement above, David connects difficulties with analysing data to a lack of classroom-level teaching initiatives.

Teachers may struggle to interpret the test results if they do not trust them. Moreover, analysis of the teacher interviews indicated that the teachers prioritise identifying student errors, misconceptions and mistakes the students might make if they misunderstand the task instructions. This could explain why it is difficult to plan interventions, as AfL builds on what students know and can do. Still, some interviewed teachers show awareness of their instruction and how this might influence student learning, as well as how it might influence the mapping test results and response patterns.

Overall, the seven teachers list many kinds of teaching innovations aimed at individual students or groups of students, in small-group or classroom teaching, including the following: engaging in learning conversations with students, setting up learning goals for individual students, using extra time when available with identified students, using more manipulatives and concrete materials when teaching, station teaching and grouping identified students with similar difficulties to work on specific topics. Moreover, differentiating task or activities during whole-class instruction, introducing peer assessment and learning partners, making courses for groups of students and focussing on mathematical concepts are also highlighted. However, most of the teachers state that they lack time to follow up on the students after the test, and thus, their main efforts have to wait until after the summer holiday. To facilitate more teaching interventions, they need time to plan (independently and in cooperation with colleagues) to identify necessary resources (time and teaching materials) and how teaching students in cooperation with colleagues could target

identified student needs. Moreover, they indicate that, in this process, they need help from the leaders in their school.

9.6 Discussion

Our analysis revealed that the mapping tests are robust; the item and test characteristics have not changed significantly over time (RQ1). Some students improved their results over time, while some did not, and some students even showed a decline in their understanding of numbers and calculation skills (RQ2). Moreover, although the teachers took care to administer the test following national and school guidelines, they struggled to interpret the test outcomes, and although a wide variety of interventions were listed, they were sometimes delayed until the fall (RQ3).

To frame our discussion, we draw on prior research on how assessment initiatives can be used to enhance equity in schools. In addition, prior research on equity (Espinoza, 2007; Zhu, 2018), AfL (Heritage et al., 2009; Wiliam, 2007), teachers' assessment literacy (Brookhart, 2011; Popham, 2009) and what teachers need to be able to do to use assessment data to improve students' opportunities to learn is used to discuss possible lessons learned from the Norwegian mapping test implementation.

9.6.1 National-Level Initiatives Such as the Mapping Tests May Contribute to Equity in Schools

For national-level assessments, such as mapping tests to contribute to equity, they need to be robust and identify students at risk of lagging behind (Brookhart, 2011; Stobart, 2008). In addition, teachers need to be able to administer the test in the same way and use the test outcomes to improve their teaching (Stobart, 2008). The IRT analysis demonstrated that the Norwegian mapping tests are robust, and judging by the interviews, the teachers managed to implement the assessment according to the national guidelines. As such, mapping tests may contribute to equity.

The analysis of the test data indicates that Norwegian teachers likely do not 'teach to the test' because the mapping tests functioned in the same way after several years of exposure. An alternative interpretation is that what the teachers practice with the students did not influence students' ability to respond to the test items. This outcome is contrary to the test score inflation that has been observed in other countries (e.g. Stobart, 2008), and it may be related to the school's ownership of test outcomes. We argue that, in situations with low-stakes national assessments, no test inflation and locally owned data, external assessments may contribute to equity because the teachers can feel more ownership to the data and influence over the use

of the assessment. Taken together this may provide more reliable measures for the identified students.

A third explanation, and a slightly less positive one, is that Norwegian teachers have not improved their instructional practices sufficiently, and over time, they have not offered better opportunities for learning for students identified as being at risk by the mapping tests. The analysis of what happens to identified students over time supports this interpretation: Some identified students (8% in total) were still at risk in the following school year. However, at the same time approximately one in two students identified as at risk in 2018 (or 7% of the total sample) scored above the cut-off score the following year. As such, we take these outcomes to mean that mapping tests can contribute to equity. Still, to improving equity classroom instruction needs to offer identified students possibilities to develop better conceptual understanding and calculation skills related to the key aspects of the mapping tests. The analysis of the interview data supports this interpretation because follow-ups were often delayed.

Previous research indicates that teachers often lack necessary assessment literacy to follow up on assessment outcomes (Heritage et al., 2009; Leighton et al., 2010). Some statements from the interviews may indicate that this is the case for some—but not all—of the interviewed teachers. As such, understanding how teachers' conceptions and beliefs about the mapping test interact with AfL initiatives is crucial.

9.6.2 Teachers' Assessment Literacy and Assessment for Learning Practices Conditions How Mapping Tests Might Contribute to Equity

According to Gersten et al. (2009) and Brookhart (2011), mapping tests need to be followed up with targeted instructions to improve learning. The tests are administered in the spring, and most of the interviewed teachers stated they experienced a lack of time to follow up with the students in the spring semester. Instead, they planned to do so after the summer break. Perhaps this notion of the mapping tests as end-of-year tests causes the teachers to view them as summative rather than as part of the on-going formative assessment they conduct during the school year. Stiggins (2005) argues that assessment that takes place during the learning process can contribute to the formative use of tests, and thus, promote student learning. The teachers' statements about following up during the autumn semester support this summative conception of the tests' purpose. In addition, teacher statements about already knowing who struggles prior to the mapping test supports the interpretation of viewing the tests as summative. We argue that teachers need to view and use the tests as formative for them to contribute to equity (e.g. Heritage et al., 2009). Still, the seven teachers had already implemented some teaching interventions in the late spring and early summer, and such activities as peer assessment, setting learning

goals and involving students in learning conversations can be viewed as AfL activities. We argue that whether teachers view the mapping tests as summative or formative depends on their assessment literacy.

The mapping test data are locally owned. The intention with the mapping tests is that schools and teachers will feel ownership, and the primary goal is teaching interventions rather than reporting. As such, the tests can function as a support tool and not an accountability measure. However, the interviews showed that we might question whether every interviewed teacher view the tests as a tool for improving teaching and learning.

The analysis of the interview data revealed that the teachers held different perceptions about classroom and external tests. According to Brookhart (2011), this could influence their assessment literacy. Leighton et al. (2010) noted that many studies have shown that teachers have somewhat negative attitudes toward large-scale national assessment. Our study may support this finding, as some interviewed teachers saw the mapping tests as an external assessment that evaluates students rather than a tool they could use to improve teaching and learning. At the same time, we found that the interviewed teachers sometimes did not trust the test results of identified students, and it may be inferred that they believed that students' test performance reflected test-taking strategies rather than numeracy skills.

Using assessment outputs to inform teaching is fundamental to formative assessment (Brookhart, 2011). The support material that accompanies the mapping tests supposedly helps teachers do this. It provides information about how the test works, how to interpret the results, what it means when students are identified as at risk and suggestions for further instruction. However, based on the interview data, it is questionable whether all teachers are actually being provided with adequate support material.

9.7   Concluding Remarks—Linking Equity, National-Level Initiatives and Assessment Literacy

The overall question in this chapter related to whether a national-level initiative—in this case, the Norwegian mapping tests—can improve equity in schools. That there is no inflation in test scores supports using the same tests over time and trusting teachers to use them as intended. Further, a large proportion of the students who were below the cut-off score one year were above it the next. This could be due to factors not included in this research, but it may also be an outcome of using the information from the mapping test, and thus, contributing to equity. Overall, these observations support the idea that mapping tests can improve equity.

In accordance with Nordic model principles for transparency (Telhaug et al., 2006) and school autonomy, the mapping test content is available to teachers and schools, and test results are locally owned. Moreover, schools are trusted to use the mapping test outcomes in accordance with national guidelines. As a result,

familiarity with the test content helps ensure formative use of assessment outcomes to improve teaching and learning. For instance, Norwegian schools are responsible for identifying professional development needs for their teachers (Imsen et al., 2017). Imsen et al. (2017) discuss a dilemma in that schools simultaneously have to deal with national-level assessment and regulations while having autonomy to interpret the curriculum and plan instruction. Our research can be seen as confirming this dilemma between national-level and local initiatives and responsibilities.

Going forward would mean developing national-level initiatives that allow for local adaption that can assist schools in further developing teachers' assessment literacy. In addition, future endeavours should provide educators with the means to develop more knowledge and strategies for targeting their teaching to students at risk, thereby enhancing equity by being better prepared to adapt teaching. We propose a three-part strategy in line with the traditions and values in the Nordic model to ensure that national-level assessments contribute to equity in primary school as follows: (1) offering high-quality assessments, (2) offering helpful and useful tutorials and support material and (3) implementing national and local initiatives that can assist teachers in further developing their assessment literacy. Norway has implemented the first two of these. However, to take full advantage of these two parts of the strategy, we argue that the third is necessary because this will help all schools and teachers improve their assessment literacy, leading to more equitable mathematics education. At the same time, to be aligned with the Nordic model principles, transparency and school autonomy must be maintained.

Each of the elements identified above seems like sound advice, but we argue that it is only when they come together that we will see development. First, quality assessments are more than mere psychometric sound assessments. They are accompanied by documents that provide teachers and school leaders with insights into how the assessments are developed, what they measure and how they should be implemented. Second, the tutorial and support material should assist teachers and schools in analysing assessment data to understand what students know and can do. Moreover, it should also help teachers to translate this knowledge into an understanding of what students should learn next and how to achieve this. Only when this is in place will the assessment operate as AfL and contribute to equity (e.g. Heritage et al., 2009). Finally, to assist teachers and schools in using the mapping tests and tutorial materials and to ensure that this initiative fosters teachers' assessment literacy, we need to offer local and national support focussed on teachers' conceptions of students and assessment.

Teachers' positive attitudes toward mapping tests are instrumental to using the assessment outcomes to improve equity in school. Based on this argument, if teachers do not believe that the mapping tests are a helpful tool for improving their instruction, and if they do not have the necessary assessment literacy, the tests are not likely to contribute to improved teaching and learning opportunities for identified students or to equity. However, at the same time, we emphasise that, as researchers, we have a primary responsibility to conduct research that can inform all three aspects of the above-mentioned strategy to promote equity in primary school mathematics instruction.

References

Andreasen, K., & Hjörne, E. (2014). Assessing children in the Nordic countries: Framing diversity and matters of inclusion and exclusion in a School for All. In U. Blossing, L. Moos, & G. Imsen (Eds.), *The Nordic education model. 'A School for All' encounters neo-liberal policy* (pp. 155–172). Dordrecht, The Netherlands: Springer.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*, 7–74. https://doi.org/10.1080/0969595980050102

Blömeke, S., & Olsen, R. V. (2018). På vei mot et sammenhengende nasjonalt kvalitetsvurderingssystem. *Acta Didactica Norge, 12*(4), 1–21. https://doi.org/10.5617/adno.6278

Børne- og undervisninsministeriet [Ministry of children and education]. (n.d.). *Om de nationale test* [About the national tests]*.* https://www.uvm.dk/folkeskolen/elevplaner-nationale-test%2D%2Dtrivselsmaaling-og-sprogproever/nationale-test/om-de-nationale-test. Accessed 18 Aug 2020.

Brookhart, S. M. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practice, 30*(1), 3–12.

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* (pp. 13–41). Cham, Switzerland: Springer.

Burkhardt, H., & Schoenfeld, A. (2018). Assessment in the service of learning: Challenges and opportunities or Plus ça Change, Plus c'est la même Chose. *ZDM, 50*(4), 571–585. https://doi.org/10.1007/s11858-018-0937-1

Campbell, C., & Collins, V. L. (2007). Identifying essential topics in general and special education. *Educational Measurement: Issues and Practice, 26*(1), 9–18.

Elstad, E., Nortvedt, G. A., & Turmo, A. (2009). The Norwegian assessment system: An accountability perspective. *Cadmo, 17*(2), 89–103.

Espinoza, O. (2007). Solving the equity–equality conceptual dilemma: A new model for analysis of the educational process. *Educational Research, 49*(4), 343–363. https://doi.org/10.1080/00131880701717198

Forskrift til opplæringslova [Regulation to the Education Act]. (2006). *Chapter II Underveisvurdering [Day-to-day assessment], § 3–1*.

Gersten, R., Beckmann, S., Clarke, B., Foeggen, A., Marsh, L., Star, J. R., & Witzel, B. (2009). *Assisting students struggling with mathematics: Response to intervention (Rtl) for elementary and middle schools (NCEE 2009–4060)*. http://ies.ed.gov/ncee/wwc/publications/practiceguides/

Ginsburg, H. P. (2016). Helping early childhood educators to understand and assess young children's mathematical minds. *ZDM, 48*(7), 941–946. https://doi.org/10.1007/s11858-016-0807-7

Gutstein, E., Fey, J. T., Heid, M. K., DeLoach-Johnson, I., Middleton, J. A., Larson, M., et al. (2005). Equity in school mathematics education: How can research contribute? *Journal for Research in Mathematics Education, 36*(2), 92–100.

Harlen, W. (2007). *Assessment for learning*. Los Angeles: Sage.

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

Hattie, J. A. C., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112.

Heritage, M., Kim, J., Vendlinski, T., & Herman, J. (2009). From evidence to action: A seamless process in formative assessment? *Educational Measurement: Issues and Practice, 28*(3), 24–31.

Heritage, M., & Wylie, C. (2018). Reaping the benefits of assessment for learning: Achievement, identity, and equity. *ZDM, 50*(4), 729–741. https://doi.org/10.1007/s11858-018-0943-3

Hopfenbeck, T. H., Lenkeit, J., El Mastri, Y., Cantrell, K., Ryan, J., & Baird, J.-A. (2017). Lessons learned from PISA: A systemic review of peer-reviewed articles on the programme for international student achievement. *Scandinavian Journal of Educational Research, 62*(3), 333–353.

Imsen, G., Blossing, U., & Moos, L. (2017). Reshaping the Nordic education model in an era of efficiency. Changes in the comprehensive school project in Denmark, Norway, and Sweden

since the millennium. *Scandinavian Journal of Educational Research, 61*(5), 568–583. https://doi.org/10.1080/00313831.2016.1172502

InfoFinland. (n.d.). *Det finländska utbildningssystemet* [The Finnish educational system]. https://www.infofinland.fi/sv/livet-i-finland/utbildning/det-finlandska-utbildningssystemet. Accessed 18 Aug 2020.

Kunnskapsdepartementet [Ministry of Education and Research]. (2006). … *og ingen stod igjen. Tidlig innsats for livslang læring (St. meld. Nr. 16(2006–2007)* [… And no one was left behind. Early intervention for lifelong learning (White paper no. 16(2006–2007))]. https://www.regjeringen.no/

Kvale, S., & Brinkmann, S. (2009). *Interviews: Learning the craft of qualitative research interviewing*. Los Angeles: Sage.

Leighton, J. P., Gokiert, R. J., Cor, M. K., & Heffernan, C. (2010). Teacher beliefs about the cognitive diagnostic information of classroom- versus large-scale tests: Implications for assessment literacy. *Assessment in Education: Principles, Policy & Practice, 17*(1), 7–21. https://doi.org/10.1080/09695940903565362

Nordahl, T., & Hausstätter, R. S. (2009). *Spesialundervisningens forutsetninger, innsats og resultater* [Conditions, efforts and outcomes of special education]. Høgskolen i Lillehammer.

Nordenbo, S. E., Allerup, P., Andersen, H. L., Dolin, J., Korp, H., Søgaard Larsen, M., et al. (2009). Pædagogisk brug av test – et systematisk review [Using tests in education – a systematic review]. In *Evidensbasen*. Copenhagen, Denmark: Danish Clearinghouse for Uddannelsesforskning, DPU, Aarhus University.

Nortvedt, G. A. (2018). «Det er et verktøy, ikke sant, for oss» – Erfaringer fra fire gjennomføringer med kartleggingsprøver i regning 2014–2017. *Acta Didactica Norge – tidsskrift for fagdidaktisk forsknings- og utviklingsarbeid i Norge, 12*(4), 1–22.

Nortvedt, G. A., & Buchholtz, N. (2018). Assessment in mathematics education: Responding to issues regarding methodology, policy, and equity. *ZDM, 50*(4), 555–570. https://doi.org/10.1007/s11858-018-0963-z

Nortvedt, G. A., Santos, L., & Pinto, J. (2016). Assessment for learning in Norway and Portugal: The case of primary school mathematics teaching. *Assessment in Education: Principles, Policy & Practice, 23*(3), 377–395. https://doi.org/10.1080/0969594X.2015.1108900

Nortvedt, G. A., & Wiese, E. (2020). Numeracy and migrant students: A case study of secondary level mathematics education in Norway. *ZDM: Mathematics Education, 52*(3), 527–539. https://doi.org/10.1007/s11858-020-01143-z

Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory Into Practice, 48*(1), 4–11. https://doi.org/10.1080/00405840802577536

Räsänen, P., Daland, E., Dalvang, T., Engström, A., Korhonen, J., Kristinsdóttir, J. V., et al. (2019). Mathematical learning and its difficulties: The case of Nordic countries. In A. Fritz, V. G. Haase, & P. Räsänen (Eds.), *International handbook of mathematical learning difficulties: From the laboratory to the classroom* (pp. 107–125). Cham, Switzerland: Springer.

Reay, D., & Wiliam, D. (1999). 'I'll be a nothing': Structure, agency and the construction of identity through assessment. *British Educational Research Journal, 25*(3), 343–354.

Roos, H. (2018). Inclusion in mathematics education: An ideology, a way of teaching, or both? *Educational Studies in Mathematics, 2019*(100), 25–41. https://doi.org/10.1007/s10649-018-9854-z

Rousseau, C., & Tate, W. F. (2003). No time like the present: Reflecting on equity in school mathematics. *Theory Into Practice, 42*(3), 210–216.

Samoff, J. (1996). Which priorities and strategies for education? *International Journal of Educational Development, 16*(3), 249–271.

Skolverket [Swedish National Agency for Education]. (n.d.). *Nationella prov i grundskola* [National tests in compulsory education]. https://www.skolverket.se/undervisning/grundskolan/nationella-prov-i-grundskolan. Accessed 18 Aug 2020.

Solli, K. A. (2005). *Kunnskapsstatus om spesialundervisning i Norge* [Special education in Norway, status report]. Utdanningsdirektoratet.

Stiggins, R. (2005). From formative assessment to assessment for learning: A path to success in standards-based schools. *Phi Delta Kappan, 87*(4), 324–328.

Stobart, G. (2008). *Testing times: The uses and abuses of assessment*. London: Routledge.

Telhaug, A. O., Mediås, O. A., & Aasen, P. (2006). The Nordic model in education: Education as part of the political system in the last 50 years. *Scandinavian Journal for Educational Research, 50*(3), 245–283.

Utdanningsdirektoratet. (2018). *Hva er kartleggingsprøver* [What are mapping tests]? https://www.udir.no/eksamen-og-prover/prover/hva-er-kartleggingsprover/

Wiliam, D. (2007). Keeping learning on track. In F. K. J. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 1053–1098). Charlotte, NC: Information Age.

Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Evaluation, 37*(1), 3–15.

Zhu, Y. (2018). Equity in mathematics education: What did TIMSS and PISA tell us in the last two decades? In G. Kaiser, H. Forgasz, M. Graven, A. Kuzniak, E. Simmt, & B. Xu (Eds.), *Invited lectures from the 13th International Congress on Mathematical Education. ICME-13 Monographs* (pp. 769–786). Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-319-72170-5_43

Part III
Focus on the Students and the Learning Environment

Chapter 10
Can Teachers' Instruction Increase Low-SES Students' Motivation to Learn Mathematics?

Ole Kristian Bergem, Trude Nilsen, Oleksandra Mittal, and Henrik Galligani Ræder

Abstract Students' motivation in mathematics has been shown to predict their achievement and whether they pursue a later career in STEM (science, technology, engineering, and mathematics). To sustain equity in education, it is important that students are motivated for the STEM fields, independent of their background characteristics (e.g., gender and SES). Previous research has revealed that students' motivation declines from primary to secondary school. The present study investigates whether this unwanted development may be related to students' SES, and more importantly, what aspects of teachers' instruction are related to student motivation for low, medium, and high-SES student groups in grade 5 and 9. We use data from students in grades 5 and 9 and their teachers who participated in TIMSS 2015 in Norway. Multilevel (students and classes), multi-group structural equation modelling is used to answer the research questions. In line with previous research from Germany and the USA, the results showed that SES is more important to student motivation in secondary than primary school, that low SES students' motivation depends more on their teachers' instructional quality than high SES students and that this dependency is stronger in secondary school than in primary school. The implications and contributions of the study are discussed.

O. K. Bergem (✉) · T. Nilsen · O. Mittal
Department of Teacher Education and School Research, University of Oslo, Oslo, Norway
e-mail: o.k.bergem@ils.uio.no

H. G. Ræder
Centre of Educational Measurement, University of Oslo, Oslo, Norway

High performance and more positive attitudes towards schooling among disadvantaged 15-year-old students are strong predictors of success in higher education and work later on (Organisation for Economic Cooperation and Development [OECD], 2018.

10.1   Introduction

Various research studies have reported strong positive correlations between students' intrinsic motivation to learn mathematics and factors such as academic proficiency, the cultivation of science, technology, engineering, and mathematics (STEM) careers, and the fostering of feelings of well-being in school (Jansen, Schroeders, & Lüdtke, 2014; Mullis, Martin, Foy, & Hooper, 2016; Watt & Eccles, 2008; Wigfield et al., 2015). Apart from intrinsic motivation's importance for improving test scores and future career choices, it is one of the preconditions for shaping a positive learning process at school (Organisation for Economic Cooperation and Development [OECD], 2012) and is a non-cognitive skill related to later success in life, determining adolescents' socio-economic outcomes (Korbel & Paulus, 2018). We also know that parental socio-economic status (SES) has a positive effect on students' academic proficiency and motivation (Kriegbaum, Jansen, & Spinath, 2015; Sirin, 2005; Tenenbaum & Leaper, 2003).

In short, enhancing *all* students' intrinsic motivation is viewed as critical to sustaining equity in education (Mullis et al., 2016; Musu-Gillette, Wigfield, Harring, & Eccles, 2015; OECD, 2018; Simpkins, Davis-Kean, & Eccles, 2006; Spinath & Steinmayr, 2012).

A robust and problematic finding across studies is that while children have high levels of intrinsic motivation to learn mathematics when they enter school—and it remains relatively high throughout elementary school—by the end of lower secondary school, they tend to have considerably less motivation (Corpus, McClintic-Gilbert, & Hayenga, 2009; Fauth, Decristan, Rieser, Klieme, & Büttner, 2014; Gottfried, Fleming, & Gotfried, 2001; Mullis et al., 2016; Steinmayr & Spinath, 2009). This drop in students' motivation may negatively affect their decision to continue with upper secondary education and discourage them from choosing STEM careers (OECD, 2012). Although this decline in intrinsic motivation is well documented, it remains unclear what factors are involved in this unwanted development. For instance, few studies have investigated whether high- and low-SES students experience the same drop in intrinsic motivation to learn mathematics during the aforementioned period. There are even fewer studies that have investigated whether students from different SES groups profit to the same extent from high-quality instruction.

Indeed, students' intrinsic motivation has been found to be affected by teachers' instructional quality (InQ), which is an important agenda embedded in educational policies (Farrington et al., 2012; Korbel & Paulus, 2018). Various studies have examined the association between dimensions of InQ and intrinsic motivation, and some promising results have been reported. Kunter, Baumert, and Köller (2007)

found that higher levels of classroom management may positively affect students' intrinsic subject-based motivational development in mathematics. Other instructional aspects, such as providing a supportive classroom climate and affording high levels of instruction clarity and cognitive challenges, have also been found to enhance student motivation to learn mathematics (Baumert et al., 2010; Klieme, Pauli, & Reusser, 2009; Scherer & Nilsen, 2016; Seidel, Rimmele, & Prenzel, 2005; Wigfield et al., 2015). Such findings suggest that aspects of InQ are important in seeking to heighten students' intrinsic motivation to learn mathematics. However, little research attention has been paid to analysing whether students from different SES groups profit to the same extent from high-quality instruction. Kyriakides, Creemers, and Charalambous (2019) argued that from an equity perspective, it is extremely important to examine whether factors that are found to contribute to better student outcomes positively affect all groups of students similarly, including those who are more disadvantaged. They claim that such analyses could make a valuable contribution to designing educational systems that improve opportunities for low-SES students to succeed in school. Our study addresses this thematic challenge.

The present study's aim is twofold: First, it investigates how the SES of students in Norway is associated with intrinsic motivation to learn mathematics in the fifth grade compared to the ninth grade. Second, for these two grade levels, it examines how InQ is associated with students' intrinsic motivation among different SES groups of students.

10.2   Theoretical Framework

In this section, we will present the key concepts used in our overall framework, namely *equity*, *InQ*, and *motivation.* We will also provide a short review of previous research relevant to our analysis, particularly research into how motivation and InQ are related to SES and student outcomes.

10.2.1   Equity

In distinguishing between the concepts of 'equality' and 'equity' as used in the educational discourse, Espinoza (2007) argues that while equality is funded upon ideas from the French Revolution (liberty, equality and fraternity), asserting sameness in treatment for all people, equity is related to aspects of fairness and justice in the provision of education, or what could also be labelled 'social justice' (see Chap. 2). He contends that the equity concept allows for individual considerations and treatment and claims that in certain situations the concepts of equality and equity may seem to be mutually 'opposed' to one another (Espinoza, 2007). For example, achieving greater equity within a school system by affording students individually

adapted support may sometimes entail a reduction of equality when understood as the same treatment for all students (see Chap. 2 for an elaboration of these concepts). In line with these considerations, Kyriakides and Creemers (2011) argue that there is general agreement that equity does not imply everyone is the same or should achieve the same outcomes. However, differences in outcomes should not be attributable to factors related to student SES.

In line with the previously described studies, equity in our chapter implies that development of motivation towards mathematics is not linked to a student's background. In order to achieve this, some students may be provided with adapted resources, such as high-quality teachers.

One of the most important objectives in many educational systems worldwide is to provide equitable opportunities and fair learning environments to *all* students to ensure that they have the chance to realize their academic potential, regardless of gender, ethnicity, or SES (Opheim, 2004). Within this context, when schools provide fair and inclusive teaching practices and fairly distribute educational tools and resources, they play a central role in compensating for unjustifiable differences in student outcomes that are attributable to their background (Field, Kuczera, & Pont, 2007; OECD, 2012). These two equity dimensions—fairness and inclusion—reflect the principal idea of effective schooling behind such large-scale international surveys as the *Trends in International Mathematics and Science Study* (TIMSS) and the *Programme for International Student Assessment* (PISA), which, among other school factors, emphasize the teacher's role in helping children overcome their socio-economic barriers to reach their full learning potential (Field et al., 2007; OECD 2012, 2018).

However, analyses of TIMSS and PISA have proved that many challenges remain in efforts to ensure equity in students' learning outcomes (Field et al., 2007; Gustafsson, Nilsen, & Hansen, 2016; OECD, 2012, 2018; Schmidt, Burroughs, Zoido, & Houang, 2015). It is important to examine the individual mechanisms that may undergird the association between students' SES and learning outcomes, as well as *whether* and *how* school-related factors—school organization, curriculum, recruitment of teachers and students, and InQ—can positively impact these mechanisms (Creemers & Kyriakides, 2008; Scheerens, 2014).

To sum up, with educational policies across several countries addressing the issue of enhancing student motivation as an outcome in itself, a need still exists for more knowledge about the relationship between students' SES and their intrinsic motivation. The present study will address this literature gap. Additionally, in light of Espinoza's (2007) definition of equity, along with the foregoing and the insight from Creemers and Kyriakides (2008) that more research is needed into how school factors may compensate for students' SES, selected InQ dimensions will be analysed to investigate *whether* and *how* they may contribute to students having equitable and fair opportunities to succeed.

10.2.2 Instructional Quality (InQ) and Its Relationship to Student Outcomes

In the educational research field, it has been acknowledged that the InQ construct should be viewed as having various aspects or dimensions (Fauth et al., 2014; Kane & Cantrell, 2010; Klette, 2015; Wagner et al., 2015). Baumert et al. (2010) and Klieme et al. (2009) have been particularly influential in developing InQ scales that have been used in several European educational studies, including PISA and TIMSS. In the present study, four InQ dimensions are measured: *classroom management,; supportive climate, clarity of instruction, and cognitive activation*.

10.2.2.1 Classroom Management

This InQ dimension focuses on classroom rules and procedures, how the teacher copes with disruptions, and how efficiently transitions are managed (Fauth et al., 2014). Such characteristics are viewed as essential to providing students with opportunities to learn (Dorfner, Förtsch, & Neuhaus, 2018; Pianta & Hamre, 2009). In several meta-studies, efficient time and classroom management have been found to be associated positively with student outcome measures, particularly achievement (Hattie, 2009; Seidel & Shavelson, 2007). Baumert et al. (2010) contend that this dimension is viewed as a particularly robust InQ measure.

10.2.2.2 Supportive Climate

The description of this InQ aspect builds on reports from motivational research studies, covering certain important aspects of the teacher–student relationship, such as constructive feedback and a generally positive approach to student misconceptions and errors. It also includes a teacher's caring behaviour toward students (Good & Brophy, 2000; Klieme et al., 2009). An important finding related to research of this dimension is that teacher support and scaffolding are crucial elements for heightening student engagement in insightful learning processes (Pianta, Nimetz, & Bennet, 1997). Thus, a supportive climate has been found to predict student interest and stimulate the development of a student's intrinsic motivation (Fauth et al., 2014; Klieme et al., 2009).

10.2.2.3 Clarity of Instruction

This InQ aspect is understood as a teacher's ability to provide clear and coherent presentations of content, goals, and tasks, which can be done through, for example, overviews, advance organizers, outlines, and periodic summaries (Brophy & Good, 1986). Another key feature of this dimension is linking instruction to students' prior

knowledge to allow new information to be integrated into existing knowledge structures (Duit, 2009). Positive relationships between clarity of instruction and student outcome measures have been reported in various studies (Creemers & Kyriakides, 2008; Scherer & Gustafsson, 2015). As for motivation, Seidel et al. (2005) found that a clear and coherent lesson structure was associated with a more positive student perception of supportive learning conditions, stimulating self-determined forms of learning motivation, including intrinsic motivation.

10.2.2.4   Cognitive Activation

Baumert et al. (2010) describe the level of cognitive activation as being determined mainly by the kinds of math problems presented to students and how the teacher implements them. An important aspect of this dimension is asking students to explain their answers and encouraging them to evaluate their solution's validity. Such classroom practices are viewed as a way to stimulate students' cognitive engagement and, consequently, lead to deeper and more elaborate knowledge (Klieme et al., 2009). Scholars have argued that cognitive activation is connected closely to subject matter (Baumert et al., 2010; Seidel & Shavelson, 2007). Results have been somewhat mixed in attempts to find associations between cognitive activation and student outcome measures (Hiebert & Grouws, 2007; Seidel & Shavelson, 2007). Regarding motivation, Fauth et al. (2014) found that primary students' cognitive activation ratings predicted their development of subject-related interest.

10.2.3   Instructional Quality (InQ) and Equity

As described above, the four InQ dimensions have been shown to influence student outcomes positively in terms of both achievement and motivation. A few studies have also investigated relationships between InQ and equity. Rjosk et al. (2014) found that cognitive activation in language instruction (German) mediated the effects of classroom SES composition on achievement. This was attributed in particular to teachers focusing less on challenging language instruction in low-SES classrooms. In a study using data from PISA 2006, Willms (2010) found that schools' SES effects were mediated by the quality of instruction and time allocated to science lessons. Using data from 50 countries that participated in TIMSS 2011, Gustafsson et al. (2016) investigated whether school characteristics, including InQ, moderated the relationship between student SES and mathematics achievement. Their findings were mixed in that InQ was found to generate compensatory effects in some countries and anti-compensatory effects in others. Compensatory national school systems tended to have relatively high achievement levels, and it was concluded that these systems can reduce the relationship between achievement and student SES through certain key factors, including high InQ.

10.2.4   *Intrinsic Motivation*

Motivational research is a broad and complex field of study. Within educational research, theories related to motivation systematically deal with one very important issue in particular: students' reasons for engaging in various kinds of achievement tasks (Eccles & Wigfield, 2002). Intrinsic motivation is a key concept frequently paired with and explained in relation to extrinsic motivation. *Intrinsic motivation* is defined as engaging in an activity for its inherent satisfactions rather than for some separable consequence (Ryan & Deci, 2000)—or to put it slightly differently, engaging in an activity for its own sake, such as for enjoyment, the challenge, interest in the activity, or natural fulfilment of curiosity (Barry & King, 2000). Thus, when a person is motivated intrinsically, learning can be viewed as a side effect of being engaged in the relevant actions (Weidinger, Steinmayr, & Spinath, 2017). In the mathematics classroom, students who are driven by a desire to learn—and who enjoy learning math—can be viewed as intrinsically motivated. Differently, extrinsic motivation is defined as activities that are pursued for expected external rewards unrelated to the activity itself (Eccles & Wigfield, 2002; Ryan & Deci, 2002). With mathematics, such external rewards can be higher grades, getting to the top of the class, or pleasing parents or teachers.

Ryan and Deci (2000) argue that humans are active, inquisitive, curious, and playful creatures who do not require extraneous incentives to learn and explore. However, it is clear that not all individuals are motivated intrinsically to engage in the same activities and tasks. Within pedagogical theory, the nurturing of a student's intrinsic motivation is a crucial part of teacher responsibilities. It is assumed that enhancing and sustaining students' intrinsic motivation for learning is critical to preparing children for successful mastery of future challenges, and such motivation should be viewed as a highly desirable developmental outcome (Ryan & Deci, 2009; Spinath & Spinath, 2015).

10.2.5   *Motivation and Equity*

To discuss and draw any causal link between students' SES and their intrinsic motivation or interest, it is necessary to understand the relationship that this theoretical construct has with other similar non-cognitive constructs, namely academic self-beliefs, which are often investigated in regard to their connection to a child's SES. Interest was initially treated as the affect component of academic self-concept or self-belief. Eccles's expectancy-value theory (EVT; Eccles et al., 1983; Eccles, 2009) separated it through a hierarchy of self-beliefs and subjective task values that are different, but positively interrelated, components of academic motivation. Subsequently, Marsh, Craven, and Debus (1999) found interest to be empirically distinguishable from academic self-concept. Further empirical studies of relationships between self-concept and intrinsic motivation found self-concept to be the

strongest factor affecting students' subsequent interest in the relevant subject (Cheung, 2018; Häussler & Hoffmann, 2000; Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2005a; Viljaranta, Tolvanen, Aunola, & Nurmi, 2014). These findings extended a variety of possible mechanisms through which parents' SES may impact their children's intrinsic motivation, thereby allowing us to discuss it in a broader theoretical and empirical context.

The development of students' intrinsic motivation takes place within multiple learning environments. The family is the first cultural and social milieu in which characteristics might exert a lasting effect on the way a child interprets other educational contexts, thereby shaping his or her academic interests and aspirations (Bandura, 2012; Boudon, 1974; Bourdieu, 1986; Eccles, 2009). For example, Eccles's EVT model refers to parents as socializers (along with teachers, peers, etc.), and children's achievement-related activities and choices are the product of a continuous negotiation of meanings in the hierarchy of learning environments. According to Bandura's socio-cognitive perspective, students' self-beliefs and academic motivation are shaped by parents' familial belief systems, which are influenced by their SES (Bandura, 2012; Bandura, Barbaranelli, Caprara, & Pastorelli, 2001). To take it further, the reproduction theories argue that high-SES parents provide their children with more stimulating environments and use complex linguistic codes that might enhance their children's motivation and ability to succeed academically (Bernstein, Bernstein, & MacRae, 1971; Bourdieu, 1986).

The somewhat limited empirical research on the association between SES and intrinsic motivation generally finds it to be significant, with some variation in effect size. This variation is mainly due to the SES indicator used, with parents influencing motivation in different academic domains to varying extents (Kriegbaum et al., 2015; Tenenbaum & Leaper, 2003). For example, fathers' SES was found to be a strong predictor of math-specific motivational constructs such as self-concept, self-efficacy, and interest.

10.3 Present Study

Against the backdrop of what was described earlier and the still-scarce knowledge about how motivation, students' home backgrounds, and InQ are linked, we examine this link within the example of Norway, using the opportunities provided through the TIMSS 2015 study. Norway was the only Nordic country that decided to include the items measuring InQ as part of their national options section in the TIMSS 2015 Student Questionnaire. Although data are not available for other Nordic countries, we see Norway as a typical representative of the principles exemplified in what is known as the Nordic model (see Chap. 2 in this volume for details). Thus, results from an analysis of the Norwegian data should be considered highly relevant in a broader Nordic perspective.

The research questions of our chapter are the following:

RQ 1: How is a student's SES associated with the intrinsic motivation to learn mathematics in fifth and ninth grade in Norway?

RQ 2: How is InQ associated with the intrinsic motivation to learn mathematics among low-SES, medium-SES, and high-SES student groups in the fifth and ninth grades in Norway?

10.4   Methodology

10.4.1   Data, Sample, and Measurements

Our study is based on achievements and questionnaire data from 4329 fifth-grade students and 4697 ninth-grade students who participated in TIMSS 2015 in Norway. Primary school in Norway encompasses grades 1 through 7 (7 years), while lower secondary school includes grades 8 through 10 (3 years). As already mentioned, Norway is the only Nordic country that measured all InQ dimensions through the national options[1] in TIMSS 2015, but some other countries, such as Germany and Belgium, included the same items for measuring InQ. These measures, based on previous research, were also piloted, and the psychometric properties worked well in Norway, Germany, and Belgium (Bellens, Van Damme, Van Den Noortgate, Wendt, & Nilsen, 2019).

In the present study, we measured InQ through four latent variables: classroom management, teacher support, cognitive activation, and clarity of instruction. These items are presented in Table 10.1.

Mathematics achievement was measured using students' achievement (gauged using five plausible values) on almost 250 mathematics items. These items capture the breadth of the domain as well as the range of cognitive dimensions: knowing, applying, and reasoning (Grønmo, Lindquist, Arora, & Mullis, 2015). The standard deviation for mathematics achievement was set at 100.

SES was measured by students' ratings of their parents' education, number of books at home, and the educational resources available at home. We used the composite variable created with item response theory.[2]

Intrinsic motivation was measured as a latent variable. Students were asked "How much do you agree with these statements about learning mathematics?" They rated items on a Likert scale that ranged from 'Disagree a lot' to 'Disagree a little'. The items included 'I enjoy learning mathematics'; 'I wish I did not have to study mathematics'; 'Mathematics is boring'; 'I learn many interesting things in mathematics'; 'I like mathematics'; 'I like any schoolwork that involves numbers'; 'I like

---

[1] Each country may include some of its own questions on the TIMSS questionnaires. These items, referred to as national options, are not part of the international questionnaire.

[2] See: http://timssandpirls.bc.edu/timss2015/international-results/timss-2015/mathematics/home-environment-support/home-resources-for-learning/

Table 10.1 The Norwegian TIMSS 2015 national option items measuring Instructional Quality (InQ)

Classroom management	Supportive climate	Clarity of instruction	Cognitive activation
Students do not listen to what the teacher says There is noise and disorder Our mathematics teacher has to wait a long time for students to quiet down Students cannot work well Students do not start working for a long time after lessons begin *Our mathematics teacher:* Makes students follow classroom rules for behaviour Calms down students who disrupt the lesson	*Our mathematics teacher:* Shows an interest in every student's learning Provides extra help when students need it Helps students with their learning Continues teaching until the students understand Gives students the opportunity to express opinions	*Our mathematics teacher:* Sets clear learning goals Explains what he or she expects us to learn Asks questions to check whether we understand the content of the lesson Explains at the start of a class how new topics relate to previous lessons Summarizes what we have covered at the end of the lesson	In our mathematics lessons, we are working on tasks that I must think about very thoroughly *Our mathematics teacher:* Asks questions that I must think about very thoroughly Gives us tasks that seem difficult at first glance Asks me what I have understood and what I have not Asks what we know about a new topic Gives us tasks that I like to think about Wants me to be able to explain my answers

to solve mathematics problems'; 'I look forward to mathematics class'; and 'Mathematics is one of my favourite subjects.'

10.4.2 Data Analysis

Three-group, multilevel structural equation models (SEMs) for low-, medium-, and high-SES student groups were estimated for both grade levels. For the cut-off points, the low-SES group included the 25% of students with the lowest SES, the medium-SES group included the 50% of students with medium-SES backgrounds, and the high-SES group comprised the 25% of students with the highest SES.

As students are nested within classes, we employed a two-level model, with students at the within level and classes at the between level.

### 10.4.3 Structural Equation Model (SEM)

SEM is a multivariate statistical analysis technique that includes confirmatory factor analyses (CFA). CFA generates the factor loadings of indicators on an underlying latent factor. Together with the model fit indices, factor loadings

provide a measure of reliability and validity (Byrne, 2012). SEM also allows for examining the relationships between multiple observed and unobserved variables, while providing explicit estimates of error variance parameters. It further enables complex modelling (e.g. multi-group) and complex patterns with intervening variables between the independent and dependent variables; independent variables may also function as dependent variables (Preacher, Zyphur, & Zhang, 2010).

A further great advantage of SEM is the possibility for multilevel approaches in which it is possible to simultaneously model at all levels.

Our main interest lies in the relationship between InQ and motivation at the class level and whether these relationships vary among different groups of students (high-SES, medium-SES, and low-SES). Additionally, we also included the relationship between InQ and motivation at the student level to remove the noise of students' variations in reporting InQ (Lüdtke, Robitzsch, Trautwein, & Kunter, 2009). We further controlled for student achievement at both levels, as shown in Fig. 10.1.

We made one model for each InQ dimension to avoid multi-collinearity. All models were estimated in Mplus Version 8 using the robust maximum likelihood (MLR) estimation. MLR also takes care of the missings (there were 93 missings). Prior to adding any structure, a CFA was conducted to ensure reliable and valid measurement models. The regression coefficients provided in the Results section of this chapter w standardized to allow for comparisons. To evaluate model fit, we referred to common guidelines (CFI $\geq$ 0.95, TLI $\geq$ 0.95, RMSEA $\leq$ 0.08, and SRMR $\leq$ 0.10 for an acceptable model fit; Marsh, Hau, & Grayson, 2005b).
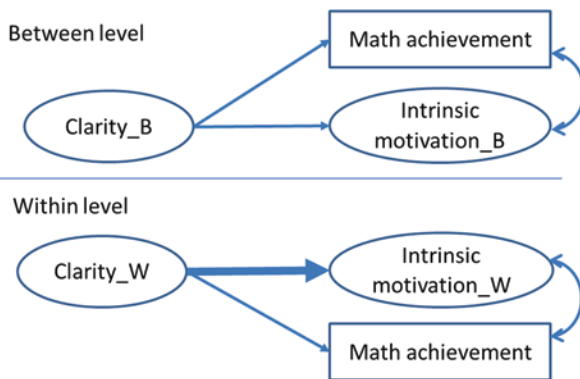
Fig. 10.1 Model of the relationship between aspects of teacher InQ (in this case, clarity of instruction) and student motivation, controlling for mathematics achievement

10.5   Results

In this section, we will present the results of our analyses of (1) the relationship between SES and student motivation, and (2) the relationship between InQ and intrinsic motivation for the three student groups (low-, medium-, and high-SES) at the between level.

First, we investigated whether SES is a predictor of intrinsic motivation to learn mathematics among Norwegian students in the fifth and ninth grades. The model fit was quite high. Our analyses revealed different results for fifth and ninth-graders. In the fifth grade, we found that the relationship between SES and intrinsic motivation was insignificant, but in the ninth grade, the relationship between SES and intrinsic motivation was 0.153 at the between level (standardized regression coefficient) and significant.

Second, we calculated the regression coefficients for the four InQ dimensions on intrinsic motivation to learn mathematics at the between level for both fifth- and ninth-grade students, controlling for achievement. These coefficients are presented in Tables 10.2, 10.3, 10.4 and 10.5.

Table 10.2 Standardized regression coefficients for classroom management on intrinsic mathematics learning motivation, by Socio-Economic Status (SES)

Class level	Low SES	Medium SES	High SES
Grade 5	0.416*	0.17	0.15
Grade 9	0.61*	0.516*	0.502*

Note. Standardized regression coefficients were calculated for classroom management's effects on intrinsic motivation to learn mathematics among the different SES groups at the class level. An * indicates significance at the.05 level

Table 10.3 Standardized regression coefficients for supportive climate on intrinsic mathematics learning motivation, by Socio-Economic Status (SES)

Class level	Low SES	Medium SES	High SES
Grade 5	0.549*	0.35*	0.414
Grade 9	0.843*	0.578*	0.422*

Note. Standardized regression coefficients were calculated for supportive climate's effects on intrinsic motivation to learn mathematics among the different SES groups at the class level. An * indicates significance at the.05 level

Table 10.4 Standardized regression coefficients for clarity of instruction on intrinsic mathematics learning motivation, by Socio-Economic Status (SES)

Class level	Low SES	Medium SES	High SES
Grade 5	0.77*	0.4*	0.46
Grade 9	0.879*	0.719*	0.702*

Note. Standardized regression coefficients were calculated for clarity of instruction's effects on intrinsic motivation to learn mathematics among the different SES groups at the class level. An * indicates significance at the.05 level

Table 10.5 Standardized regression coefficients for cognitive activation on intrinsic mathematics learning motivation, by Socio-Economic Status (SES)

Class level	Low SES	Medium SES	High SES
Grade 5	0.815*	0.427	0.276
Grade 9	0.928*	0.944*	0.885*

Note. Standardized regression coefficients were calculated for cognitive activation's effects on intrinsic motivation to learn mathematics among the different SES groups at the class level. An * indicates significance at the.05 level

As revealed in Table 10.2, the regression coefficient for classroom management on intrinsic motivation to learn mathematics was generally higher for ninth-grade students in comparison to fifth-grade students for all three SES groups. In addition, for both grade levels, the regression coefficient was highest for the low-SES student groups. Furthermore, the regression coefficients for the medium- and high-SES student groups in fifth grade were quite low and insignificant at the .05 level. The regression coefficients for the low-SES student group in fifth grade and the three SES groups in the ninth grade were all significant at the.05 level.

Table 10.3 presents the corresponding regression coefficients for the dimension supportive climate. As can be seen in the diagram, the overall picture was quite similar to the preceding one. First, the regression coefficients are generally somewhat higher for ninth grade (G9) than for fifth grade (G5). Second, the regression coefficients have a declining tendency from low-SES, via medium-SES, to high-SES student groups and are particularly high for low-SES students within each grade level. The regression coefficient for high-SES students in fifth grade is insignificant.

Table 10.4 gives the regression coefficients for the dimension clarity of instruction on students' intrinsic motivation to learn mathematics in relation to student SES groups. The same pattern as the previously presented dimensions can be seen. The regression coefficients are particularly high for the low-SES student groups in both grades, and a declining tendency exists from low-SES via medium-SES to high-SES student groups. This tendency is more distinct in the fifth grade than in the ninth.

In Table 10.5, the regression coefficients for the dimension cognitive activation on intrinsic motivation are presented. The regression coefficients are extremely high for the ninth-grade student SES groups, and they are also quite high for the low-SES student group in the fifth grade. As for the medium- and high-SES student groups in the fifth grade, the regression coefficients were insignificant at the.05 level. In fifth grade, the difference between the regression coefficients for low-SES students and medium/high-SES students is considerable, but this is not the case for the ninth-grade students.

10.6   Discussion

The discussion of our findings will be done in relation to our research questions. Our first research question addresses how a student's SES is associated with their intrinsic motivation to learn mathematics in the fifth and ninth grades in Norway.

In our analyses, we found a significant association between students' SES and intrinsic motivation among ninth-grade students, but not among fifth-grade students. With ninth-graders, this association had a standardized regression coefficient of 0.153 and was significant at the .05 level. Even though this association is not very strong, these results clearly indicate that SES is more strongly associated with students' intrinsic motivation in lower secondary school than in primary school. We know from previous research that intrinsic motivation is generally quite high in primary school and considerably lower in secondary school (Fauth et al., 2014; Mullis et al., 2016). This goes for most nations participating in TIMSS and is also reported in the Norwegian 2015 TIMSS report (Bergem, Kaarstein, & Nilsen, 2016a). Interpreting our findings in light of this well-established knowledge, we can conclude that all students in the fifth grade in Norway, regardless of family background (SES), enjoy a relatively high intrinsic motivation to learn mathematics. However, this seems to change during the period between the fifth and ninth grades. When students are in ninth grade, their intrinsic motivation to learn mathematics is not only substantially lower than in elementary school but is also significantly associated with family background (SES). *Why is this so? Why does family background predict students' intrinsic motivation to learn mathematics in the ninth grade, but not in the fifth grade?* We would like to point out a few factors that seem relevant in trying to interpret these findings. First, in an international context, Norwegian classrooms are rather heterogeneous in terms of both SES and achievement. There is no streaming in either elementary or lower secondary school. However, marks are introduced in eighth grade, so this makes a difference between fifth-grade and ninth-grade students. Several international studies have reported a positive relationship between intrinsic motivation and marks (e.g. Corpus et al., 2009; Gottfried, 1990). Another robust finding in international studies is the positive correlation between students' SES and achievement in all countries (Mullis et al., 2016; OECD, 2016). Therefore, the introduction of marks between fifth and ninth grade in Norway may positively influence the correlation between students' SES and their intrinsic motivation to learn mathematics in the ninth grade as compared to the fifth grade, and ninth-grade low-SES students may lose their intrinsic motivation to a greater extent after receiving lower marks than high-SES students.

Second, both Eccles's EVT model and Bandura's socio-cognitive perspective accentuate the important role of parents in socializing their children (Bandura, 2012; Bandura et al., 2001; Eccles, 2009; Eccles et al., 1983). A key element of this process is influencing and shaping children's interest in learning. It has been noted that families' value systems, which are linked closely to family SES, are fundamental in these processes. Taken together with Bourdieu's (1986) reproduction theory, there are reasons to assume that the importance of a family value system that

stimulates and encourages academic work and perseverance and positively evaluates the effort that children put into their schoolwork will increase from elementary to secondary school, in line with the higher demands that students face as they enter higher grades. However, such value systems characterize high-SES families to a larger extent than low-SES families (Bandura, 2012; Bourdieu, 1986; Eccles, 2009) and, therefore, can be assumed to affect the correlation between students' SES and their intrinsic motivation, making it higher in lower secondary school than in primary school.

Through the formulation of our second research question, we set out to investigate whether Norwegian schools can possibly influence and counteract the unwanted trajectory of students' intrinsic motivation to learn mathematics over the school years, with family SES being more important for this motivation in ninth grade than in fifth grade. We did this by examining the association between the four InQ dimensions and students' intrinsic motivation to learn mathematics for low-SES, medium-SES, and high-SES student groups in these two grades. Although our analyses of these relationships between InQ, intrinsic motivation, and SES revealed some similarities of the fifth and ninth grades, some distinct differences were also found. In the following, these traits will be presented and discussed.

First, for the high-SES students in the fifth grade, none of the calculated regression coefficients was found to be significant. As mentioned above, at this grade level, we know that high-SES students' level of intrinsic motivation to learn mathematics is quite high in Norway (Kaarstein & Nilsen, 2016). Our analyses indicated that InQ is not a decisive factor in determining these levels. However, for fifth-grade low-SES students, we found a positive relationship between intrinsic motivation to learn mathematics and each of the measured InQ aspects. This finding will be further elaborated on later.

Second, as seen in Tables 10.2, 10.3, 10.4 and 10.5, the regression coefficients for the four InQ dimensions on intrinsic motivation to learn mathematics were generally higher for ninth-grade students than those in the fifth grade. Our interpretation of this finding is the following: *High InQ seems to be particularly important for strengthening and consolidating the intrinsic motivation to learn mathematics in lower secondary school and much more so than in primary school.*

Third, and perhaps most importantly, the regression coefficients for the InQ dimensions on intrinsic motivation to learn mathematics are highest for low-SES students, and with a few exceptions, as related to the medium-SES group, are lowest for high-SES students in both fifth and ninth grade. This goes for all four dimensions. We interpret this finding as follows: *High InQ is particularly important for low-SES students in both the fifth and ninth grades in relation to strengthening and consolidating their intrinsic motivation to learn mathematics.* This means that in both elementary and lower secondary school, a teacher's InQ can contribute to higher levels of equity. In other words, to provide equitable opportunities for all students to succeed in mathematics, which is a prominent aim in Norway's educational system (Kunnskapsdepartementet, 2006), it is particularly important to ensure that low-SES students receive high-quality mathematics instruction. If teacher education contributed to enhancing the InQ of teachers, it would boost both high- and

low-SES students, but according to our results, low-SES students would benefit more from such circumstances. More high-quality teachers would thus result in reducing the gap between low- and high-SES students. It would also reduce the gap between schools, as classes in low-SES schools would benefit more from such teachers. This is also in line with previous studies (Gustafsson et al., 2016).

This last finding corresponds well with reports from other studies related to associations between InQ and student outcomes. Using achievement as the outcome measure, Baumert et al. (2010) found that differences in teachers' pedagogical content knowledge in mathematics, mediated mainly by levels of cognitive activation and learning support (supportive climate), made the greatest impact in low-SES classes. Other studies have also reported positive associations between InQ dimensions and student outcome measures, but this mainly entails measures of student achievement, not the aspect of motivation (Bergem, Nilsen, & Scherer, 2016b; Rjosk et al., 2014; Willms, 2010).

10.7   Limitations and Future Research

We would like to point out a few limitations to the conclusions that can be drawn from our study. Our data set was taken from a study with a cross-sectional design; thus, no causal inferences should be drawn. In addition, as only TIMSS data from Norway have been used, we do not know whether our findings could be repeated with data sets from other nations. As the decline in students' intrinsic motivation to learn mathematics from primary to lower secondary school is an international phenomenon, using our study design on data sets from other countries would make for highly interesting research. Furthermore, associations between dimensions of the InQ construct, SES, and intrinsic motivation are investigated in relation to only one subject: mathematics. It remains to be seen whether the current findings could be replicated in other subject areas. While the current analyses focus on fifth and ninth-grade students, further investigation is required to determine whether the same relationships hold in other age groups.

To strengthen the claims made in the current study, the aforementioned limitations could be addressed in future research. Our study design would then need to be copied using representative data sets from other nations and analysed for other age groups and subjects. Most importantly, if our research design were used in longitudinal studies, more robust inferences could be drawn. One cannot draw causal inferences from cross-sectional data, which only capture a moment in time, and inferences made from cross-sectional data may be invalid due to challenges related to omitted variables and reversed causality (Gustafsson, 2013). Longitudinal data reduce such risks. With longitudinal studies, it would be possible to investigate, for instance, whether InQ is related to *changes* in student outcomes. Examples of such studies include longitudinal extensions of PISA (e.g., Krauss, Baumert, & Blum, 2008) and TIMSS with additional classroom observations (Nilsen, 2019).

In TIMSS 2019, more emphasis is put on InQ, and more extensive scales that measure different InQ dimensions are included on the student questionnaire. This

will allow for better InQ validity, and countries need not include national options to measure this construct anymore.

10.8   Implications

Few studies have investigated the relationships between InQ and equity in Nordic countries (Nilsen, Scherer, & Blömeke, 2018). Therefore, the present study's findings will extend knowledge about relations between InQ, SES, and key student outcome measures that are particularly pertinent in school equity debates.

Our main findings indicate that high InQ is more important for stimulating and maintaining students' intrinsic motivation in ninth grade than in fifth grade, and it is especially critical for low-SES student groups, regardless of grade level. These findings should be highly relevant within various strands of educational research, including mathematics education, teacher education, and the field of educational equity. It seems particularly important that teachers and teacher students get introduced to results from research that indicate a close association between students' SES and the development of their intrinsic motivation to learn mathematics, as well as high InQ's key role in compensating for this association. A comprehensive understanding of these issues may motivate teachers to prioritize aspects of their teaching to ensure that all children, regardless of their SES, can tap into their potential and succeed in school to a greater extent. This is also in line with Espinoza (2007), whose understanding of equity is the distribution of resources according to students' needs (see Chap. 2). In our case, resources refers to high InQ.

Additionally, but closely related to the above argument, our findings could be used to inform discussions about education on a policy level. Our findings provide evidence in support of those educational policies that aim to recruit well-qualified teachers who can implement high-quality instruction in their classrooms. This can be done by prioritizing advanced teacher education and high-quality professional development courses. Our findings suggest that such measures not only have the potential to counteract declining intrinsic motivation to learn mathematics in lower secondary school but would also be highly relevant for addressing one of the most important issues in education in Norway at all levels: providing equitable opportunities for *all* students to succeed in school.

References

Bandura, A. (2012). On the functional properties of perceived self-efficacy revisited. *Journal of Management, 38*(1), 9–44. https://doi.org/10.1177/0149206311410606

Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (2001). Self-efficacy beliefs as shapers of children's aspirations and career trajectories. *Child Development, 72*(1), 187–206. https://doi.org/10.1111/1467-8624.00273

Barry, K., & King, L. (2000). *Beginning teaching and beyond* (3rd ed.). Wentworth Falls, NSW: Social Science Press.

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., et al. (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 2s015 data in Flanders (Belgium), Germany, and Norway. *Large-Scale Assessments in Education, 7*(1). https://doi.org/10.1186/s40536-019-0069-2

Bergem, O. K., Kaarstein, H., & Nilsen, T. (2016a). *Vi kan lykkes i realfag. (We Can Succeed in Mathematics and Science)*. Norwegian national TIMSS 2015 report. Universitetsforlaget.

Bergem, O. K., Nilsen, T., & Scherer, R. (2016b). Undervisningskvalitet i matematikk (instructional quality in mathematics). In O. K. Bergem, H. Kaarstein, & T. Nilsen, *Vi kan lykkes i realfag. (We Can Succeed in Mathematics and Science)* (Norwegian national TIMSS 2015 report). Universitetsforlaget.

Bernstein, P., Bernstein, B., & MacRae, D. G. (1971). *Class, codes and control: Theoretical studies towards a sociology of language* (Vol. 1). London: Routledge & K. Paul.

Boudon, R. (1974). *Education, opportunity, and social inequality: Changing prospects in Western society*. New York: Wiley.

Bourdieu, P. (1986). The forms of capital. In J. G. Richardson (Ed.), *Handbook of theory and research for the sociology of education* (pp. 241–258). Connecticut, CT: Greenwood Press.

Brophy, J., & Good, T. L. (1986). Teacher behavior and student achievement. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 328–375). New York: McMillan.

Byrne, B. M. (2012). *Structural equation modeling with mplus: Basic concepts, applications, and programming*. New York: Routledge.

Cheung, D. (2018). The key factors affecting students' individual interest in school science lessons. *International Journal of Science Education, 40*(1), 1–23. https://doi.org/10.1080/09500693.2017.1362711

Corpus, J. H., McClintic-Gilbert, M. S., & Hayenga, A. O. (2009). Within-year changes in children's intrinsic and extrinsic motivational orientations: Contextual predictors and academic outcomes. *Contemporary Educational Psychology, 34*(2), 154–166. https://doi.org/10.1016/j.cedpsych.2009.01.001

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

Dorfner, T., Förtsch, C., & Neuhaus, B. J. (2018). Effects of three basic dimensions of instructional quality on students' situational interest in sixth-grade biology instruction. *Learning and Instruction, 56*, 42–53. https://doi.org/10.1016/j.learninstruc.2018.03.001

Duit, R. (2009). Bibliography—Students and teachers' conceptions and science education (STCSE). http://www.ipn.uni-kiel.de/aktuell/stcse/stcse.html

Eccles, J. (2009). Who am I and what am I going to do with my life? Personal and collective identities as motivators of action. *Educational Psychologist, 44*(2), 78–89.

Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., et al. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives* (pp. 75–146). San Francisco: Freeman.

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

Espinoza, O. (2007). Solving the equity–equality conceptual dilemma: A new model for analysis of the educational process. *Educational Research, 49*(4), 343–363.

Farrington, C. A., Roderick, M., Allensworth, E., Nagaoka, J., Keyes, T. S., Johnson, D. W., & Beechum, N. O. (2012). *Teaching adolescents to become learners: The role of noncognitive factors in shaping school performance—A critical literature review*. Consortium on Chicago School Research.

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.

Field, S., Kuczera, M., & Pont, B. (2007). *No more failures: Ten steps to equity in education*. Paris: OECD.

Good, T. L., & Brophy, J. E. (2000). *Looking in classrooms* (8th ed.). New York: Longman.

Gottfried, A. E. (1990). Academic intrinsic motivation in young elementary school children. *Journal of Educational Psychology, 82*(3), 525–538. https://doi.org/10.1037/0022-0663.82.3.525

Gottfried, A. E., Fleming, J. S., & Gotfried, A. W. (2001). Continuity of academic intrinsic motivation from childhood through late adolescence: A longitudinal study. *Journal of Educational Psychology, 93*, 3–13.

Grønmo, L. S., Lindquist, M., Arora, A., & Mullis, I. V. (2015). TIMSS 2015 mathematics framework. In I. V. Mullis & M. O. Martin (Eds.), *TIMSS 2015 Assessment Framework* (pp. 11–27). Chestnut Hill, MA: TIMSS & PIRLS International Study Center.

Gustafsson, J. E. (2013). Causal inference in educational effectiveness research: A comparison of three methods to investigate effects of homework on student achievement. *School Effectiveness and School Improvement, 24*(3), 275–295.

Gustafsson, J.-E., Nilsen, T., & Hansen, K. Y. (2016). School characteristics moderating the relation between student socioeconomic status and mathematics achievement in grade 8. Evidence from 50 countries in TIMSS 2011. *Studies in Educational Evaluation, 57*, 16–30.

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

Häussler, P., & Hoffmann, L. (2000). A curricular frame for physics education: Development, comparison with students' interests, and impact on students' achievement and self-concept. *Science Education, 84*, 689–705.

Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. Lester (Ed.), *Second Handbook of research on mathematics teaching and learning* (pp. 371–404). Charlotte, NC: Information Age.

Jansen, M., Schroeders, U., & Lüdtke, O. (2014). Academic self-concept in science: Multidimensionality, relations to achievement measures, and gender differences. *Learning and Individual Differences, 30*, 11–21. https://doi.org/10.1016/j.lindif.2013.12.003

Kaarstein, H., & Nilsen, T. (2016). Motivasjon (Motivation). In O. K. Bergem, H. Kaarstein, & T. Nilsen (Eds.), *Vi kan lykkes i realfag (We Can Succeed in Mathematics and Science)*. Norwegian national TIMSS 2015 report. Universitetsforlaget.

Kane, T., & Cantrell, S. (2010). *Learning about teaching: Initial findings from the measures of effective teaching project* (MET Project Research Paper, Bill & Melinda Gates Foundation, 9).

Klette, K. (2015). Introduction. In K. Klette, O. K. Bergem, & A. Roe (Eds.), *Teaching and learning in lower secondary schools in the era of PISA and TIMSS*. Cham, Switzerland: Springer Publishing Company.

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

Korbel, V., & Paulus, M. (2018). Do teaching practices impact socio-emotional skills? *Education Economics, 26*(4), 337–355.

Krauss, S., Baumert, J., & Blum, W. (2008). Secondary mathematics teachers' pedagogical content knowledge and content knowledge: Validation of the COACTIV constructs. *ZDM, 40*(5), 873–892.

Kriegbaum, K., Jansen, M., & Spinath, B. (2015). Motivation: A predictor of PISA's mathematical competence beyond intelligence and prior test achievement. *Learning and Individual Differences, 43*, 140–148.

Kunnskapsdepartementet. (2006). … og ingen sto igjen. Tidlig innsats for livslang læring. https://www.regjeringen.no/no/dokumenter/stmeld-nr-16-2006-2007-/id441395/

Kunter, M., Baumert, J., & Köller, O. (2007). Effective classroom management and the development of subject-related interest. *Learning and Instruction, 17*(5), 494–509.

Kyriakides, L., & Creemers, B. P. (2011). Can schools achieve both quality and equity? Investigating the two dimensions of educational effectiveness. *Journal of Education for Students Placed at Risk, 16*(4), 237–254.

Kyriakides, L., Creemers, B. P., & Charalambous, E. (2019). Searching for differential teacher and school effectiveness in terms of student socioeconomic status and gender: Implications for promoting equity. *School Effectiveness and School Improvement, 30*(3), 286–308.

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 modeling. *Contemporary Educational Psychology, 34*(2), 120–131. https://doi.org/10.1016/j.cedpsych.2008.12.001

Marsh, H., Craven, R., & Debus, R. (1999). Separation of competency and affect components of multiple dimensions of academic self-concept: A developmental perspective. *Merrill Palmer Quarterly, 45*(4), 567–601.

Marsh, H. W., Hau, K., & Grayson, D. (2005b). Goodness of fit evaluation in structural equation modeling. In A. Maydeu-Olivares & J. McArdle (Eds.), *Mahwah contemporary psychometrics* (pp. 275–340). Mahwah, NJ: Erlbaum.

Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005a). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development, 76*(2), 397–416.

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

Musu-Gillette, L. E., Wigfield, A., Harring, J. R., & Eccles, J. S. (2015). Trajectories of growth in students' expectancies and values in math and college major choice. *Educational Research and Evaluation, 21*(4), 343–370.

Nilsen, T. (2019). *Conceptualizations of instructional quality in the longitudinal project TESO, and previous findings from TIMSS* [Paper presentation]. Symposium at the QUINT conference 2019: Analyzing Teaching Quality: Perspectives, Potential, and Pitfalls, CITY, STATE.

Nilsen, T., Scherer, R., & Blömeke, S. (2018). The relation of science teachers' quality and instruction to student motivation and achievement in the fourth and eighth grade: A Nordic perspective. In Nordic Council of Ministers (Eds.), *Northern Lights on TIMSS and PISA 2018* (pp. 61–94). Nordic Council of Ministers.

Opheim, V. (2004). *Equity in education: Country analytical report Norway* (NIFU STEP, rapport 7).

Organisation for Economic Cooperation and Development. (2012). *Equity and quality in education: Supporting disadvantaged students and schools*. OECD Publishing. https://doi.org/10.1787/9789264130852-en

Organisation for Economic Cooperation and Development. (2016). PISA 2015 results (Vol. 1): Excellence and equity in education. OECD Publications.

Organisation for Economic Cooperation and Development. (2018). Equity in education: Breaking down barriers to social mobility. OECD Publications.

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.

Pianta, R. C., Nimetz, S. L., & Bennet, E. (1997). Mother–child relationships, teacher–child relationships and adjustment in preschool and kindergarten. *Early Childhood Research Quarterly, 12*(3), 263–280.

Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods, 15*(3), 209–233. https://doi.org/10.1037/a0020141

Rjosk, C., et al. (2014). Socioeconomic and language minority classroom composition and individual reading achievement: The mediating role of instructional quality. *Learning and Instruction, 32*, 63–72.

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

Ryan, R. M., & Deci, E. L. (2002). An overview of self-determination theory: An organismic-dialectical perspective. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of self-determination research* (pp. 3–33). Rochester, NY: University of Rochester Press.

Ryan, R. M., & Deci, E. L. (2009). Promoting self-determined school engagement. Motivation, learning, and well-being. In K. R. Wentzel & D. B. Miele (Eds.), *Handbook of motivation at school* (pp. 171–195). New York: Routledge.

Scheerens, J. (2014). School, teaching, and system effectiveness: Some comments on three state-of-the-art reviews. *School Effectiveness and School Improvement, 25*(2), 282–290. https://doi.org/10.1080/09243453.2014.885453

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*, 1550.

Scherer, R., & Nilsen, T. (2016). The relations among school climate, instructional quality, and achievement motivation in mathematics. In T. Nilsen & J.-E. Gustafsson (Eds.), *Teacher quality, instructional quality and student outcome: Relationships across countries, cohorts, and time* (pp. 51–80). Cham, Switzerland: Springer.

Schmidt, W. H., Burroughs, N. A., Zoido, P., & Houang, R. T. (2015). The role of schooling in perpetuating educational inequality: An international perspective. *Educational Researcher, 44*(7), 371–386.

Seidel, T., Rimmele, R., & Prenzel, M. (2005). Clarity and coherence of lesson goals as a scaffold for student learning. *Learning and Instruction, 15*(6), 539–556.

Seidel, T., & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research, 77*(4), 454–499. https://doi.org/10.3102/0034654307310317

Simpkins, S. D., Davis-Kean, P. E., & Eccles, J. S. (2006). Math and science motivation: A longitudinal examination of the links between choices and beliefs. *Developmental Psychology, 42*, 70–83. https://doi.org/10.1037/0012-1649.42.1.70

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

Spinath, B., & Spinath, F. M. (2015). Longitudinal analysis of the link between learning motivation and competence beliefs among elementary school children. *Learning and Instruction, 15*(2), 87–102.

Spinath, B., & Steinmayr, R. (2012). The roles of competence beliefs and goal orientations for change in intrinsic motivation. *Journal of Educational Psychology, 104*(4), 1135–1148.

Steinmayr, R., & Spinath, B. (2009). The importance of motivation as a predictor of school achievement. *Learning and Individual Differences, 19*(1), 80–90.

Tenenbaum, H. R., & Leaper, C. (2003). Parent–child conversations about science: The socialization of gender inequities? *Developmental Psychology, 39*(1), 34–47.

Viljaranta, J., Tolvanen, A., Aunola, K., & Nurmi, J.-E. (2014). The developmental dynamics between interest, self-concept of ability, and academic performance. *Scandinavian Journal of Educational Research, 58*, 734–756.

Wagner, W., Göllner, R., Werth, S., Voss, T., Schmitz, B., & Trautwein, U. (2015). Student and teacher ratings of instructional quality: Consistency of ratings over time, agreement, and predictive power. *Journal of Educational Psychology., 108*, 705–721. https://doi.org/10.1037/edu0000075

Watt, H. M. G., & Eccles, J. S. (Eds.). (2008). *Gender and occupational outcomes: Longitudinal assessments of individual, social, and cultural influences*. American Psychological Association. https://doi.org/10.1037/11706-000

Weidinger, A. F., Steinmayr, R., & Spinath, B. (2017). Math grades and intrinsic motivation in elementary school: A longitudinal investigation of their association. *British Journal of Educational Psychology, 87*(2), 187–204.

Wigfield, A., Eccles, J. S., Fredricks, J., Simpkins, S., Roeser, R., & Schiefele, U. (2015). Development of achievement motivation and engagement. In M. Lamb & C. Garcia Coll (Eds.), *Handbook of Child Psychology and Development Science* (Vol. 3, 7th ed., pp. 657–700). Hoboken, NJ: Wiley.

Willms, J. D. (2010). School composition and contextual effects on student outcomes. *Teachers College Record, 112*(4), 1008–1037.

Chapter 11
Resilient and Nonresilient Students in Sweden and Norway—Investigating the Interplay Between Their Self-Beliefs and the School Environment

Jelena Radišić and Andreas Pettersen

Abstract Using TIMSS 2015 data and a person-centred approach, the chapter focuses on academically resilient students in Norway and Sweden in grade eight. The self-belief profiles of academically resilient students compared with the nonresilient groups (i.e., low SES/low achievement, high SES/low achievement and high SES/high achievement) are investigated. Further, we evaluated the characteristics of the classroom environment for each of the profiles. After accounting for student SES and achievement, personal characteristics, advantages and disadvantages in the classroom and the school environment, we identified distinctive student profiles that might be more prone to risk. In the context of the equality–inequality paradigm, recognition of these profiles can strengthen the possibility to reduce the gap in battling different aspects of inequality across social groups. Concurrently, although we distinguish the same student groups across Sweden and Norway, their distribution within the countries differs. The latter results contribute to the ongoing debate on the dissolution/unification of the Nordic model, especially regarding particular trends within the Swedish education system.

Keywords TIMSS · Students' self-beliefs · Students at risk · School environment

A strong relationship between students' socio-economic background and school achievement has been reported in various cases (Nilsen, Blömeke, Hansen, & Gustafsson, 2016). Research has shown that across different educational systems,

J. Radišić (✉) · A. Pettersen
Department of Teacher Education and School Research, University of Oslo, Oslo, Norway
e-mail: Jelena.radisic@ils.uio.no

students from families with more social and economic resources (SES) may have higher chances of succeeding in school (e.g., Nilsen et al., 2016; OECD, 2018; Xie & Ma, 2019). Similarly, students from the lower socioeconomic spectrum are more likely to perform poorly at school and have less of a chance to complete secondary and tertiary education (see Reardon, 2011). This strong relationship between student background and school achievement implies that educational systems may not be equally ensuring the success of every child (Doll, 2013; Pianta & Walsh, 1998).

However, despite the reports of a consistent relationship between SES and achievement, many students with socioeconomically disadvantaged backgrounds still succeed in school (Masten, 2014). Some are even among the top-performing students in their schools. These students are commonly labelled as academically resilient because they are successful in school despite being situated in an environment linked to poorer outcomes (Martin & Marsh, 2006). Trying to understand more about resilient students and what might have contributed to their success is regarded as critical to both educators and policy makers. Indeed, targeted actions and support could ensure that more students can succeed in school, equipping them with the different tools needed to obtain positive outcomes. Also, such analyses can provide valuable insights of the possible differentiation between resilient and nonresilient groups because the latter may include students who face different challenges and difficulties and who often lack support in battling adverse outcomes.

The current study adds to this field by focusing on the distinct differences between resilient students (i.e., high-achieving students with low SES) and nonresilient student groups (i.e., low-achieving students irrespective of SES and successful students with high SES) in connection to their self-beliefs related to mathematics (i.e., confidence, interest and value) and sense of school belonging. Our primary assumption rests on the premise of a person-centred approach of the heterogeneity of the student population, which is often overlooked when we observe the relationship between different variables alone (Bergman & Trost, 2006). We take this idea a step further by examining whether distinct belief patterns can be extracted across resilient and nonresilient student groups or whether some belief patterns can be regarded as unique to a particular student category (e.g., resilient students). Furthermore, we investigate the school and classroom environment of students with distinct self-belief patterns and how these relate back to students initially categorised as resilient or not. To address these issues, we utilise the 2015 data from the Trends in International Mathematics and Science Study (TIMSS) for students in grade eight in Norway and Sweden. The multiple-group analyses within a broader person-centred perspective will enable us to observe different nuances across the resilient and nonresilient student groups in both countries (Morin, Meyer, Creusier, & Biétry, 2016).

11.1 Academic Resilience as a Mirror of Systems' Inequality

Although earlier definitions of resilience have focused more on observing resilience as an individual characteristic (Masten, 2018), the literature has slowly moved in the direction of the view that resilience can originate from factors external to the

individual. This idea allows for converging towards the so-called risk and protective paradigm, which explores those influences that can predict resilience (e.g., family or school environment; Abelev, 2009; Franklin, 2000; Masten, 2014, 2018; Rutter, 2006). Over time, the complexity of the lenses used to understand resilience has only increased, introducing the ecological view (Bronfenbrenner & Morris, 2006), developmental-systemic perspective (Masten, 2018) or view of resilience as a dynamic process of interaction between the contexts and the individual's agency (Hernandez-Martinez & Williams, 2013). At the same time, these multiple lenses have not aided in finding a more unified operationalisation of resilience (Sattler & Gershoff, 2019). Despite this, the construct tends to travel across disciplines, especially in the social sciences and humanities (Masten, 2018).

Within the educational milieu, resilience is often discussed in the context of academic resilience or educational resilience. As such, it is described as '*heightened likelihood of success in school* (…) *despite environmental adversities*' (Wang, Haertel, & Walberg, 1994, p. 46). Hence, academically resilient students are those who succeed in school even though they are subjected to unfavourable surroundings or despite having a disadvantaged background. According to Sattler and Gershoff (2019), the different definitions of resilience can be divided into two main categories. The first focuses on '*the processes between risk and protective factors in promoting or hindering positive adjustment*', while the latter focuses on '*the criteria used for judging competence following adversity*' (Ibidem, 2019, p. 88). Furthermore, two different criteria are used in distinguishing resilient from the nonresilient students. These include (1) doing better than peers experiencing similar risks (low-threshold resilience) and (2) doing as well as peers not experiencing risk (high-threshold resilience).

Over the past few years, constructs similar to academic resilience have emerged. One such example is academic buoyancy, which refers to students' ability to deal with everyday setbacks, challenges and pressures, such as exam pressure, low grades and difficulties related to schoolwork (Martin & Marsh, 2008). Thus, although resilience is related to more 'chronic' adversities, buoyancy is related to 'everyday hassles and coping' (Martin & Marsh, 2008). Independent of its conceptualisation, academic resilience can be seen as a by-product of inequality in the school system. Per the definition, resilient students are subjected to inequality because they come from socially, culturally or economically disadvantaged milieus. Thus, resilience is often viewed in light of and related to the concepts of educational (in)equality and equity (e.g., OECD, 2018; Reyes, Elias, Parker, & Rosenblatt, 2013). Across the literature, the terms are frequently used in both a similar and dissimilar fashion, provoking some disagreements regarding its meaning (Espinoza, 2007). According to Espinoza (2007), equity is usually '*associated with fairness or justice in the provision of education*', while equality is associated with *sameness in treatment*. Rather than striving for a unique and straightforward conception of equity and equality in education, Espinoza (2007) argues that multiple definitions are needed. His 'equality–equity model' aims to clarify and differentiate educational equity and equality regarding the different stages and features of the educational process, allowing for a broader perspective of the equity–equality continuum. Similarly, academic resilience can also be related to the different stages and features

of the educational process and, thus, can be associated with different conceptions and definitions of equity and equality.

Data from the international large-scale assessment studies (ILSA), such as TIMSS and the Programme in International Student Assessment (PISA), are often used to investigate equity in education from a variable-centred approach. In these, students' gender, socioeconomic status, immigrant background or school characteristics are related to achievement as a measure of equity (e.g., Agasisti, Avvisati, Borgonovi, & Longobardi, 2018; Erberber, Stephens, Mamedova, Ferguson, & Kroeger, 2015; Nilsen et al., 2016; OECD, 2018; Zhu, 2018). From this perspective, more equitable school systems are those in which less variance in student outcomes can be attributed to their background.

In recent years, using data from the ILSAs researchers have also focused on investigating resilience in countries around the world. These investigations provide understanding of the factors that underlie students' success across sometimes very different systems (e.g., Erberber et al., 2015; OECD, 2018). In these studies, academically resilient students are defined in a somewhat different fashion; they are viewed as scoring above a certain threshold in achievement and below the specified limit on students' background measures (i.e., related to the social, economic and cultural resources of the students' homes) and are thus very much aligned with the earlier described categorisation by Sattler and Gershoff (2019). However, the criteria for judging competence within ILSAs can take both a national and international perspective. Although both views have their merits, they do produce different population draws distinguishing between resilient or nonresilient students (OECD, 2018). In the current study, the former is chosen under the assumption that it provides a more comparative approach across different reference groups within the national systems we observe.

11.2   Factors Linked to Resilience in Mathematics

With the notion that academic resilience occurs at the crossroads of the individual, family and school (Doll, 2013), numerous studies have tried to map out both student and school characteristics that may support resilience. At the student level, positive student attitude towards mathematics, confidence, high self-esteem, commitment and sense of control are reported to endorse resilience (Martin & Marsh, 2006; Sandoval-Hernández & Białowolski, 2016; Wayman, 2002). Similarly, Kalender (2015) argues that resilient students have mostly positive attitudes towards school and their teachers compared with low-achieving students. The latter, in comparison, perceive that they could not be successful even if they tried. Also, resilient students show confidence in using their resources and ask for help when it is needed. They establish and preserve positive relationships with their teachers and peers (Eisenberg et al., 2003; Lessard, Butler-Kisber, Fortin, & Marcotte, 2014). The quality and nature of these relationships are consistently reported among the essential protective elements needed to succeed (Doll, Zucker, & Brehm, 2004). At the same time,

teacher confidence in student performance was also associated with higher chances of academic success (Sandoval-Hernández & Białowolski, 2016). Yet although the comparisons established across the studies are usually between resilient students and low achievers, they do not differentiate between low achievers from the higher and lower SES bands. Thus, because particular constituents are attributed to the resilient students compared with nonresilient ones, further exploration is needed to explore these student categories regarding SES.

In a position paper by Ungar, Connelly, Liebenberg, and Theron (2019), access to supportive relationships, experiences of social cohesion with others and access to material resources are listed as examples of what schools can do to support student resilience. These school characteristics are sustained even when student background is controlled for (for details, see Borman & Dowling, 2010; Perry & McConney, 2010; Wiberg, 2019). School climate has been steadily seen as a school feature that fosters the conditions for optimal learning environments, which lead to positive student outcomes (Kyriakides, Creemers, Antoniou, & Demetriou, 2010; Maxwell, Reynolds, Lee, Subasic, & Bromhead, 2017). Among its key aspects are the school's emphasis on academic success (Hoy, Tarter, & Hoy, 2006) and a safe and orderly climate (Wang & Degol, 2016). Although both constructs are mutually connected (Hoy et al., 2006; Thapa, Cohen, Guffey, & Higgins-D'Alessandro, 2013), they are also linked to students' outcomes and engagement (Martin, Foy, Mullis, & O'Dwyer, 2013; O'Brennan & Furlong, 2010; Wang & Degol, 2016). However, some authors do caution on the need for exploring these constructs' differential effects across low- and high-SES schools (Lee & Smith, 1999) and countries (Sandoval-Hernández & Białowolski, 2016).

ILSAs also contribute with some crucial insights. Using the TIMSS 2011 data, Erberber et al. (2015) observe the factors associated with resilience in 28 education systems participating in TIMSS. They find that students' educational aspirations, valuing of mathematics and experiencing less frequent bullying emerged as predictors of resilience in several education systems, coupled with students' beliefs about their teachers' confidence in their abilities. At the school level, across the board, schools' emphasis on academic success and schools having a lower percentage of economically disadvantaged students could be linked to resilience (Erberber et al., 2015). The authors, however, conclude that despite some similarities in their cross-country analyses, there is no universal recipe that could be applied to all the 28 examined cases. Similar results are obtained in PISA 2015 in connection to the field of science. Here, the school socioeconomic profile and the disciplinary climate in school are the two school factors most frequently associated with resilience in the national context, and students' motivation to achieve the best they can was the most critical student factor linked to it (OECD, 2018). Although one of the significant affordances of the ILSA data lies in its opportunity for cross-country comparisons, one can argue for a more focused approach in the selection of the countries involved in such analyses. Thus, although no universal recipes are found when examining particular relationships (e.g., Erberber et al., 2015), a more focused choice can be a first step.

11.3   Provision of Education in Norway and Sweden

After World War II, both Norway and Sweden—as well as other Nordic countries—
saw significant advances in the introduction of comprehensive school systems. This
system allowed all children and young people to be enrolled in a standard structure
across different stages in the educational system (Telhaug, Mediås, & Aasen, 2006).
The policy differed from some other Western countries, such as Germany, France,
the Netherlands or the UK. In the early days, equality of education for all students
rarely extended beyond ages 10 or 11, corresponding to grades four and five within
the system of compulsory public schooling. At the same time, students from the
upper social classes often did not attend the state schools (Telhaug et al., 2006).

According to Blossing, Imsen, and Moos (2014), the Nordic model of education
has been, historically speaking, '*based on a vision that schools should be inclusive,
comprehensive, with no streaming and with easy passages between the levels'.*
(p. 1). Also, they argue that in the Nordic countries '*school* (…) *was considered to
be an extension of the state's duty to provide equality of opportunity for all members
of society* (…) *regardless of social background, abilities, gender and place of living'*
(Blossing et al., 2014, p. 1). Such vision has provided a chance for all students to
develop their potentials and aims, given the goal to supply all with the same quality
of education provision. At the same time, the principle also envisions that compe-
tent students, irrespective of their background (i.e., low or high SES), are assumed
to be among the most successful. The main goal of the comprehensive school sys-
tem in the Nordic countries as a whole has been to abolish the class-based society.

In both Sweden and Norway, compulsory education is free and ranges from
grades one through ten (Norway) and one through nine (Sweden). In Norway, upper
secondary school (in Norwegian *videregående skole/opplæring*) is voluntary but
legally accessible to all students, while in Sweden, upper secondary school (in
Swedish *gymnasieskolan*) is voluntary and free. Thus, at both levels, there is *equal-
ity of opportunity, providing* all students with access to prescribed educational lev-
els, irrespective of whether they use the opportunity or not. At the same time, over
the years, notable differences have appeared between the systems in Norway and
Sweden. The latter has gone through extensive decentralisation reforms (Blossing
& Söderström, 2014). The change has been coupled with more severe marketisation
and privatisation practices (Lundahl, 2016) and because of which Sweden has
somewhat lost its position as one of the most equitable school systems (Lundahl,
2016; Skolverket, 2013).

In addition, the importance of students' socioeconomic backgrounds has
increased. Although Wiberg (2019) shows that both the school context and the stu-
dents' background has had an impact on the students' TIMSS results in Sweden,
Broer, Bai, and Fonseca (2019) observe a substantial increase in the gap between
high- and low-SES students' achievement in mathematics. In contrast, an overall
decrease in the achievement gap in mathematics for the same period was reported
for Norway (Broer et al., 2019). The results in PISA also demonstrate this shift. The
results from the 2018 cycle show that for students in Sweden, 13.2% of the variation

in mathematics performance can be explained by SES, which is similar to the OECD average (OECD, 2019). In PISA 2012, a similar trend can be found—10.6% in Sweden and 14.8% in OECD countries. In Norway, this was 8.4% in 2018 and 7.4% in 2012, respectively. Despite the differences in the described trends, the results indicate that this unfavourable background does contribute to the poorer outcomes of some students. It also illustrates what Espinoza (2007) describes as the (*in*)*equality on average across the social groups* relative to student output and survival in the system. The former indicates students' adverse outcomes are linked to differentiation in the available resources between the SES groups, while the latter implies a higher dropout rate of the lower SES band (Farrell, 2013). Conversely, although such trend analyses are informative in keeping track of different educational processes in the system, what they often disregard is the heterogeneity of the student body (e.g., diversity in self-beliefs), even if students belong to the same SES categories. Focusing on such distinctive features may aid in providing a more differentiated portrait of students and their outcomes, even across settings.

11.4   Current Study

Against the background described in the previous sections, we focus on distinct differences between academically resilient students and nonresilient student groups (i.e., low-achieving students irrespective of SES and successful students with high SES) in connection to their self-beliefs related to mathematics (i.e., confidence, interest and value) and their sense of school belonging. Following this, we investigate the school and classroom environment of students who have distinct self-belief patterns and how these relate to students initially categorised as resilient or not. For this purpose, we utilise TIMSS 2015 grade eight data for Norway and Sweden within the context of the person-centred approach (Bergman & Trost, 2006) and multiple-group analyses (Morin et al., 2016). Both these methods enable us to better understand the different nuances across the student body in Sweden and Norway, not disregarding diversity in the applied thresholds when discerning resilient and nonresilient students (Sattler & Gershoff, 2019), either from a national or cross-national standpoint (OECD, 2018). Two research questions are central to this investigation:

(1) What are the characteristics of academically resilient students compared with the nonresilient groups in connection to the students' perceived confidence in mathematics, them valuing and liking mathematics as a subject and their sense of school belonging? Here, we expect optimal self-belief profiles to attract both resilient students and other high-achieving students (Erberber et al., 2015; Kalender, 2015; Martin & Marsh, 2006; Sandoval-Hernández & Białowolski, 2016; Wayman, 2002), while the low-achieving students, irrespective of the risk factors, will be more frequently found in the disfavourable self-belief profiles.

(2) What are the typical features of the school and classroom environment of students with distinct self-belief patterns, and how do these relate to students initially categorised as resilient or not? Based on previous studies, we postulate optimal self-belief profiles, saturated by resilient and non-risk-achieving students will be conducive to environments with strong school emphasis on academic success (Erberber et al., 2015; Hoy et al., 2006) and will be in a safe and orderly climate (Wang & Degol, 2016) with less frequently reported experiences of bullying (Erberber et al., 2015).

In addition to these two questions, we will observe whether the same patterns are discernible in both Sweden and Norway (Sandoval-Hernández & Białowolski, 2016) given the latest developments in the Swedish education system (Lundahl, 2016; Skolverket, 2013). Finally, we explore to what extent the patterns are transferable across low-achieving students given their differences in SES (Lee & Smith, 1999).

11.5 Methods

11.5.1 Participants

In the analyses, TIMSS mathematics 2015 data for grade eight in Norway and Sweden were used. The TIMSS framework implements strict sampling procedures at the country level, here following a two-step sequence. In the first step, a school sample is selected from a complete list of schools. The targeted population is grade eight students. In the second step, a random class is chosen in each of the schools (for details, see Mullis & Martin, 2013). The full data set for Norway totalled 4733 students and 4090 in Sweden, respectively.

Finally, to build the sample used in this investigation, both data sets were further stratified into four student categories. The *academically resilient* category comprised students from the lowest 25% on the SES scale who are at the same time among the 25% highest achieving students in the TIMSS mathematics test within their own country. The three comparison categories involved the *failing under risk* students (the lowest 25% on SES/the lowest 25% in mathematics achievement), the *low-achieving group* (the highest 25% on SES/the lowest 25% in mathematics achievement) and *the nonrisk achievers* (the highest 25% on SES/the highest 25% in mathematics achievement). The four categories were obtained for each country separately. Please see Table 11.1 for more details.

All later analyses were performed with these four categories as the principal sample constituents. At the same time, by defining these four categories in such a way, we could include the criteria of both low and high resilience thresholds (i.e., 'doing better than peers experiencing similar risks' and 'doing as well as peers not experiencing risk') in our investigation (Sattler & Gershoff, 2019).

Table 11.1 The four student categories comprising the final sample

	Sweden	Norway
Academically resilient group	94	109
Failing under risk students	511	463
Low-achieving group	78	136
Nonrisk achievers	459	531
Total	1142	1239

11.5.2 Measures

The TIMSS procedures have students first take a 90-minute test followed by a contextual questionnaire that captures different indices in connection to attitudes, beliefs and learning environment related to mathematics. The TIMSS provides an incomplete block design for the mathematics test (and science), while all students receive the same items for the contextual questionnaire (Mullis & Martin, 2013).

Mathematics and science teachers also receive a block of questions related to the various features of the classroom and the school environment. In the analyses, only mathematics teachers' data were used and deaggregated to the student level. Please see Table 11.2 for an overview.

11.5.3 Analyses

Upon preliminary descriptive analyses across the constructs in SPSS, the primary analyses were performed in Mplus, version 8.4 (Muthén & Muthén, 1998–2017). All missing data were handled using the FIML option. Following the pragmatics of a person-centred approach, the analyses were performed using latent profile analyses (LPA), which allows for the use of continuous indicators aligned with the nature of the constructs used. The LPA works by producing solutions with maximally different groups. Within each tested solution, it will assign individuals (i.e., students) who are similar across the examined indicators (i.e., student sense of school belonging, students like learning mathematics, student confidence in mathematics and students' value of mathematics) in one group. The individuals who are less similar to each other across the examined indicators will be assigned to different groups. The final outcome leads to homogeneous, but mutually exclusive, latent groups within a larger heterogeneous population, where each student is assigned to a single group. Each group represents a unique self-belief profile. Neither group composition nor the number of groups is known in advance (Geiser, 2013).

Because the profile analyses included two distinctive populations (i.e., Norway and Sweden), it was essential to investigate whether these samples could be treated as one or if the analyses were necessary for each country separately. In doing so, we were guided by the principles of the multiple-group analyses of similarity in latent profiles solutions proposed by Morin et al. (2016). In the first step, configural

Table 11.2 Constructs used in the study

Construct	Description	Mean (Standard Error) and missing rate (%)[a]	
		Norway	Sweden
	Achievement data		
Student achievement in mathematics	First plausible value was used	487 (2.0); 0%	501 (2.8); 0%
	Student background data		
Home educational resources	A composite score comprised of the number of books at home, the highest level of education of either parent and number of home study supports	11.4 (0.02); 1.3%	11.2 (0.02); 1.5%
	Students' self-belief and school environment constructs		
Students sense of school belonging	A seven-item composite score. Example items: 'I feel safe when I am at school'	10.7 (0.05) 0.2%	9.7 (0.06) 0.2%
Students like learning mathematics	A nine-item composite score. Example items: 'I enjoy learning mathematics'	9.7 (0.05) 0.3%	9.6 (0.06) 0%
Students confident in mathematics	A nine-item composite score. Example items: 'I learn things quickly in mathematics'	11.0 (0.07) 0.4%	10.5 (0.08) 0.4%
Students value mathematics	A nine-item composite score. Example items: 'I think learning mathematics will help me in my daily life'	10.3 (0.06) 1%	9.4 (0.05) 0.4%
Student bullying	A nine-item composite score. Example behaviour 'spread lies about me'	10.9 (0.05) 0.2%	10.4 (0.05) 0.1%
Engaging teaching in math lesson	A 10-item composite score. Example items: 'My teacher is easy to understand'	10.0 (0.05) 0.4%	9.6 (0.05) 0%
	School environment constructs (teachers)		
School emphasis on academic success	A 14-item composite score. Example of an aspect: 'Teachers' expectations for student achievement'	10.1 (0.10) 9.6%	9.6 (0.11) 2.9%
Safe and orderly schools	An eight-item composite score. Example items: 'I feel safe at this school'	11.2 (0.12) 9.6%	9.7 (0.13) 2.9%
School conditions and resources	A seven-item composite score. Example items: 'Teachers do not have adequate workspace'	10.3 (0.11) 9.6%	9.6 (0.11) 2.9%
Challenges facing teachers	An eight-item composite score. Example items: 'There are too many students in the classes'	9.4 (0.15) 9.6%	9.7 (0.12) 2.4%
Teaching limited by student needs	A six-item composite score. Example of the need: 'Disruptive students'	10.6 (0.12) 12.1%	10.5 (0.14) 6.3%

[a]Achievement and student background data refer to full national samples because these were used to construct the final sample used for the analyses

similarity of the profiles was validated, that is, whether the same number of latent profiles when using the same overarching model can be identified in both countries. The assumption was tested through a series of latent profile solutions that were estimated separately for Norway and Sweden by using the same set of profile indicators. In the next step, the structural similarity of the profiles was tested. This step determines whether the profiles for both countries are similar and represent a basis to explore other types of similarities or possible differences. The third step tests the dispersion similarity (i.e., if the within-profile variability of the indicators is similar across countries), followed by the distributional similarity of the profiles (i.e., if the relative size of the profiles differs or not across the nations). The final stages focused on explanatory similarity, allowing us to observe the profiles not just from the perspective of resilient and nonresilient groups in both countries, but also to investigate the school/classroom features surrounding each. All models were estimated using 5000 random start values sets with 100 iterations and the 200 best solutions retained for the final stage of optimisation.

11.6 Results

The results section is divided into three major components. First, we discuss the basis for establishing the joint self-belief profiles, which is followed by the saturation of resilient and nonresilient student groups. Finally, we observe the different aspects of the school environment in connection to the established self-belief profiles and how these relate to the four initial student categories.

11.6.1 Students' Self-Belief Profiles

In identifying the number of profiles, we examined solutions with up to seven profiles separately for Norway and Sweden. Please see the fit indices in Table 11.3. The statistical adequacy and interpretability of the solution guided our final choice for an optimal profile solution.

In the case of Norway, with the addition of the profiles, most indices continued decreasing, except for the BLRT, which remained unchanged across the inspected models. The LMR values supported five profiles, but after inspecting the neighbouring four- and the five-profile solutions, the former was accepted. The four-profile solution provided a more meaningful interpretation in relation to both the data and previous research (e.g., Kalender, 2015). Entropy was also satisfactory. In the Swedish sample, both the fit indices (AIC, BIC, SABIC, LMR) and the model interpretability supported a four-profile solution. Again, entropy was satisfactory.

A multiple-group model for the four-profile solution was then simultaneously estimated for both country samples to test for a cross-national similarity. We first tested for the configural and then the structural similarity. Both the models were

Table 11.3 Fit results from the latent profiles analyses on student self-belief profiles

	k	LL	#fp	AIC	BIC	SABIC	LMR	BLRT	VL-LRT	Entropy	SM%
Class enumeration—Norway											
Two profiles	2	−10328.535	13	20683.071	20749.658	20708.364	0.0000	0.0000	0.0000	0.714	45.5%
Three profiles	3	−10140.619	18	20317.237	20409.434	20352.258	0.0000	0.0000	0.0000	0.779	17%
Four profiles	4	−10095.588	23	20237.175	20354.983	20281.924	0.0091	0.0000	0.0081	0.727	11.4%
Five profiles	5	−10065.945	28	20187.889	20331.307	20242.366	0.0053	0.0000	0.0047	0.780	7.6%
Six profiles	6	−10044.816	33	20155.632	20324.660	20219.838	0.2929	0.0000	0.2846	0.798	0.6%
Seven profiles	7	−10026.685	38	20129.371	20324.009	20203.304	0.5804	0.0000	0.5731	0.763	0.9%
Class enumeration—Sweden											
Two profiles	2	−9397.271	13	18820.541	18886.068	18844.776	0.0000	0.0000	0.0000	0.726	44.5%
Three profiles	3	−9258.539	18	18553.079	18643.808	18586.635	0.0000	0.0000	0.0000	0.764	17.1%
Four profiles	4	−9185.197	23	18416.395	18532.327	18459.272	0.0000	0.0000	0.0000	0.777	12.2%
Five profiles	5	−9151.108	28	18358.216	18499.351	18410.415	0.1167	0.0000	0.1116	0.812	11.7%
Six profiles	6	−9127.419	33	18320.839	18487.176	18382.358	0.6193	0.0000	0.6149	0.783	7.9%
Seven profiles	7	−9093.331	38	18262.662	18454.203	18333.503	0.0431	0.0000	0.0410	0.799	0.7%
Cross-cultural similarity											
Configural	4	−20929.192	47	41952.384	42223.822	42074.493	/	/	/	0.834	/
Structural (means)	4	−21070.267	31	42202.533	42381.567	42283.073	/	/	/	0.826	/
Dispersion (means and variances)	4	−21072.793	27	42199.586	42355.519	42269.734	/	/	/	0.826	/
Distributional countries (means, variances, probabilities)	4	−21086.408	24	42220.816	42359.423	42283.170	/	/	/	0.830	/
Explanatory similarity: Student perceptions											
Freely estimated across countries	4	−30249.073	46	60590.146	60855.808	60709.657	/	/	/	0.849	/
Equality across countries	4	−30242.964	38	60561.928	60781.389	60660.655	/	/	/	0.826	/

Explanatory similarity: School characteristics

Freely estimated across countries	4	−41312.077	64	82752.154	83125.174	82921.829	/	/	/	0.851	/
Equality across countries	4	−41410.504	48	82917.007	83196.772	83044.264	/	/	/	0.823	/

Note: *LL* model log likelihood, *#fp* number of free parameters, *AIC* Akaike information criterion, *BIC* Bayesian information criterion, *SABIC* sample-size adjusted BIC, *LMR* Lo, Mendell and Rubin likelihood ratio test, *BLRT* bootstrap likelihood ratio test, *VL-LRT* Vuong-Lo-Mendell-Rubin likelihood ratio test, *SM%* smallest group frequency

confirmed. Please see Table 11.3 for more details. The next model, testing the dispersion similarity, showed somewhat lower values for the AIC, BIC and SABIC compared with the structural similarity model. These results support the dispersion similarity; that is, the within-profile variability of the indicators is similar across Norway and Sweden. The four self-belief profiles from the dispersion similarity model are shown in Fig. 11.1. We discuss each in connection to students' sense of school belonging, students liking to learn mathematics, students being confident in mathematics and students' valuing of mathematics.

Among the profiles, the largest share of the students (38%) do not seem to enjoy learning mathematics, find it boring or the topic to be of little interest. At the same time, the students perceive themselves as confident when it comes to mathematics as a subject. In their view, they are not lagging behind their peers and have experienced praise from their mathematics teacher. At the same time, these students do value mathematics and see it as a tool that can contribute to their success later in life. Across the dimensions, their sense of belonging to school, teachers and peers seems to be the most distinctive feature. We labelled this group as the *nonadmirers* because compared with the other profiles, these students do not seem to enjoy learning mathematics but still to some extent see the value of mathematics, are somewhat confident and report a high sense of school belonging.

The second-largest group (36%) is labelled *confident*. Across the dimensions that entered our analyses, students' confidence concerning mathematics is their strongest characteristic. The confidence is related to their perception of doing well in mathematics or mastering difficult tasks, along with the absence of negative emotions in relation to mathematics. Almost equally strongly students report on their feeling of belonging to the school and valuing mathematics concerning their later
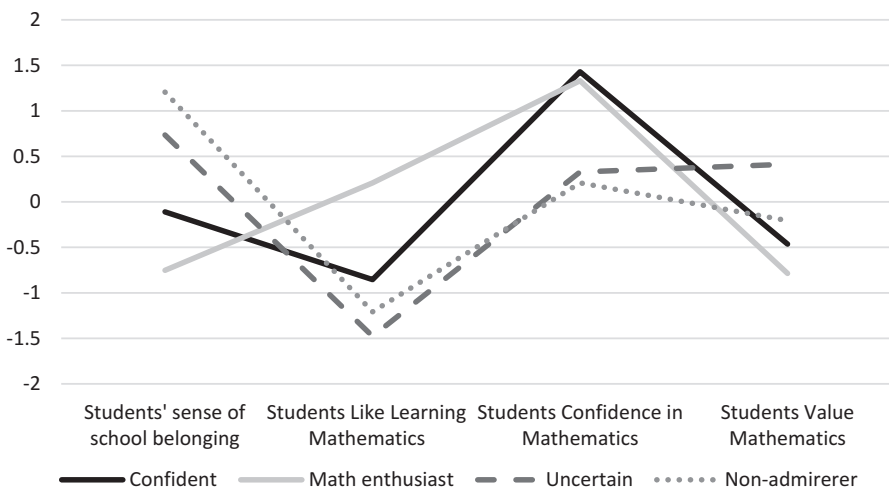

Fig. 11.1 Characteristics of the four identified self-belief profiles. (Note: The results were standardised to a mean of 0 and a standard deviation of 1 for visualisation purposes)

life choices or daily lives. These students enjoy learning content in mathematics, yet the subject does not necessarily represent their favourite domain of interest.

The next student profile, labelled *math enthusiasts,* comprise 15% of all students in our sample. Across all the dimensions, these students score the highest. These include their perception of belonging to the school environment they are part of, highly regarding mathematics as a domain that will aid them in their daily life, being successful in other school subjects or later being successful in obtaining a job they aspire. These students very much like and enjoy learning content related to mathematics. Mathematics is one of their favourite subject domains. Finally, compared with the other profiles, they perceive themselves as highly confident in relation to different aspects of dealing with the content of mathematics.

The last profile gathers students that, compared with others, enjoy or like learning mathematics the least. Also, these students are the least confident when grappling with the mathematical content. To an extent, they value mathematics and see it as useful for their future or success in other domains. Across the observed dimensions, they are most favourable in relation to how they perceive their sense of school belonging. We have labelled this group as the *uncertain* (11%).

11.6.2   Resilient and Nonresilient Students and Their Characteristics

The following steps in the analyses have allowed us to test the similarity in the size of the profiles across Sweden and Norway (the distributional similarity). Compared with previous results on dispersion, where we tested whether the within-profile variability of the indicators is similar across countries, the values across the observed criteria have increased (i.e., AIC, BIC, SABIC), suggesting that the sizes of the profiles somewhat differ across both Sweden and Norway. In both countries, *nonadmirers* and *confident* profiles are dominant. Both mount to nearly 37% in Norway and 41% for the former and 34% for the latter in Sweden. In absolute numbers, the *nonadmirers* have a greater share in the overall population compared with Norway, whereas for the *confident* profile, this share is about the same. Compared with Sweden, in Norway, we find more students in the *math enthusiasts* profile (17% and 12%, respectively). Finally, in both countries, the saturation within the *uncertain* profile is the least but somewhat higher for Sweden (9.5% and 13% respectively).

To further shed light on the presence of each of the four profiles, we observe the profiles concerning the categories that present the building blocks of our sample, that is, the academically resilient category, failing under risk students, low-achieving category and the nonrisk achievers (Table 11.4). Significant differences were registered regarding the occurrence of the student categories in relation to the examined profiles (χ^2 (21) = 826.634, p < 0.001). In both countries, the *math enthusiast* and *confident* profiles have a higher occurrence among the nonrisk achievers. Both of the profiles are viewed as optimal because both in different ways suggest positive

Table 11.4 Students' self-belief profiles and categories

Sample categories		Student profiles			
		Confident	Math enthusiast	Uncertain	Nonadmirer
Norway					
Academically resilient group	%	55.0%	18.3%	5.5%	21.1%
	St. res.	3.3	0.9	−1.7	−2.9
Failing under risk students	%	27.4%	4.3%	17.3%	51.0%
	St. res.	−3.1	−5.9	4.1	4.5
Low-achieving group	%	33.8%	2.2%	16.2%	47.8%
	St. res.	−0.4	−3.8	1.8	1.9
Nonrisk achievers	%	48.0%	33.1%	1.7%	17.1%
	St. res.	4.6	10.8	−6.4	−7.8
Sweden					
Academically resilient group	%	47.9%	11.7%	6.4%	34.0%
	St. res.	1.9	−0.8	−1.3	−0.6
Failing under risk students	%	13.1%	2.3%	22.7%	61.8%
	St. res.	−8.6	−7.4	8.0	8.7
Low-achieving group	%	16.7%	1.3%	23.1%	59.0%
	St. res.	−2.9	−3.1	3.2	3.0
Nonrisk achievers	%	53.6%	24.6%	0.9%	20.9%
	St. res.	6.2	5.4	−6.5	−5.9

Note: Frequency is provided for within the category. The standard residual with values over 1.9 indicate a statistically significant difference

attitudes and emotions towards a variety of aspects in connection to mathematics learning or valuing one's own competence. Although the pattern is stronger for the *confident* profile in Sweden, in Norway, this is the case for the *math enthusiasts* (see Table 11.4, standardised residual values). As expected, the failing under risk category of students is underrepresented in both of these profiles, but again, the pattern is stronger than in Sweden. As to the low-achieving category, we find significantly fewer students in Sweden belonging to these two profiles. For Norway, the trend is only noticeable for the *confident* profile.

Overall, the failing under risk and low-achieving categories are overrepresented in the *uncertain* and the *nonadmirer* profiles. Both profiles are linked to negative perceptions of mathematics as a subject. Students within the latter group do exhibit some confidence in their competence and, to an extent, may value mathematics. With some exceptions for Norway, the pattern is very strong for both these groups in Sweden.

As for the academically resilient students, in Norway, they are underrepresented in the *nonadmirer* profile and overrepresented in the *confident* profile. The latter is similar to the results for Sweden although the pattern is somewhat weaker. Interestingly, when observing the resilient students, no distinctive pattern is found for the *math enthusiast* profile in either of the countries. Although these students may be found among those that profoundly enjoy and like learning the content of mathematics, value the subject and are confident when grappling with the

mathematical content, they are rather overrepresented in the other profile that we also consider to be optimal. At the same time, the mere fact we find significantly more resilient students within the *confident* profile can also be seen as a distinctive characteristic of resilience itself. Resilient students succeed despite an adverse background or when met with a set of unfavourable factors. To do so, they need to believe in themselves and the abilities they possess. For the confident profile, this is the very thing that sets them apart from the other profiles we found.

11.6.3 The School and the Classroom Environment

An essential aspect of our investigation was also related to examining classroom and school environment features distinctive to the profiles and how the profile membership is differentially associated with each of these. To achieve this aim, we tested an explanatory similarity model across the samples, starting with student perceptions of student bullying and engagement of the teaching in the math lesson. We first conducted a model that allows within-profile levels of both these aspects to be freely estimated across the samples and then a model in which these levels were constrained to be equal across the samples. The latter model resulted in lower values for AIC, BIC and SABIC (Table 11.3), thus supporting the explanatory similarity. Systemic tests of mean level differences across the pairs of profiles (Table 11.5) revealed significant differences between all the profiles regarding students' perceptions of engaging teaching in mathematics. The *math enthusiast* group, which is saturated by resilient and nonrisk achievers, holds the most positive perceptions of the instruction they are exposed to. In their view, the teachers are clear with the instruction, engaging and provide them with feedback that is attuned to their needs.

Table 11.5 Students' self-belief profiles and classroom and school environment

Sample categories	Student profiles				Tests of significance
	Math enthusiast	Confident	Uncertain	Nonadmirer	
Student bullying	11.160	10.094	9.881	11.124	1 > 2; 1 > 3; 1 = 4; 2 = 3; 2 < 4; 3 < 4
Engaging teaching in math lesson	11.402	9.522	7.976	9.899	1 > 2; 1 > 3; 1 > 4; 2 > 3; 2 < 4; 3 < 4
School emphasis on academic success	10.558	9.180	9.313	10.506	1 > 2; 1 > 3; 1 = 4; 2 = 3; 2 < 4; 3 < 4
Safe and orderly schools	11.072	8.969	9.961	11.566	1 > 2; 1 = 3; 1 = 4; 2 = 3; 2 < 4; 3 < 4
School conditions and resources	10.317	9.184	9.959	10.584	1 = 2; 1 = 3; 1 = 4; 2 < 3; 2 < 4; 3 = 4
Challenges facing teachers	9.248	9.718	9.513	9.341	1 < 2; 1 = 3; 1 = 4; 2 = 3; 2 = 4; 3 = 4

Note: See Appendix for details on significance tests

The least positive perceptions are typical of the *uncertain* group. For them, the instruction is difficult to follow, uninteresting and without clear evidence to show them if they have mastered the subject. If we take into account that the *uncertain* profile is very much saturated by failing under risk students, this finding creates an opportunity to observe the further particular needs this group may have in relation to the instruction they may or may not be receiving. The mean level differences across the profiles have also been captured concerning students' perceptions of bullying experienced. However, neither of these values falls under 9.3, which is a critical score that shifts a student from the category 'almost never' to 'about monthly' when observing this scale within the existing TIMSS framework (Mullis & Martin, 2013; see also Appendix).

Finally, we tested whether the relations between the student profiles and particular distal features—school emphasis on academic success, safe and orderly schools, school conditions and resources and the challenges facing teachers and teaching limited by students' needs are replicated across the profiles. The assumption was not confirmed (the AIC, BIC and SABIC values are lower for equality across the countries model, Table 11.3). Furthermore, the mean level test differences across the pairs of profiles (Table 11.5) reveal a distinctive pattern.

According to teacher perceptions, all students are placed in school environments in which their math teachers do face some challenges related to the organisation of their own teaching, or minor problems are reported as to the school's conditions and resources. Yet the challenges seem to be somewhat more significant for students belonging to the *confident* profile. Given the fact that both resilient and nonrisk achievers saturate the profile, it is essential to understand how these students manage to compensate for the possible barriers related to these challenges and the extent their self-belief capacities aid them in the process. The finding is even more important in light of the result that these students are also in school environments in which the school's emphasis on academic success is perceived to be at a medium level. Similarly, students within the *uncertain* profile face the same challenge. Although the profile itself is the least optimal among all the profiles, the *uncertain* profile is also very much saturated by students in the category failing under risk and low-achieving students. Thus, the finding, together with students' perception of less-than-engaging teaching in mathematics, may imply some of the students in this profile could be facing adverse conditions both at school and home.

Although the nonadmirer profile is also saturated by students in the categories failing under risk and low-achieving, the school's emphasis on academic success is ranked very high, and students have perceived the teaching as engaging. This could be viewed as a compensatory mechanism that aids the students who fail in maintaining certain levels of confidence and valuing mathematics, which is distinctive of the profile. Despite some mean differences across the four profiles for the safe and orderly school as reported by the teachers, overall, all students may be tied to at least safe and orderly category. In the case of *math enthusiasts* and *nonadmirer* profiles, the teachers report very safe and orderly school milieus.

11.7   Discussion

Ensuring the success of every child in an equal way is an essential aspect of the agenda for many education providers across the world (Doll, 2013; Pianta & Walsh, 1998). In particular, this includes the Nordic countries, which are often viewed as among the top leading countries with such an agenda (Blossing et al., 2014), despite the argument regarding how Sweden has somewhat lost its position among these countries (Blossing & Söderström, 2014; Lundahl, 2016; Skolverket, 2013). At the same time, a significant strand of researchers has been trying to capture and examine the mechanisms that could explain the achievement gap between low- and high-SES students (e.g. Broer et al., 2019; OECD, 2019; Wiberg, 2019) and how these may relate different education reforms (e.g., Lundahl, 2016). Conversely, others focus on the adverse background students may be facing, here mapping out both the student and school characteristics that could support students' academic success despite the adversities they face (Doll, 2013). Against this background, we conducted a study aiming to examine the distinct differences between resilient and nonresilient student groups (Lee & Smith, 1999; OECD, 2018; Sattler & Gershoff, 2019) in connection to their self-beliefs related to mathematics (i.e., confidence, interest and value) and sense of school belonging. Furthermore, we investigated the school and classroom environment of students with distinct self-belief patterns and how these relate to students initially categorised as resilient or not.

We expected the optimal self-belief profiles to attract both resilient students (i.e., high-achieving students with low SES) and nonrisk achievers (i.e., high-achieving students with high SES) (Erberber et al., 2015; Kalender, 2015; Martin & Marsh, 2006; Sandoval-Hernández & Białowolski, 2016; Wayman, 2002). We also expected low-achieving students, irrespective of their SES, to be more frequently found in the less-optimal profiles. The assumption was partially confirmed. Both the *math enthusiast* and *confident* profiles gathered substantially more nonrisk achievers. At the same time, in both Norway and Sweden, a large fraction of the resilient students were found in the *confident* profile but far less were found in the *math enthusiast* profile. The fact that we do capture more resilient students in a profile associated with high levels of confidence in mathematics resonates with the findings in existing studies (e.g., Martin & Marsh, 2006; Sandoval-Hernández & Białowolski, 2016; Wayman, 2002). However, the aspect of liking mathematics and genuinely enjoying grappling with the mathematical content, which was more a characteristic of the math enthusiast profile, was strongly linked with the nonrisk achievers. For students saturating each of the two profiles, the higher proportion of resilient students being found in the confident profile may be expected. If we regard confidence as one of the major correlates of resilience, its higher levels could be viewed as an aid or even a compensatory mechanism students develop in battling adverse circumstances as they strive to succeed. Enjoyment in an actual activity may become secondary, but it is essential for students' perseverance.

Both the *nonadmirer* and the *uncertain* profiles were overrepresented by the low-achieving and failing under risk categories. At the same time, students in the

nonadmirer profile exhibit some confidence in their competence and, to an extent, value mathematics, even though they may find it uninteresting. Thus, the finding contradicts the expected pattern reported in previous studies (e.g., Kalender, 2015; Sandoval-Hernández & Białowolski, 2016), indicating further investigation is necessary for contrasting the resilient and nonresilient student categories. This also underlines the importance of taking into account the heterogeneity of the student population and points to the importance of the different criteria used in distinguishing resilient from the nonresilient students (i.e., low-threshold and high-threshold resilience; Sattler & Gershoff, 2019) and how their combination adds to the complexity in assessing the needs of diverse students.

The other part of our investigation focused on the constituents of the classroom and school environment. Based on previous research, we expected optimal profiles (i.e., *math enthusiasts* and the *confident* profile) to be found within environments with a strong school emphasis on academic success (Erberber et al., 2015; Hoy et al., 2006) and a safe and orderly climate (Wang & Degol, 2016) with less frequently reported experiences of bullying (Erberber et al., 2015). Again, these assumptions were only partially confirmed. Across profiles, a safe and orderly climate and almost no experience with bullying was reported by both teachers and students. Although this contradicts some previous results (Erberber et al., 2015; Wang & Degol, 2016), the finding is in line with the results for Sweden and Norway on the lower frequency of reported bullying overall and safe school environments (Jensen et al., 2019; OECD, 2018). However, the findings in connection to school emphasis on academic success and school conditions and resources point to some distinctive patterns across the profiles, shedding light on the very idea of equality of opportunities and outcome, as proposed by Espinoza (2007). Across the investigated profiles, students within the *confident* profile seem to be affected the most by the reported challenges related to the organisation of teaching or minor difficulties with the school conditions and resources. The finding is coupled with the perception of a school environment not strongly focusing on academic success. Importantly, resilient and nonrisk achiever categories saturate the *confident* profile, and both are strong achievers. Students within the uncertain profile face the same challenges, but compared with the previous profile, they report less-engaging teaching. The *uncertain* profile is also very much saturated by students in the category failing under risk and low-achieving students. Thus, although the school environment is the same for both, within their immediate surrounding—the classroom—students from the *confident* profile do experience some compensatory mechanism through instruction they perceive as engaging (Ungar et al., 2019) and are supported by their own strong performance. Regarding the *uncertain* profile, the adverse characteristics pertain both in the school and classroom, and for many, these extend to the home environment (i.e., low SES). If we take into account that education in the Nordic welfare system has been regarded as a crucial instrument for social justice and security, the perceived differences for the two described profiles contradict this very idea. Although both are examples of inequality of opportunity, the uncertain profile is also under risk of inequality related to the output when considering the profile's

saturation with failing and low SES students, as well as availability in the school and classroom resources (Espinoza, 2007).

As reported, the *nonadmirers* profile is also overrepresented by students in the categories failing under risk and low achieving. However, their teachers reported strong school emphasis on academic success, whereas the students themselves perceived the teaching as engaging. Both aspects could be observed as compensatory mechanisms that aid students who fail in maintaining certain levels of confidence and valuing mathematics, which is distinctive of the profile. Thus, in the context of support that schools can provide to students (Hoy et al., 2006; Kyriakides et al., 2010; Maxwell et al., 2017), we postulate that interventions addressing these two aspects could be a fruitful ground in ensuring more equal chances across different at-risk groups, supporting both their outcomes and engagement (Martin et al., 2013; O'Brennan & Furlong, 2010; Wang & Degol, 2016).

Finally, some distinctive differences between Norway and Sweden were found in the frequency of each student category, despite the existence of the same four profiles in both countries. In the context of the necessity to primarily support students with adverse social background (i.e., resilient and failing under risk categories), in Norway, these students are represented more frequently in the optimal self-beliefs profiles. Given the latest trends in the Swedish education system (e.g., Lundahl, 2016) and a higher degree of SES variation within compulsory schools in Norway (OECD, 2019), this warrants further investigation into school characteristics and existing practices, going beyond school SES (Lee & Smith, 1999) and again controlling for country differences (Sandoval-Hernández & Białowolski, 2016).

11.7.1   Limitation and Future Research

One important limitation of our work stems from the nature of the data used, that is, cross-sectional data. Although the data allow for diverse analyses such as latent profiling, it is not possible to determine if the students remain in the same profiles across time. Besides the opportunity to follow student trajectories, latent transition analyses would allow for an even more in-depth understanding of the interplay between the individual, classroom and school-level characteristics in the context of these phenomena. A study of this nature would also aid in tracing the possible effects of targeted intervention towards both resilient and nonresilient student categories. In the context of our results, one such example would be following whether students from the *uncertain* profile may shift to the *nonadmirers* after being supported by an intervention at the school and classroom level or whether the same transition may occur between the *confident* and *math enthusiast* profile. Second, although the TIMSS data have allowed us to run models for Norway and Sweden simultaneously, we were limited by the constructs provided in the study and its overall organisation. Thus, it can be argued that more or other school and classroom characteristics may have been added to this investigation, both at the teacher and the student level. Although this may hold, our current choice from the variable poll was

anchored and supported by previous research. Finally, aligned with the person-centred approach, students and their characteristics were the focus of this investigation, irrespective of the actual schools they may have been enrolled in. The former was the reason for merging student and teacher data, that is, deaggregating them to the student level. Although the process can lead to a type II error, thus underestimating some patterns in the data, we could also observe evidence where the student and teacher data were corroborating with each other. An example could be found in the findings on the perception of a safe and orderly climate.

11.8   Concluding Remarks

Apart from the optimal outcomes concerning achievement, students are also expected to develop an optimal set of self-beliefs that will aid them in the process of education and allow them to persevere once their interests are further profiled. Such expectations are set for all students, irrespective of their social backgrounds. Following Erberber et al. (2015), who caution proposing universal recipes in comparative research, a more focused country choice approach was adopted in the current study. The method was coupled with combining low- and high-threshold criteria when distinguishing between resilient and nonresilient student categories (Sattler & Gershoff, 2019), allowing us to identify particular patterns about the classroom and school characteristics after accounting for students' self-beliefs. The method has also aided in moving beyond the resilient and nonresilient divide, capturing fine-grained differences in the student population, respective also to the cross-country differences. In addition, we were able to identify particular student profiles that might be more prone to risk after accounting not merely for students' SES, but also for the individual strengths and hindrances in the classroom, along with the school setting. In the context of the equality–inequality paradigm (Espinoza, 2007), recognition of such student subgroups strengthen the possibility to reduce the gap in battling different aspects of inequality across social groups.

   Also, our results speak both in favour and against the very idea of the Nordic model. Although we have been able to distinguish the same student groups across Sweden and Norway, confirming some commonalities across both countries, their distribution within each differs. Observing more students with an adverse social background in the optimal self-beliefs profiles was not replicated in Sweden the same way it was in Norway. This result speaks of some diverse pathways although both countries are considered representative of the Nordic model. Having in mind the latest developmental trends in the Swedish education system and student composition across the schools in Norway creates space for a more focused investigation into the existence of particular school practices and mechanisms catering to diverse students within the education system of both countries. Their existence could provide clearer evidence of the possible dissolution of the Nordic model.

Appendices

Appendix 1: List of Constructs with Items and Scale Range

Construct	Construct items	Range
Student questionnaire		
Home educational resources	(1) Number of books in the home	Scale ranging from *Many Resources* (score of at least 12.4) to *Few Resources* (score no higher than 8.3)
	(2) Highest level of education of either parent	
	(3) Number of home study supports	
Sense of school belonging scale, eight grade	(1) I like being in school	Scale ranging from *High Sense of School Belonging* (score of at least 10.3) to *Little Sense of School Belonging* (score no higher than 7.5)
	(2) I feel safe when I am at school	
	(3) I feel like I belong at this school	
	(4) I like to see my classmates at school	
	(5) Teachers at my school are fair to me	
	(6) I am proud to go to this school	
	(7) I learn a lot in school	
Students like learning mathematics	(1) I enjoy learning mathematics	Scale ranging from *Very Much Like Learning Mathematics* (score of at least 11.4) to *Do Not Like Learning Mathematics* (score no higher than 9.4)
	(2) I wish I did not have to study mathematics*	
	(3) Mathematics is boring*	
	(4) I learn many interesting things in mathematics	
	(5) I like mathematics	
	(6) I like any schoolwork that involves numbers	
	(7) I like to solve mathematics problems	
	(8) I look forward to mathematics class	
	(9) Mathematics is one of my favorite subjects	

(continued)

Construct	Construct items	Range
Students confident in mathematics	(1) I usually do well in mathematics	Scale ranging from *Very Confident in Mathematics* (score of at least 12.1) to *Not Confident in Mathematics* (score no higher than 9.5)
	(2) Mathematics is more difficult for me than for many of my classmates*	
	(3) Mathematics is not one of my strengths*	
	(4) I learn things quickly in mathematics	
	(5) Mathematics makes me nervous*	
	(6) I am good at working out difficult mathematics problems	
	(7) My teacher tells me I am good at mathematics	
	(8) Mathematics is harder for me than any other subject*	
	(9) Mathematics makes me confused*	
Students value mathematics	(1) I think learning mathematics will help me in my daily life.	Scale ranging from *Strongly Value Mathematics* (score of at least 10.3) to *Do Not Value Mathematics* (score no higher than 7.7)
	(2) I need mathematics to learn other school subjects	
	(3) I need to do well in mathematics to get into the university of my choice	
	(4) I need to do well in mathematics to get the job I want	
	(5) I would like a job that involves using mathematics	
	(6) It is important to learn about mathematics to get ahead in the world	
	(7) Learning mathematics will give me more job opportunities when I am an adult	
	(8) My parents think that it is important that I do well in mathematics	
	(9) It is important to do well in mathematics	

Construct	Construct items	Range
Student questionnaire		
Student bullying	(1) Made fun of me or called me names	Scale ranging from *Almost Never* (score of at least 9.3) to *About Weekly* (score no higher than 7.3)
	(2) Left me out of their games or activities	
	(3) Spread lies about me	
	(4) Stole something from me	
	(5) Hit or hurt me (e.g. shoving, hitting, kicking)	
	(6) Made me do things I didn't want to do	
	(7) Shared embarrassing information about me	
	(8) Posted embarrassing things about me online	
	(9) Threatened me	
Engaging teaching in math lesson	(1) I know what my teacher expects me to do	Scale ranging from *Very Engaging Teaching* (score of at least 10.4) to *Less Than Engaging Teaching* (score no higher than 8.2)
	(2) My teacher is easy to understand	
	(3) I am interested in what my teacher says	
	(4) My teacher gives me interesting things to do	
	(5) My teacher has clear answers to my questions	
	(6) My teacher is good at explaining mathematics	
	(7) My teacher lets me show what I have learned	
	(8) My teacher does a variety of things to help us learn	
	(9) My teacher tells me how to do better when I make a mistake	
	(10) My teacher listens to what I have to say	

Construct	Construct items	Range
Teacher questionnaire		
School emphasis on academic success	(1) Teachers' understanding of the school's curricular goals	Scale ranging from *Very High Emphasis* (score of at least 13.4) to *Medium Emphasis* (score no higher than 9.8)
	(2) Teachers' degree of success in implementing the school's curriculum	
	(3) Teachers' expectations for student achievement	
	(4) Teachers working together to improve student achievement	
	(5) Teachers' ability to inspire students	
	(6) Parental involvement in school activities	
	(7) Parental commitment to ensure that students are ready to learn	
	(8) Parental expectations for student achievement	
	(9) Parental support for student achievement	
	(10) Parental pressure for the school to maintain high academic standards	
	(11) Students' desire to do well in school	
	(12) Students' ability to reach school's academic goals	
	(13) Students' respect for classmates who excel in school	
	(14) Collaboration between school leadership and teachers to plan instruction	
Safe and orderly schools	(1) This school is located in a safe neighborhood	Scale ranging from *Very Safe and Orderly* (score of at least 10.6) to *Less than Safe and Orderly* (score no higher than 7.2)
	(2) I feel safe at this school	
	(3) This school's security policies and practices are sufficient	
	(4) The students behave in an orderly manner	
	(5) The students are respectful of the teachers	
	(6) The students respect school property	
	(7) This school has clear rules about student conduct	
	(8) This school's rules are enforced in a fair and consistent manner	

(continued)

Construct	Construct items	Range
Teacher questionnaire		
School conditions and resources	(1) The school building needs significant repair	Scale ranging from *Hardly Any Problems* (score of at least 10.9) to *Moderate to Severe Problems* (score no higher than 8.5)
	(2) Teachers do not have adequate workspace (e.g., for preparation, collaboration, or meeting with students)	
	(3) Teachers do not have adequate instructional materials and supplies	
	(4) The school classrooms are not cleaned often enough	
	(5) The school classrooms need maintenance work	
	(6) Teachers do not have adequate technological resources	
	(7) Teachers do not have adequate support for using technology	
Challenges facing teachers	(1) There are too many students in the classes	Scale ranging from *Few Challenges* (score of at least 10.3) to *Many Challenges* (score no higher than 6.7)
	(2) I have too much material to cover in class	
	(3) I have too many teaching hours	
	(4) I need more time to prepare for class	
	(5) I need more time to assist individual students	
	(6) I feel too much pressure from parents	
	(7) I have difficulty keeping up with all of the changes to the curriculum	
	(8) I have too many administrative tasks	
Teaching limited by student needs	(1) Students lacking prerequisite knowledge or skills	Scale ranging from *Not Limited* (score of at least 11.4) to *Very Limited* (score no higher than 7.4)
	(2) Students suffering from lack of basic nutrition	
	(3) Students suffering from not enough sleep	
	(4) Disruptive students	
	(5) Uninterested students	
	(6) Students with mental, emotional or psychological disabilities	

Note: Constructs and items from the TIMSS 2015 contextual questionnaire, reproduced from Martin et al. (2016). Items marked with an asterisk (*) are reverse-coded

Appendix 2: Significance Tests on Students' Profiles and Classroom and School Environment Constructs

	Student profiles			
	Math enthusiast	Confident	Uncertain	Nonadmirer
	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)
Sample Categories	p-value	p-value	p-value	p-value
	Student Bullying			
Math enthusiast		1.066 (0.235)	1.280 (0.148)	0.036 (0.118)
		0.000	0.000	0.760
Confident			0.214 (0.260)	−1.030 (0.258)
			0.412	0.000
Uncertain				−1.244 (0.147)
				0.000
	Engaging Teaching in Math Lesson			
Math enthusiast		1.880 (0.132)	3.427 (0.129)	1.504 (0.139)
		0.000	0.000	0.000
Confident			1.547 (0.149)	−0.376 (0.171)
			0.000	0.028
Uncertain				−1.923 (0.116)
				0.000
	School Emphasis on Academic Success			
Math enthusiast		−1.378 (0.337)	−1.246 (0.618)	−0.052 (0.292)
		0.000	0.044	0.857
Confident			0.133 (0.335)	−1.326 (0.199)
			0.692	0.000
Uncertain				−1.193 (0.410)
				0.004
	Safe and Orderly Schools			
Math enthusiast		−2.104 (0.557)	−1.112 (1.137)	0.493 (0.418)
		0.000	0.328	0.238
Confident			0.992 (0.619)	−2.597 (0.249)
			0.109	0.000
Uncertain				−1.605 (0.784)
				0.041
	School Conditions and Resources			
Math enthusiast		−1.133 (0.587)	−0.358 (0.594)	0.266 (0.287)
		0.054	0.546	0.353
Confident			0.775 (0.129)	−1.400 (0.356)
			0.000	0.000
Uncertain				−0.625 (0.383)
				0.103

(continued)

Sample Categories	Student profiles			
	Math enthusiast	Confident	Uncertain	Nonadmirer
	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)
	p-value	p-value	p-value	p-value
	Challenges Facing Teachers			
Math enthusiast		0.470 (0.154)	0.264 (0.302)	0.092 (0.311)
		0.002	0.382	0.766
Confident			−0.206 (0.228)	0.378 (0.225)
			0.367	0.093
Uncertain				0.172 (0.162)
				0.287

References

Abelev, M. S. (2009). Advancing out of poverty: Social class worldview and its relation to resilience. *Journal of Adolescent Research, 24*(1), 114–141.

Agasisti, T., Avvisati, F., Borgonovi, F., & Longobardi, S. (2018). *Academic resilience: What schools and countries do to help disadvantaged students succeed in PISA* (OECD education working papers) (Vol. 167). OECD Publishing.

Bergman, L. R., & Trost, K. (2006). The person-oriented versus the variable-oriented approach: Are they complementary, opposites, or exploring different worlds? *Merrill-Palmer Quarterly, 52*, 601–632.

Blossing, U., Imsen, G., & Moos, L. (2014). Nordic schools in a time of change. In U. Blossing, G. Imsen, & L. Moos (Eds.), *The Nordic education model. Policy implications of research in education* (Vol. 1, pp. 1–14). Springer.

Blossing, U., & Söderström, Å. (2014). A school for every child in Sweden. In U. Blossing, G. Imsen, & L. Moos (Eds.), *The Nordic education model. Policy implications of research in education* (Vol. 1, pp. 17–34). Springer.

Borman, G. D., & Dowling, M. (2010). Schools and inequality: A multilevel analysis of Coleman's equality of educational opportunity data. *Teachers College Record, 112*(5), 1201–1246.

Broer, M., Bai, Y., & Fonseca, F. (2019). Socio-economic achievement gaps: Trend results for education systems. In M. Broer, Y. Bai, & F. Fonseca (Eds.), *Socioeconomic inequality and educational outcomes. IEA research for education* (A series of in-depth analyses based on data of the international association for the evaluation of educational achievement (IEA)) (pp. 35–70). Springer.

Bronfenbrenner, U., & Morris, P. A. (2006). The bioecological model of human development. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology: Vol. 1. Theoretical models of human development* (6th ed., pp. 793–828). Wiley.

Doll, B. (2013). Enhancing resilience in classrooms. In S. Goldstein & R. B. Brooks (Eds.), *Handbook of resilience in children* (pp. 399–409). Springer.

Doll, B., Zucker, S., & Brehm, K. (2004). *Resilient classrooms: Creating healthy environments for learning*. Guilford Publications.

Eisenberg, N., Valiente, C., Fabes, R. A., Smith, C. L., Reiser, M., Shepard, S. A., et al. (2003). The relations of effortful control and ego control to children's resiliency and social functioning. *Developmental Psychology, 39*, 761–776.

Erberber, E., Stephens, M., Mamedova, S., Ferguson, S., & Kroeger, T. (2015). Socioeconomically disadvantaged students who are academically successful: Examining academic resilience cross-nationally. *IEA's Policy Brief Series, 5.*

Espinoza, O. (2007). Solving the equity–equality conceptual dilemma: A new model for analysis of the educational process. *Educational Research, 49*(4), 343–363.

Farrell, J. P. (2013). Equality of education: Six decades of comparative evidence seen from a new millennium. In R. F. Arnove, C. A. Torres, & S. Franz (Eds.), *Comparative education: The dialectic of the global and the local* (4th ed., pp. 149–174). Rowman & Littlefield Publishers.

Franklin, W. (2000). Students at promise and resilient: A historical look at risk. In M. G. Sanders (Ed.), *Schooling students placed at risk research, policy, and practice in the education of poor and minority adolescents* (pp. 3–16). Lawrence Erlbaum Associates, Publishers.

Geiser, C. (2013). *Data analyses with Mplus*. Guilford Press.

Hernandez-Martinez, P., & Williams, J. (2013). Against the odds: Resilience in mathematics students in transition. *British Educational Research Journal, 39*(1), 45–59.

Hoy, W. K., Tarter, C. J., & Hoy, A. W. (2006). Academic optimism of schools: A force for student achievement. *American Educational Research Journal, 43*(3), 425–446.

Jensen, F., Pettersen, A., Frønes, T. S., Kjærnsli, M., Rohatgi, A., Eriksen, A., et al. (2019). *PISA 2018. Norske elevers kompetanse i lesing, matematikk og naturfag*. Universitetsforlaget.

Kalender, I. (2015). An analysis of the resilient students' profile based on PISA. *Journal of Measurement and Evaluation in Education and Psychology, 6*(1), 158–172.

Kyriakides, L., Creemers, B., Antoniou, P., & Demetriou, D. (2010). A synthesis of studies searching for school factors: Implications for theory and research. *British Educational Research Journal, 36*(5), 807–830.

Lee, V. E., & Smith, J. B. (1999). Social support and achievement for young adolescents in Chicago: The role of school Academic Press? *American Educational Research Journal, 36*(4), 907–945.

Lessard, A., Butler-Kisber, L., Fortin, L., & Marcotte, D. (2014). Analysing the discourse of dropouts and resilient students. *The Journal of Educational Research, 107*(2), 103–110.

Lundahl, L. (2016). Equality, inclusion and marketisation of Nordic education: Introductory notes. *Research in Comparative and International Education, 11*(1), 3–12.

Martin, A. J., & Marsh, H. W. (2006). Academic resilience and its psychological and educational correlates: A construct validity approach. *Psychology in the Schools, 43*(3), 267–281.

Martin, A. J., & Marsh, H. W. (2008). Academic buoyancy: Towards an understanding of students' everyday academic resilience. *Journal of School Psychology, 46*, 53–83.

Martin, M. O., Foy, P., Mullis, I. V., & O'Dwyer, L. M. (2013). Effective schools in reading, mathematics, and science at fourth grade. In M. O. Martin & I. V. Mullis (Eds.), *TIMSS and PIRLS 2011: Relationships among reading, mathematics, and science achievement at the fourth grade—Implications for early learning* (pp. 109–180). International Association for the Evaluation of Educational Achievement (IEA) and Lynch School of Education, Boston College.

Martin, M. O., Mullis, I. V. S., & Hooper, M. (Eds.) (2016). Methods and Procedures in TIMSS 2015. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Masten, A. S. (2014). Global perspectives on resilience in children and youth. *Child Development, 85*, 6–20.

Masten, A. S. (2018). Resilience theory and research on children and families: Past, present, and promise. *Journal of Family Theory & Review, 10*, 12–31.

Maxwell, S., Reynolds, K., Lee, E., Subasic, E., & Bromhead, D. (2017). The impact of school climate and school identification on academic achievement: Multilevel modeling with student and teacher data. *Frontiers in Psychology, 8.* https://doi.org/10.3389/fpsyg.2017.02069

Morin, A. J. S., Meyer, J. P., Creusier, J., & Biétry, F. (2016). Multiple-group analysis of similarity in latent profile solutions. *Organizational Research Methods, 19*(2), 231–254.

Mullis, I. V. S., & Martin, M. O. (Eds.). (2013). *TIMSS 2015 assessment frameworks*. Boston College, TIMSS & PIRLS International Study Center.

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide* (8th ed.). Muthén & Muthén.

Nilsen, T., Blömeke, S., Hansen, K. Y., & Gustafsson, J.- E. (2016). *Can schools contribute to enhance equity? The answer may depend on a country's developmental level*. IEA Policy Brief.

O'Brennan, L. M., & Furlong, M. J. (2010). Relations between students' perceptions of school connectedness and peer victimisation. *Journal of School Violence, 9*(4), 375–391.

OECD. (2018). *Equity in education: Breaking down barriers to social mobility*. OECD Publishing.

OECD. (2019). *TALIS 2018 results (volume I): Teachers and school leaders as lifelong learners*. OECD Publishing.

Perry, L., & McConney, A. (2010). School socio-economic composition and student outcomes in Australia: Implications for educational policy. *Australian Journal of Education, 54*(1), 72–85.

Pianta, R. C., & Walsh, D. J. (1998). Applying the construct of resilience in schools: Cautions from a developmental systems perspective. *School Psychology Review, 27*, 407–417.

Reardon, S. F. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In R. Murnane & G. Duncan (Eds.), *Whither opportunity? Rising inequality in schools, and children's life changes* (pp. 91–116). Russel Sage Foundation Press.

Reyes, J. A., Elias, M. J., Parker, S. J., & Rosenblatt, J. L. (2013). Promoting educational equity in disadvantaged youth: The role of resilience and social-emotional learning. In S. Goldstein & R. B. Brooks (Eds.), *Handbook of resilience in children* (2nd ed., pp. 349–370). Springer.

Rutter, M. (2006). Implications of resilience concepts for scientific understanding. *Annals of the New York Academy of Sciences, 1094*, 1–12.

Sandoval-Hernández, A., & Białowolski, P. (2016). Factors and conditions promoting academic resilience: A TIMSS-based analysis of five Asian education systems. *Asia Pacific Education Review, 17*, 511–520.

Sattler, K., & Gershoff, E. (2019). Thresholds of resilience and within- and cross-domain academic achievement among children in poverty. *Early Childhood Research Quarterly, 46*, 87–96.

Skolverket. (2013). *Skolverkets lägesbedömning 2013*. Rapport 387. Skolverket.

Telhaug, A. O., Mediås, O. A., & Aasen, P. (2006). The Nordic model in education: Education as part of the political system in the last 50 years. *Scandinavian Journal of Educational Research, 50*(3), 245–283.

Thapa, A., Cohen, J., Guffey, S., & Higgins-D'Alessandro, A. (2013). A review of school climate research. *Review of Educational Research, 83*(3), 357–385.

Ungar, M., Connelly, G., Liebenberg, L., & Theron, L. (2019). How schools enhance the development of young people's resilience. *Social Indicators Research, 145*, 615–627.

Wang, M., & Degol, J. (2016). School climate: A review of the construct, measurement, and impact on student outcomes. *Educational Psychology Review, 28*, 315–352.

Wang, M. C., Haertel, G. D., & Walberg, H. J. (1994). Educational resilience in inner cities. In M. C. Wang & E. W. Gordon (Eds.), *Educational resilience in inner-city America: Challenges and prospects* (pp. 45–72). Lawrence Erlbaum Associates, Inc.

Wayman, J. C. (2002). The utility of educational resilience for studying degree attainment in school dropouts. *The Journal of Educational Research, 95*(3), 167–178.

Wiberg, M. (2019). The relationship between TIMSS mathematics achievements, grades, and national test scores. *Education Inquiry, 10*(4), 328–343.

Xie, C., & Ma, Y. (2019). The mediating role of cultural capital in the relationship between socio-economic status and student achievement in 14 economies. *British Educational Research Journal, 45*, 838–855.

Zhu, Y. (2018). Equity in mathematics education: What did TIMSS and PISA tell us in the last two decades? In G. Kaiser, H. Forgasz, M. Graven, A. Kuzniak, E. Simmt, & B. Xu (Eds.), *Invited lectures from the 13th international congress on mathematical education* (ICME-13 monographs). Springer.

Chapter 12
Equity and Diversity in Reading Comprehension—A Case Study of PISA 2000–2018

Tove Stjern Frønes, Maria Rasmusson, and Jesper Bremholm

Abstract This chapter studies equity in reading performance in PISA 2000–2018 in three Nordic countries: Denmark, Sweden and Norway. Using regression analyses, the study investigates how the reading performance trend for groups of students with different genders, home backgrounds and minorities has developed. The study is contextualised through an up-to-date description of reading comprehension instruction in the countries. In addition to trend analyses of general reading performance, the study examines if the differences between groups of students are consistent across different text formats in the digital version of the PISA test, distinguishing between static text types (e.g., articles, letters, stories) and dynamic text types (e.g., websites, forums and e-mails, etc.). We find a consistently high reading literacy performance in all Scandinavian countries compared with international development. There are large gender differences in the average reading performance in all three countries, disfavouring boys, especially low-performing boys from low SES home backgrounds. We find a huge and stable gap between minority and majority students' reading achievement, even when corrected for SES. Taking these findings into account, we assert that there is no basis for concluding that the school systems give more equitable learning conditions for groups of students now than when the PISA assessments started. However, it appears that the new online text formats in PISA 2018 might shrink the differences between student groups. Based on our findings, we argue that it is highly doubtful if one can still speak of a Nordic model of education, both as an idea of equity and fairness and as a model that is united across the Nordic countries.

T. S. Frønes (✉)
Department of Teacher Education and School Research, University of Oslo, Oslo, Norway
e-mail: t.s.frones@ils.uio.no

M. Rasmusson
Department of Education, Uppsala University, Uppsala, Sweden

J. Bremholm
National Centre for Reading, Copenhagen, Denmark

In the Nordic educational systems, equity and equal opportunities are among the primary aims of schooling; they are based on the belief that if all students are given equal learning opportunities, equity will be obtained (Telhaug, Mediås, & Aasen, 2006). Indeed, the Nordic education model has been regarded as exemplary in ensuring social cohesion, justice and security, with equal access and learning opportunities for all (Telhaug, Aasen, & Mediås, 2004). However, several researchers have noted that it is debatable whether we can still talk about a joint Nordic model regarding policies, school choice and school competition (Klette, 2018) and that the differences in these areas starting in the early 2000s to today threaten the Nordic model (Lundahl, 2016). In this chapter, we examine the question of equity in the Nordic school system from the perspective of reading literacy because it is a centrepiece of basic schooling in Nordic countries and worldwide. Reading literacy is a key competency and life skill because a certain level of reading proficiency is needed to learn other subjects at school, undertake further education and work and participate in societal life (OECD, 2019a; UNESCO, 2004). Likewise, longitudinal studies indicate that students with insufficient levels of reading in the PISA assessment have a higher risk of not completing further schooling or education (Piacentini & Pacileo, 2019). In the current study, we perceive students' reading proficiency as an indicator of equity in the school systems, here by looking at the assessments of students' reading literacy in PISA. In the chapter, we set out to examine Nordic students' reading performance and the trend of reading performance development over time for different student population groups while also looking at the new text formats introduced in the reading assessment test in PISA 2018.

When the first results of PISA 2000 were published (OECD, 2001), it became clear that not all students in the Nordic countries were proficient readers; furthermore, the results revealed that in addition to individual variations between students, there were systematic differences in reading proficiency between groups of students (e.g., in relation to gender, socioeconomic status and language background; Andersen et al., 2001; Lie, Kjærnsli, Roe, & Turmo, 2001; Molander, Pettersson, Skarlind, & Taube, 2001). The PISA results were one factor leading to important changes in Nordic educational policies that moved towards improving the students' reading proficiency and reducing the proportion of low-performing readers (Mejding, 2019).

The PISA study design makes it possible to compare students' reading performance over time and between different groups of students. These comparisons provide valuable information on the extent to which the educational system in different countries supports equality and equity for students. Twenty years have passed since the first PISA study, so we consider it relevant and important to investigate the impact of the reforms and initiatives, examining whether they have led to changes

in student reading trends in the Nordic countries and whether these changes can be translated into a larger degree of equality and equity.

Regarding the text format for the reading assessment, significant changes were made to the test design in PISA 2018. In the previous PISA studies, the reading assessment was constructed based on traditional reading materials in single texts displayed in paper booklets. Although the digitisation of the PISA assessments began in PISA 2015, substantial changes did not take place in the reading domain before the development of new reading material for PISA 2018 (Støle, Mangen, Frønes, & Thomson, 2018). In PISA 2018, the assessment was expanded with two new text formats: *multiple texts* and *dynamic texts*. Multiple texts are several texts on the same theme organised into a text unit (OECD, 2019a). For our purposes, dynamic texts refer to interactive hypertext where the readers choose their own reading path, here including texts designed for the Internet and social media. Studies have shown that gender and home background seem to have less of an effect on students' performance in Nordic countries, among others, when reading digital texts compared with reading traditional texts (Fraillon, Ainley, Schulz, Friedman, & Gebhardt, 2014; Frønes & Narvhus, 2011; Olsen, Hatlevik, & Loi, 2015; Rasmusson, 2016). However, none of these studies have investigated whether the weakened effect on performance is covaried with the digital format, the text types or the reading tasks. Hence, it is relevant to investigate whether different groups read the new text formats in PISA differently or, in other words, if the new dynamic and multiple texts promote or hinder equity in the school system in a Nordic context.

In this chapter, we set out to answer two research questions. First, how has student equity in the Nordic countries, as indicated by reading performance in PISA, developed between groups of students of different socioeconomic (SES), gender and language backgrounds in the period 2000–2018? Second, do the new text formats in PISA 2018 (dynamic and/or multiple texts) strengthen equity in reading performance for the same groups of students?

A couple of aspects regarding the research questions need clarification. We focus on Norway, Sweden and Denmark as representatives of the Nordic school systems. We chose these three Scandinavian countries because of their similarities in language, culture, school systems and curriculum (Imsen, Blossing, & Moos, 2017). Finland and Iceland have less in common linguistically, culturally and regarding the school systems than the chosen Scandinavian countries.

Our understanding of equity in relation to reading literacy is in line with Espinoza (2007), whose ideas are explained in detail in Chap. 2 in this book. Thus, we consider that equity is obtained when students with similar abilities reach the same level of reading proficiency at a defined point in the educational system, here measured as educational achievement based on test performance (Espinoza, 2007, p. 353). In other words, the effect of home background and gender will be diminished in totally equitable educational systems, and the distributions of, for example, test scores would overlap between these subgroups. In the same way, equality is obtained when all formal obstacles (legal, political, social, cultural or economic) to achieve at the same level have been eliminated (Espinoza, 2007). As underlined by

Buchholtz, Stuart and Frønes in Chap. 2 of this book, the concepts of equity and equality are inextricably linked with the concept of diversity because diversity adds the perspective of 'being different but of equal worth' (Blossing, Imsen, & Moos, 2014, p. 7), which philosophically and ethically are necessary to approach equality and equity in educational contexts. They also point to the long Nordic tradition of legal rights to participate in free, public school for all students, unlike in many other school systems. In the present study, we include the perspective of diversity by focusing on and comparing different groups (gender and minority backgrounds) in the PISA population.

To contextualise the methodological and analytical parts of the study, we start the chapter with an outline of the major educational reforms and initiatives related to reading and literacy in Denmark, Sweden and Norway for the period 2000–2020; this is followed by a theoretical and research-informed account of online reading.

12.1 The Nordic Educational Context and Trends in Reading Development

Since the new millennium and the first PISA study (OECD, 2001), all three Scandinavian countries have witnessed significant changes to their educational systems that began with a number of political reforms and initiatives. Despite national differences regarding the specific nature of these reforms, according to Imsen et al. (2017), they share the same overall characteristics: a new and strong emphasis on competences, learning goals and learning outcomes, assessment and accountability with a corresponding downgrade of teaching, curricular content, and democratic *Bildung* (i.e., education, formation). Important here is that these characteristics are both part of and influenced by a strong general trend across Western countries in the first part of the twenty-first century (Antunes, 2012; Hodgson, Rønning, Skogvold, & Tomlinson, 2010; Moos, 2014; Sivesind, Akker, & Rosenmund, 2012); indeed, various Scandinavian scholars have analysed how this international reform trend poses serious challenges to the Nordic model of education (Imsen et al., 2017; Lundahl, 2016).

In all three Scandinavian countries, the national results of the first PISA studies, which placed the Scandinavian students around the OECD average, gave rise to disappointment and alarm, especially at the political level and among the public. This was popularly termed 'the PISA shock' (Mejding, 2019). Subsequently, the unsatisfactory national PISA results were a regular part of governments' arguments for the necessity of educational reforms, thus playing a legitimising role regarding these reforms (Imsen et al., 2017). A part of the ambition behind the reforms, as well as various other educational initiatives, have been to improve students' skills in the three subject domains tested in PISA (reading, mathematics and science). In the case of reading, a considerable number of different initiatives have been implemented in the Scandinavian countries over the past 20 years to improve literacy

instruction and students' literacy skills. A particular incentive behind most of these initiatives has been to reduce the number of students with insufficient reading skills (below level 2 in PISA) because the large proportion of students at this level in PISA 2000 challenged the values of equality and equity on which the Nordic educational model is based (Mejding, 2019).

Below, we enumerate the most important reforms and initiatives related to the domain of reading in the three Scandinavian countries since 2000.

12.1.1 Reforms and Initiatives in Denmark, 2000–2018

In this time period, three curricular reforms for compulsory school (grade Kindergarten to grade 9) have passed: in 2001 (Undervisningsministeriet 2001), 2009 (Undervisningsministeriet 2009) and 2014 (Undervisningsministeriet 2014 2014). All three reforms have been based on learning goals, and each has had a stronger emphasis on reading as part of the curriculum for Danish language arts. In the last and current reform in 2014, reading constitutes one of the four main competences for Danish as first language (L1) across all grade levels. In addition, reading and literacy have become a cross-disciplinary 'theme' for all subject areas and across all grade levels. The approach towards reading in the 2014 curriculum corresponds to a large degree to PISA's definition of reading literacy.

2006–2007. Introduction of a mandatory national test of reading and other subject areas (math, English and science). The students take reading tests in the 3rd, 6th and 8th grades, focusing on basic technical skills (based on the simple view of reading) yet aligning poorly with the national curriculum for reading and with PISA's conception of reading (Bremholm & Bundsgaard, 2019).

2007. Introduction of a national written exam in reading proficiency at the end of compulsory school (grade 9). The exam focuses on reading speed and basic technical skills, but alignment with the national curriculum and the PISA's definition of reading is weak (Bremholm & Bundsgaard, 2019).

2007–2009. Implementation of an in-service training programme for teachers to be certified as reading counsellors. The training programme is managed by the six Danish university colleges, and to begin with, the programme was supported by substantial governmental funding (Kuhlman & Rydén, 2011). Today, almost all compulsory schools in Denmark have a reading counsellor, and many schools have more than one (EVA, 2009).

2006 and 2012. The latest two reforms of national teacher education have put a stronger emphasis on reading and literacy, reading development and reading instruction. Furthermore, they have introduced grade-level specialisation, which includes reading. In Denmark, teacher education is regulated at the national level, and it is managed by the six university colleges across the country.

2006. The National Centre for Reading was founded by a governmental initiative. The purpose was to promulgate research-based knowledge on reading and

literacy to schools, teachers and teacher education, as well as to do research and developmental projects in the field of reading and literacy.

12.1.2   Reforms and Initiatives in Sweden, 2000–2018

2011. Following a new school law (SFS, 2010:800), the latest curricular reform for compulsory school (Lgr11) was introduced. This curriculum (and current) is based on proficiency levels and key subject matter content instead of learning objectives, as was the case in the previous curriculum from 1994 (Lpo94). Reading comprehension is given a much more prominent role in the 2011 curriculum compared with its predecessors.

2013 and onwards. Initiation of *Boost for Reading* [*Läslyftet*], an in-service training literacy programme for teachers. The programme was organised by the Swedish National Agency for Education and was fully implemented between 2015 and 2018 (Carlbaum, Andersson, & Hanberger, 2016). In 2017, about 30,000 teachers had enrolled in the programme.

2015 and onwards. Implementation of *Cooperation for Better Schools* [*Samverkan för bästa skola*], a governmental initiative to support low-performing schools with an explicit aim to raise student achievement. The initiative is led by the Swedish National Agency for Education. By 2019, 252 schools have been involved in this school development project.

2017. Adoption of an amendment to the school legislation to digitalise the national tests and strengthen the influence of the tests on the students' grades to increase equity in grading. In Sweden, all students take national tests in both reading and writing in Swedish language arts in the 3rd, 6th and 9th grades. This digitisation and new framework for assessment is planned to take effect in 2022.

2019 and onwards. A guarantee for early support was added to Swedish school law. Schools are obliged to map the students' reading, writing and mathematical abilities in the preschool class and in the first grade to ensure that students with special needs will get support at an early stage in their schooling.

12.1.3   Reforms and Initiatives in Norway, 2000–2018

2003 and onwards. Increased emphasis on students' early reading development through close monitoring by teachers and the use of mapping tests (Roe, 2012). Students who show signs of reading or numeracy difficulties receive help at an early stage, and starting in 2018, a responsibility to provide intensive instruction for students in danger of being left behind in the 1st to 4th grades was established by law.

2003 and onwards. A number of national reading initiatives have been implemented. The first and costliest, *Opportunities to Read* [*Gi rom for lesing*!], was launched by the Ministry of Education in 2003 and completed in 2007. The main

goals were to improve the reading skills of children and adolescents, motivate them to read more, strengthen teachers' competence in literacy education and raise awareness of reading as a gatekeeper for learning, cultural competence, quality of life and community participation (UFD, 2003).

2003. Several educational research centres were established, among these the Norwegian Reading Centre [*Lesesenteret*] and the Norwegian Centre for Writing Education and Research [*Skrivesenteret*].

2006. Implementation of a comprehensive national curriculum reform (LK06) known as the *Knowledge Promotion Reform* [*Kunnskapsløftet*] (Aasen et al., 2012). The Knowledge Promotion Reform is often characterised as a literacy reform because of its explicit focus on the use of oral and written language as tools in all subjects (Berge, 2005).

2007. The *Quality Assessment System* (NKVS) was established as a part of the Knowledge Promotion Reform, and national reading tests were developed and implemented with an explicit focus on the formative role of the tests (Jensen, Frønes, Kjærnsli, & Roe, 2020). From 2007 onwards, all students at the beginning of the 5th and 8th grades take national tests in reading, numeracy and English.

2010. Initiation of *Assessment for Learning* [*Vurdering for læring*], a nationwide initiative where school owners, schools and learning enterprises receive support to further develop their assessment culture. Assessment for learning was introduced as an educational principle and as part of the Knowledge Promotion Reform in 2006. It promotes criterion-based assessment, linking the criteria to curriculum goals and with the characteristics of mastery levels.

12.1.4   Trends in Reading Development

Despite the national differences, there are interesting common traits behind the initiatives and reforms. We argue that these traits can be characterised as an embedded or integrated approach to literacy instruction as opposed to the approach applied before 2000, which considered reading as primarily a technical skill pedagogically limited to the primary grades. The embedded approach to literacy considers reading and writing as an integrated part of all subjects and all communicative practices across grades. We find aspects of this approach towards literacy instruction in elements such as literacy-based curriculum reforms, the introduction of standardised and validated tests and mapping tools for formative assessment, early efforts, reading stimulating campaigns and the widespread use of reading counsellors.

The brief descriptions of the educational context in the Scandinavian countries will be used in this chapter to discuss reading literacy development. Likewise, in the final part of this section, we give a quick overview of the PISA results in reading for Denmark, Norway and Sweden as background knowledge for the analyses. In Fig. 12.1, the overall performance in reading literacy in Denmark, Norway and Sweden from 2000 to 2018 is shown.
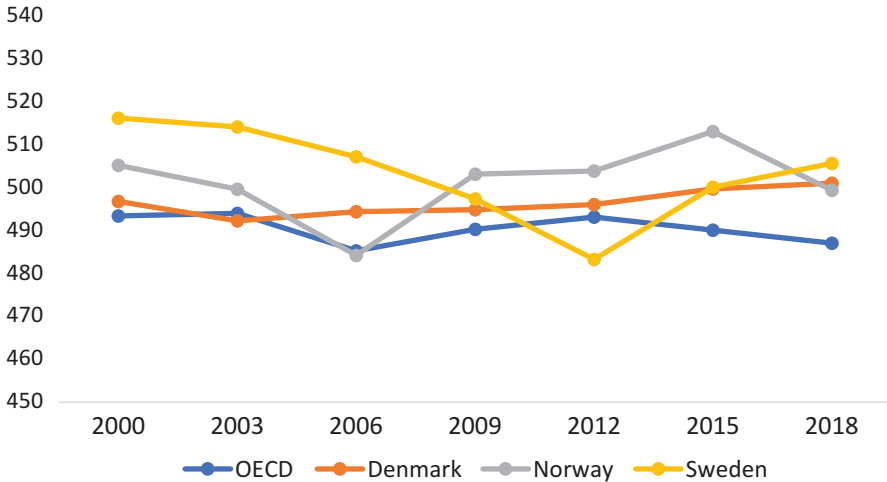
Fig. 12.1 Average performance in reading literacy in PISA 2000–2018 (see also, e.g., OECD, 2019b)

Table 12.1 Average performance in reading literacy by gender in PISA 2000–2018

		Female		Male		
		Average	S.E.	Average	S.E.	Diff F-M
2018	OECD	502	0.5	472	0.5	30[a]
	Denmark	516	2.3	486	2.3	30[a]
	Norway	523	2.6	476	2.6	47[a]
	Sweden	523	3.4	489	3.2	34[a]
2009	OECD	510	0.5	471	0.6	39[a]
	Denmark	509	2.5	480	2.5	29[a]
	Norway	527	2.9	480	3	47[a]
	Sweden	521	3.1	475	3.2	46[a]
2000	OECD	510	0.8	478	0.9	32[a]
	Denmark	510	2.9	485	3	25[a]
	Norway	529	2.9	486	3.8	43[a]
	Sweden	536	2.5	499	2.6	37[a]

[a]$p < .05$

As shown in Fig. 12.1, Denmark, Norway and Sweden have statistically significantly higher results than the OECD average in PISA 2018 and with no significant difference between the three countries. According to recent trend analyses, Norway and Denmark are among the few OECD countries that have stable performance close to the OECD average in all PISA cycles, while Sweden has a negative trend line (Jensen et al., 2020; OECD, 2019b).

In all three countries, the performance differs between groups of students. Table 12.1 gives an overview of boys' and girls' average performance for all PISA cycles.

In all three countries, as well as for the OECD average, girls perform statistically above boys in reading literacy. In PISA 2018, girls in all three countries performed significantly higher than the OECD average for girls. The same was the case for Danish and Swedish boys, who performed above the OECD average for boys. Norwegian boys performed at the OECD average.

12.2 New Reading Challenges in the Digitised World

According to Coiro (2003), reading and understanding online texts can set new literacy practices in motion, and when this occurs, readers need to activate both traditional and fundamentally new thought processes. Expert readers use their usual strategies when reading online: they activate prior knowledge on the text and topic, identify the main themes and monitor their own understanding (Coiro, 2011). In addition, good readers are experts in doing web searches, reviewing search results and managing and comparing multiple text representations. In this section, we point to previous research on how online reading is related to the features of text, the reader's cognitive processes, prior knowledge and ability to spatial orientation and the reader's reading comprehension strategies.

Texts organised as hypertext impose a greater cognitive burden on readers, and the ability to effectively use strategies is crucial to avoid cognitive overload and, thus, confusion and disorientation (Lawless & Kulikowich, 1996; Shapiro & Niederhauser, 2004). Theories of cognitive flexibility have indicated that the lack of a supportive structure in dynamic texts raises the demands on the reader, who must devote more resources and metacognitive effort to adapt to new and ever-changing texts with multiple representations of information (Coiro, 2011; Spiro, Feltovich, Jacobson, & Coulson, 1992; Spiro, Klautke, & Johnson, 2015). Wylie et al. (2018) showed how the reading of online dynamic texts puts additional demands on executive functions, potentially threatening comprehension and learning because of shallow processing. Extensive research on the additional evaluation and sourcing processes related to reading dynamic texts has agreed that such processes raise the demands on the reader (e.g., Bråten et al., 2011; Kiili, Laurinen, & Marttunen, 2008; Salmerón, Strømsø, Kammerer, Stadtler, & van den Broek, 2018).

Other studies have shown that readers need to develop corresponding online comprehension strategies. When reading hypertext online, the reader encounters layers of 'possible links, possible texts, possible decisions and possible interactions' (Afflerbach & Cho, 2009, p. 81). It is clear that even proficient readers with satisfactory reading strategies for single and static texts experience the interaction with the text as more demanding and complex. Afflerbach and Cho pointed to three areas where the reading process of static and dynamic texts differ: (a) the process of constructing a text while reading, (b) the need for strategies that can help manage the information load on the working memory and (c) special strategies for self-regulation (2009, p. 81).

As mentioned in the introduction, the change in the delivery mode starting with PISA 2015 has led to new text types in the reading assessments, with texts inspired by online genres that can be labelled as dynamic texts. In PISA 2018, dynamic texts were a part of the regular reading assessment for the first time (OECD, 2019a). In addition, *text source* was introduced as a text format dimension, dividing single texts from multiple texts (several texts from different sources on the same topic). A multiple text unit might contain unique, overlapping and/or conflicting information and incorporate reading processes such as evaluating the veracity of texts, seeking information, detecting and evaluating conflicting information and integrating/synthesising information across sources (OECD, 2019a, p. 24). By incorporating these new text formats, a number of new genres were also introduced in the PISA 2018 reading assessment, including webpage, online forum, e-mail, blogs, newspaper, online search and chat. With our second research question, we examine if digital reading as represented by the new text formats in PISA 2018 influences equity regarding the students' reading performance. Before we present results for the first research question – how reading performance in PISA has developed between groups of students – we will account for the methods used.

12.3   Methods

Measuring equity in educational systems in general is a complex issue (Chap. 3)., and it is not investigated thoroughly enough by only reporting achievement gaps. Therefore, in our study, we have taken SES, gender and minority background into account. Furthermore, we have specifically focused on equity aspects in the field of reading literacy. We argue that the subject-specific aspects of equity are important to consider, and in the field of reading literacy, the recent change towards more digital reading needs to be appraised.

In this section, we firstly provide an account of the PISA data and of the Danish, Norwegian and Swedish samples used in the current study. Furthermore, we describe the analytical tools and procedures we applied, along with our methodological choices and reflections.

12.3.1   PISA Data

The major domain in the PISA studies shifts between the three standard subject domains (reading, mathematics and science) at each 3-year cycle. Hence, reading is the major domain every ninth year, and we have chosen to use data from these cycles: PISA 2000, 2009 and 2018. In PISA, the students' performance in reading literacy is reported as plausible values and computed as a proficiency distribution around a reported value by assigning a set of values drawn from this distribution (OECD, 2009). This method reduces errors in the analysis on the population level (Braun & von Davier, 2017; Rutkowski, Gonzalez, von Davier, & Zhou, 2014). In the current study, we used the plausible values for students' performance on the overall reading

performance and for the subcategories of reading multiple and single texts. In addition, an index of students' socioeconomic background was used. In PISA 2000, this index was the international socioeconomic index of occupational status (HISEI), which is derived from items on parents' occupation in the student questionnaire. In the following PISA studies, the new index of economic, social and cultural status (ESCS) was used; this index is derived from a number of items in the student questionnaire about parents' education and occupation, home possessions (such as possession of a car, the existence of a quiet room to work, access to the Internet, the number of books and other educational resources). In sum, we used the ESCS index, immigrant background and gender from the student questionnaire (OECD, 2019c).

In addition to the plausible values in reading literacy and background variables, we wanted to analyse new text formats in PISA 2018. Plausible values for static and dynamic items were not available, so we used the proportion of items answered correctly and omitted items. Because of a new multistage adaptive test (MSAT) implemented in PISA 2018 for the computer-based reading assessment, the values for the correct proportions are computed in a different way than before. These new equated proportion correct statistics were used to compare the performance on items classified as dynamic and static for Denmark, Norway and Sweden. The main idea behind MSAT is that students will have to answer fewer items, but the items they answer are better adjusted to their proficiency level. In total, the test included 245 reading items belonging to 45 units in three blocks. The new method for computing the equated correct proportion was based on item response theory and mean deviation statistics (ETS, 2019; OECD, 2020).

12.3.2   Sample

Table 12.2 presents an overview of the samples used in PISA 2000, 2009 and 2018 in the three countries. Hereafter, the first- and second-generation students are labelled as minority students and the native students as majority students. Since

Table 12.2   The samples in the three countries in PISA 2000, 2009 and 2018

	All	Female	Male	Native/ Majority	Second generation	First generation	Minority
2000							
Denmark	4212	2099	2113	3835	85	198	283
Norway	4082	2014	2068	3746	74	193	267
Sweden	4383	2153	2230	3828	195	280	475
2009							
Denmark	5924	3038	2886	4478	931	358	1289
Norway	4660	2285	2375	4305	167	146	313
Sweden	4567	2256	2311	3993	344	163	507
2018							
Denmark	7657	3816	3841	5858	1282	269	1551
Norway	5813	2880	2933	4882	348	347	695
Sweden	5504	2763	2741	4283	556	499	1055

2009, Denmark has had an oversample of immigrant students (Beuchert & Christensen, 2019; Egelund, 2010). In PISA 2000, there was not an index labelled 'immig', but instead, we computed an index from questions in the student questionnaire in PISA: Was the student born in the country? Was the mother and/or the father born in the country? According to PISA, native students are those born in the country in which they were assessed by PISA or who have at least one parent who was born in that country. Immigrant students are those with an immigrant background, and they can be either first generation (those who are foreign born and whose parents are also foreign born) or second generation (those who were born in the country of assessment but whose parents are foreign born) (OECD, 2011, p. 1).

12.3.3   Analyses

The current study comprises groupings based on different criteria, including gender (boys and girls), socioeconomic and language background (majority and minority) and a case analysis of the trend development in the three Nordic countries. The analyses include both average reading results for the groups through PISA 2000, 2009 and 2018—when reading was the main area of research—and their subscores on text types in PISA 2018. The analyses were performed using Stata, SPSS, IEA IDB Analyzer and PISA Data Explorer. Using descriptive statistics and regression analysis, estimates of the contribution of gender and immigrant background to the overall performance in reading literacy and performance on multiple and single texts were separately calculated for each country. In the models, socioeconomic status (the index HISEI in PISA 2000 and ESCS in the following PISA surveys) was considered. There is reason to exhibit caution when comparing these indicators across countries and over time (OECD, 2019c). Studies have shown that a comparison raises several challenges (Rutkowski & Rutkowski, 2013, 2017), so we have chosen not to compare ESCS trends between cycles but rather to compare the contribution of socioeconomic background to reading literacy performance in separate regression models. The standardised beta (β) coefficients were used to estimate the difference between the regression models. The β coefficient gives an estimate of the strength of the effect of each individual independent variable to the dependent variable. The higher the absolute value of the beta coefficient, the stronger the effect. To answer RQ 2, we used the equated proportion's correct values for each item. The reading items were classified as either static or dynamic. The average proportion of both correct answers and omitted tasks for the items categorised as static and dynamic, respectively, was computed per country. In the analyses of the proportion of correct and omitted items, we used descriptive statistics to compare the dynamic and static items answered correctly or that were omitted for each country.

12.4   Results

In this section, we present the results for each of the two research questions. In the first part, we inspect the trend development for different groups of students based on gender, home background and immigrant status (RQ1). In part two, we compare the reading results from PISA 2018 in the new and old formats: dynamic items vs. static items andmultiple vs. single texts (RQ2).

12.4.1   Main Trends for Groups: Gender Differences Controlled for SES

Girls outperformed boys in all three countries and for the OECD average in all PISA cycles. This is the case in most participating countries. In the first PISA survey in 2000, the socioeconomic index (HISEI) had a similar association with performance in reading literacy in the three countries when gender was accounted for. As seen in Table 12.3, Norway has a slightly smaller β-value (0.28) than Denmark and Sweden. Girls performed better than boys in all three countries, but the disadvantage to boys was smaller in Denmark (β=−0.14).

In PISA 2009, the socioeconomic indexes (HISEI in 2000 and ESCS in 2009 and 2018) had a larger association with the reading results in Denmark and Sweden than in PISA 2000 when gender was accounted for. In Norway and Denmark, the negative effect of being a boy increased compared with in 2000. However, the association between SES and reading performance decreased in all three countries in 2018 compared with 2009 when gender was accounted for (see Table 12.3). Moreover, the disadvantage for boys when SES is accounted for also decreased in 2018 compared with 2009 (see Fig. 12.2).

Table 12.3 Regression analysis with plausible values in reading as the dependent variable, PISA 2000, 2009 and 2018

		Constant	SES B	SES β	Boy B	Boy β	R^2
2000	Denmark	424.95	1.82	0.31	−25.62	−0.14	0.11
	Norway	430.94	1.85	0.28	−43.53	−0.21	0.12
	Sweden	450.79	1.72	0.31	−38.01	−0.21	0.13
2009	Denmark	500.09	36.65	0.38	−29.81	−0.18	0.18
	Norway	511.62	37.20	0.30	−49.21	−0.27	0.16
	Sweden	507.82	43.74	0.37	−45.12	−0.23	0.19
2018	Denmark	497.04	37.90	0.31	−28.78	−0.16	0.12
	Norway	505.75	35.14	0.27	−46.09	−0.22	0.12
	Sweden	510.47	38.76	0.33	−31.59	−0.15	0.13

Note: All coefficients are statistically significant at p<.05
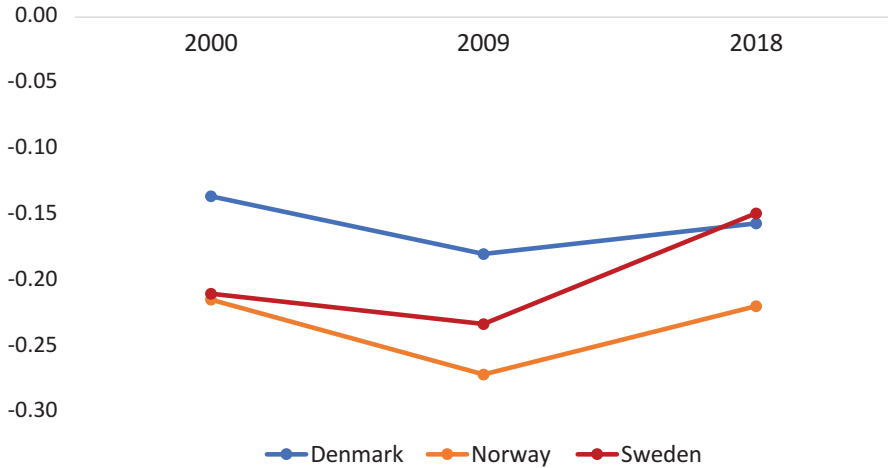
Fig. 12.2 The standardised b value (β) for the effect of gender (negative effect for boys) on reading performance when SES is accounted for

12.4.2 Main Trends for Groups: Language Background

Denmark, Norway and Sweden have students with a first language that differ from the language tested in the PISA reading assessment. Because linguistic comprehension skills such as vocabulary and grammar are closely associated with reading comprehension, the performance gap between language groups could be an indicator of equity. To investigate how the three school systems support the development of reading literacy in Danish, Norwegian and Swedish, we looked further into students with a majority background (native students) and minority background (first- and second-generation students).

In Fig. 12.3, there is a large performance gap between the majority and minority students in 2000, 2009 and 2018. This gap is larger in Sweden in 2018 compared with Norway and Denmark, and it is larger than the previous gaps in Sweden. There is a need for a cautionary note here because of the small sample of minority students in the Norwegian and Swedish samples. Denmark, however, oversampled minority students (first and second generation) in 2009 and 2018 and obtained a sample large enough to generalise from.
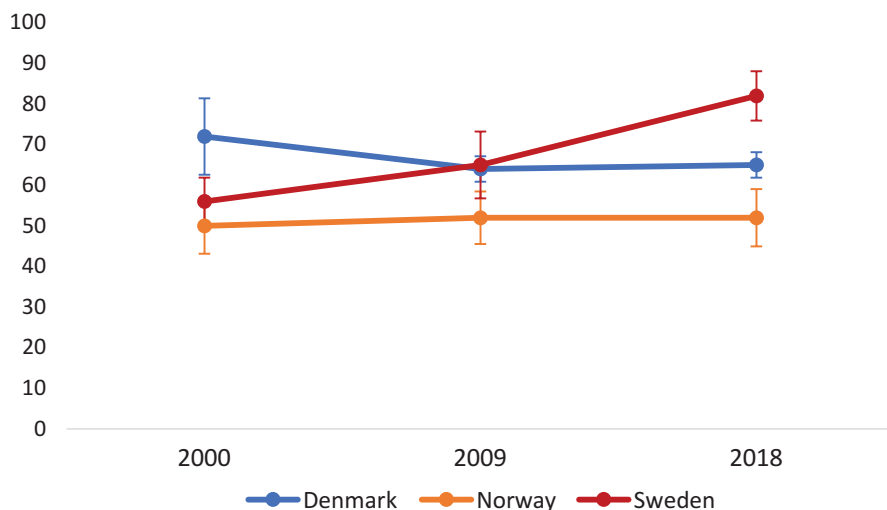
Fig. 12.3 The difference between majority and minority students' reading performance in PISA 2000, 2009 and 2018. Confidence interval (95%): Difference +/− 1.96 * S.E

12.4.3   Main Trends for Groups: Minority Students from Underprivileged Backgrounds

Being a minority student in Scandinavian countries often covaries with low socio-economic background. Many of the newly arrived immigrants have few home possessions, and their parents have not yet entered the workforce. Even if the relationship between both home background and performance and language background and performance is fairly stable, there are still many students with socioeconomically disadvantaged backgrounds who succeed in school (Masten, 2018). Some students from lower SES homes and with non-native language backgrounds are among the middle and top performers. These students are commonly labelled academically resilient because they are successful in school despite being situated in an environment linked to poorer outcomes (Martin & Marsh, 2006). Figure 12.4 shows the average reading performance for minority students in the bottom quarter of SES.

Figure 12.4 indicates that in PISA 2000, the Swedish minority students in the bottom quarter of SES had the highest reading performance among the three countries but dropped to having the lowest performance in 2018. Even though the only statistically significant difference is between the Danish and Swedish students in 2000, in all three countries, there are differences between years. Danish students have a higher average in PISA 2018, while Norway has had the most stable trend. Annex B shows the proportion of minority students in the lowest SES quarter in PISA 2000, 2009 and 2018.

When comparing the β-values for minority students in Table 12.4, the negative effect on reading performance has been stable in PISA 2000, 2009 and 2018 (β =
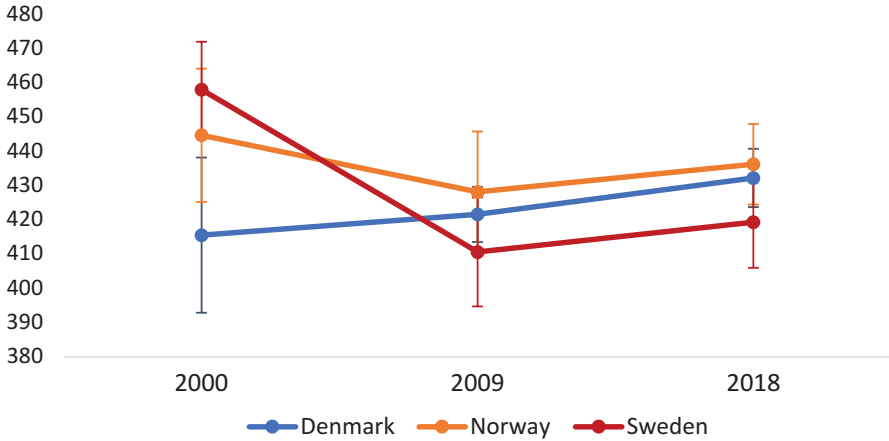
Fig. 12.4 Average reading performance for minority students, bottom quarter of SES. Confidence interval (95%): Difference +/− 1.96 * S.E

Table 12.4 Regression analyses of the effect of minority background on reading performance when SES is accounted for in all three countries in PISA 2000, 2009 and 2018

		2000		2009		2018	
		B	β	B	β	B	β
Denmark	Constant	418.95		489.63		489.83	
	SES	1.77	0.30	33.68	0.35	33.21	0.28
	Minority	−55.95	−0.15	−35.75	−0.12	−40.63	−0.14
	R²	0.12		0.16		0.12	
Norway	Constant	413.91		490.70		489.28	
	SES	1.80	0.27	34.01	0.28	32.19	0.25
	Minority	−35.13	−0.08	−32.87	−0.09	−31.51	−0.10
	R²	0.08		0.09		0.09	
Sweden	Constant	441.22		491.70		511.13	
	SES	1.63	0.29	39.84	0.34	30.68	0.26
	Minority	−42.07	−0.14	−39.95	−0.13	−61.22	−0.23
	R²	0.11		0.15		0.16	

Note: All coefficients are statistically significant at p<.05

0.08–0.15). The exception was Sweden in PISA 2018, where the disadvantage for minority students increased dramatically. In Norway, the gap between majority and minority students is smaller compared with Denmark and Sweden in all three PISA surveys. The explained variance (R^2) in the models differs among the countries. The three regression models for Norway have a lower level of explained variance; thus, SES and minority background have a smaller influence on students' performance in reading than in Denmark and Sweden. The results for Denmark in 2000, Norway (all years) and Sweden (all years) must be interpreted with caution because of the small sample sizes of minority students.

12.4.4 Text Effect: Student Diversity When Reading Dynamic or Multiple Texts

To answer our second research question, we conducted a number of analyses of the effect of dynamic/static and multiple/single texts on students' reading performance and of the differences between groups.

To investigate the possible difference between items with static texts compared with items with dynamic texts in PISA 2018, the average proportion of correct answered items was calculated. No clear pattern was evident. The static and dynamic items were correctly answered by approximately the same share of students in all three countries (see Table 12.5).

We also compared the equated proportion of correctly answered items with single and multiple texts. The multiple text items and the single text items were answered correctly to the same degree in all three countries (see Table 12.6). However, when we computed the average performance using the plausible values for girls and boys, it became evident that the boys performed particularly well on multiple items (Fig. 12.5). In Norway, the girls stood out because they performed almost equally well on the two item types. In all three countries, the boys showed a larger margin than the girls between their performance on multiple and single items.

We also compared students' performance on items with single and multiple texts. The students performed slightly better on multiple text items than on single text items in all three countries, with the biggest difference being found in Sweden. The Swedish students performed better than the Norwegian and Danish students and better than the OECD average (see Table 12.6). Figure 12.5 shows that the Swedish boys performed particularly well on the multiple items. In Norway, the girls stood out because they performed almost equally well on the two item types. In all three countries, the boys showed a larger margin than the girls between their performance on multiple items and on single items.

When SES is accounted for in the analysis of gender differences for multiple and single texts, the difference shown in Fig. 12.5 remains. The gender gap was found to be smaller for multiple texts than for single texts in all three countries. In other words, the boys' disadvantage is smaller for multiple texts than for single texts when SES is accounted for (see Table 12.7).

As Table 12.7 shows, boys performed better on multiple texts than single texts when SES was accounted for.

Table 12.5 Average equated proportion of correct answered static and dynamic items in PISA 2018

	Static (n=86)	SD	Dynamic (n=70)	SD	Diff.	S.E.
DNK	0.656	0.20	0.658	0.18	0.002	0.03
NOR	0.641	0.18	0.645	0.18	0.004	0.029
SWE	0.653	0.19	0.65	0.18	−0.003	0.029

Note: No significant differences between the static and dynamic texts

Table 12.6 Average equated proportion of correct answered items (P-values) and plausible values for single and multiple text items in PISA 2018 (see also e.g. OECD, 2019b)

	Single items				Multiple items			
	P-value (n=61)	SD	Plausible values	S.E.	P-value (n= 95)	SD	Plausible values	S.E.
DNK	0.644	0.19	496	2	0.665	0.19	503	1.8
NOR	0.638	0.18	498	2.4	0.645	0.18	502	2.3
SWE	0.648	0.18	503	3.1	0.655	0.18	511	3.1

Note: No significant differences between the single and multiple items
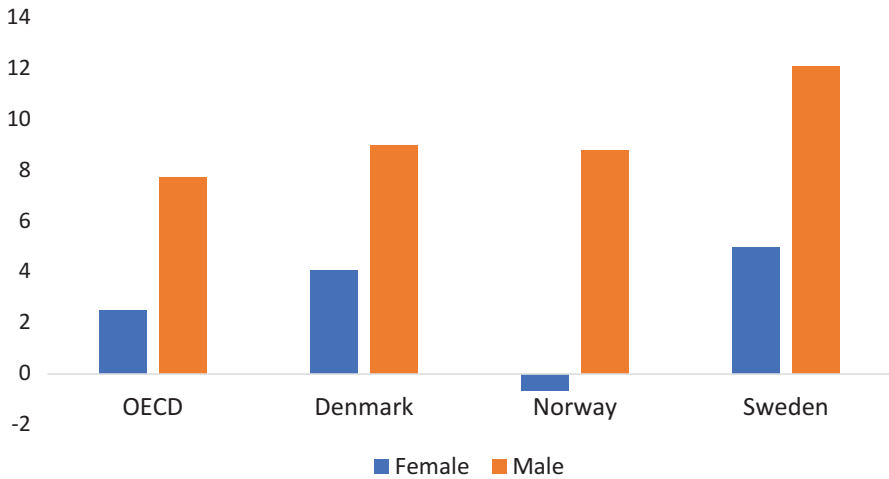
Fig. 12.5 Difference in reading performance between multiple and single texts by gender in PISA 2018

12.4.5   *Text Effect: Students' Coping Strategies Through Task Omission*

To give nuance to the analyses, we also analysed the students' omission of tasks. We consider omission as a strategy the students applied to cope with hard items, which is somewhat opposite of strategic flexibility. A larger percentage of the dynamic items than the static items are omitted in all three countries, and the same is true for the tasks connected to multiple texts (see Fig. 12.6). However, the performance was higher on multiple items than on single items in all three countries, despite the larger share of omitted multiple items.

A larger percentage of dynamic items than static items are omitted in all three countries, and this is the case for both single and multiple items. However, the results for single dynamic items should be interpreted with caution because they only include four items.

Table 12.7 Separate regression analysis of gender differences for multiple and single texts when SES is accounted for in PISA 2018

		Single		Multiple	
		B	β	B	β
Denmark	Constant	492.6		497.5	
	SES	40.0	0.32	38.4	0.32
	Boys	−31.8	−0.17	−26.7	−0.14
	R²	0.13		0.12	
Norway	Constant	504.9		505.3	
	SES	37.8	0.29	35.0	0.27
	Boys	−49.7	−0.23	−40.1	−0.19
	R²	0.13		0.11	
Sweden	Constant	508.3		513.4	
	SES	39.2	0.33	39.0	0.32
	Boys	−33.6	−0.16	−26.5	−0.12
	R²	0.14		0.12	

Note. All coefficients are statistically significant $p < .05$
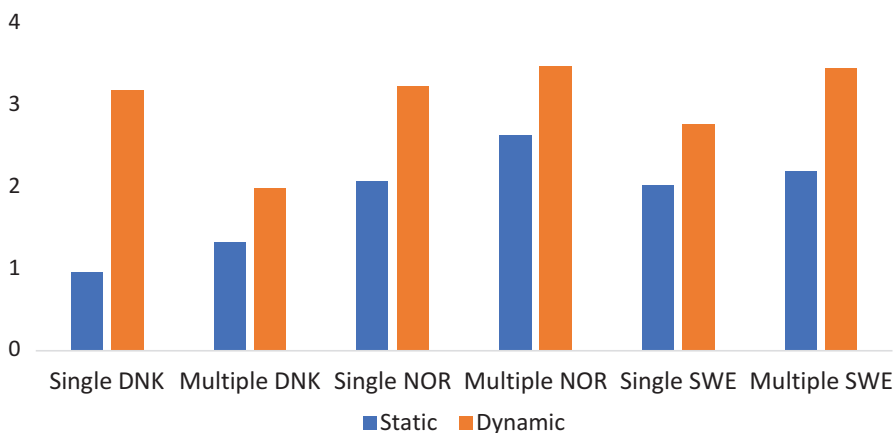
Fig. 12.6 Average percent omitted static/dynamic and multiple/single text items. Annex C also describes the percentage of omitted items

12.5   Discussion

There are four main findings that we want to highlight. First, we found a consistently high reading literacy performance in all the Scandinavian countries compared with international development although the Swedish trend between PISA 2000 and 2018 is slightly negative. Second, there are large gender differences in the average reading performance in all three countries, and in PISA 2018, the difference disfavouring boys is particularly large in Norway. Third, there is a huge and stable gap between minority and majority students' reading achievement, even when

correcting for SES. There has been a marked development in Sweden, with distinctly weaker reading results for this group of students in PISA 2018. Thus, the effect of home background is similar, and there is no reason to conclude that the school systems give more equitable learning conditions for groups of students now than when the PISA assessments started. However, fourth, it appears that the new online text formats might shrink the differences between student groups; albeit, at the same time, we also see a larger proportion of students skipping these items.

12.5.1   Equity and Reading Literacy Opportunities

There is reason to positively interpret the stable trend of Scandinavian students' reading skills since 2000. Indeed, the pervasive digitisation of society has given students' reading interest completely different preconditions than for previous generations. In addition, the trend in many participating OECD countries and the international average in PISA have had a negative trend throughout the same period. The decline in the international average still applies, even if we focus on the 27 original OECD countries that participated in PISA 2000 and in all subsequent cycles (Jensen et al., 2020). PISA 2018 also indicates that reading interest and habits have dropped dramatically during this period, and here, it is remarkable that the reading results have not fallen in line with this (OECD, 2019b). The most obvious theory as to why the Scandinavian countries have managed to achieve stable results is that high-quality reading instruction is given at school. However, as the results also show, the measures taken to ensure high-quality education do not seem to affect all students.

Even though Imsen et al. (2017) found that all three countries have had a comparable educational development emphasising learning outcomes, assessment and accountability, there is also reason to emphasise the renewed weight on embedded literacy education. We found common traits among the educational initiatives in Denmark, Sweden and Norway, where reading and writing instruction are integrated in the school subjects and are supported by pervasive implementation in the literacy practices of teachers and schools. We have identified various educational initiatives aimed at embedded literacy, such as literacy-based curriculum reforms; the introduction of standardised and validated tests and mapping tools for formative assessment; early efforts to identify students at risk; reading campaigns for engagement; and the widespread use of reading counsellors. We cautiously conclude that all these measures are probably related to the stable and high reading performances and that there is good reading education in Scandinavia. However, there are nuances. Both Denmark and Norway initiated earlier and similar measures, while Sweden had curriculum reforms at a later stage and to a lesser extent. The Swedish in-service training programme in literacy for teachers, *Läslyftet*, is unique in the Scandinavian context, but as an evaluation has shown (Skolverket, 2020), it is not considered sufficient in terms of being a rigorous implementation ensuring comparable effects across schools and municipalities.

From an equity perspective, this positive general picture of reading literacy in the Scandinavian countries is being nuanced when considering and comparing the

reading performance of majority and minority students. When it comes to the performance gap between majority and minority students in PISA 2018, Sweden stands out by having the lowest minority performance and a larger gap compared with Norway and Denmark. In Norway, the gap between majority and minority students is smaller compared with Denmark and Sweden in all three PISA surveys. The regression analyses also show that being a minority student has a stable negative effect on reading performance, except in Sweden in PISA 2018, where we find a dramatically larger disadvantage for minority students. There is a remarkably small gap between the groups of students in Norway that is stable over time. Even though the sample sizes are small, there is reason to put weight on these findings. Thus, the results of our study show that equity related to language background has not improved in any of the three countries between 2000 and 2018; they also indicate that Norway is doing markedly better than Sweden and Denmark, while in the case of Sweden, they indicate a weakening tendency regarding equity, which should raise concern.

Another factor to consider when discussing underprivileged minority groups is the size and composition of the countries' minority populations. In this respect, Sweden differs from the other countries. In Sweden, there has been a stepwise increase in immigration, with an especially high number of immigrants starting in 2013 (Swedish Migration Agency, 2020). Even though the same general movement can be found in Denmark and Norway, with a steady rise of immigrants peaking after the crisis in 2015 (Statistics Denmark, 2020; Statistics Norway, 2020), it is far from proportional to the rise in Sweden. Sweden also has the highest number of humanitarian migrants in the 2009–2018 period, while Denmark and Norway have a higher proportion of immigrant workers. In all OECD countries, humanitarian migrants have difficult integration processes (OECD, 2015). However, the number of immigrants or the different reasons for migration cannot be treated as a matter of equity. With a considerable number of newly arrived immigrants in all three countries, the crucial question is as follows: How do the systems compensate for these underprivileged students in school?

In all three countries, newly arrived students are typically enrolled in school introduction programmes, most commonly after some time and after obtaining a residence permit. Most of these student programmes last for up to 2 years and have intensive language training. In Sweden, the introduction programme is decentralised and differs between the municipalities; newly arrived students are sometimes placed in an ordinary class and sometimes in preparation classes.

As pointed out above, Norway stands out as having a higher degree of equity in reading performance regarding language background. A possible contributing cause for this, we argue, could be that Norway remains most in line with the traditional Nordic model of schooling. As touched on briefly in the introduction, Lundahl concluded that it is highly doubtful whether one can still speak of a Nordic model of education when considering the development in Sweden from the perspective of extensive marketisation and privatisation practices (2016, p. 9). Likewise, Klette (2018) discussed how the emergence of new models of individualism and competition in both private and public schools in the Nordic countries pose a challenge to

education as a foundation for a cooperative and fair society. She found that although all the Scandinavian countries have a strong decentralisation of school governance, there are some differences (2018, p. 67). Denmark and Sweden stand out by having free choice of schools, while this is only possible in some municipalities in Norway. In both countries, this leads to educational segregation because many students enrol in private schools rather than in local neighbourhood schools. Norway, by contrast, has eschewed this tendency towards increased educational segregation and, thus, remains most in line with the traditional Nordic model of schooling. It is not unlikely that having less educational segregation could be part of the reason why Norway has a higher degree of equity in reading performance between majority and minority students. Further studies are needed to examine this hypothesis.

12.5.2   *Equitable New Reading Challenges?*

To answer our second research question, we compared students' performance when reading traditional texts in PISA 2018 with their reading of dynamic and multiple texts. We did not find any average performance difference between dynamic and static texts in any of the three countries. However, the students performed slightly better on multiple text items than on single text items. We found the greatest performance difference between the text types in Sweden, where the students performed better on the multiple items than Norwegian and Danish students and the OECD average. The gender differences were smaller for multiple texts than for single texts in all three countries when SES was accounted for, and boys were less disadvantaged when reading texts in new formats. Thus, our results indicate that the new digital formats strengthen equity in reading performance, reducing the gender difference between boys and girls, which has been a constant throughout all PISA surveys. Please note that we have not investigated differences between reading on paper or screen, only reading different genres on screen. However, although reading in new formats seems to give more equitable conditions, students' completion of tasks varied considerably. We treated student omission of items as a student strategy for coping with hard items, and in all three countries, a larger percentage of the dynamic items and multiple items were omitted compared with the static and multiple items, respectively.

Indeed, the impact of SES weakens in online reading and digital competence compared with traditional reading proficiency studies (Frønes & Narvhus, 2011; Olsen et al., 2015; Rasmusson, 2016), and our findings confirm this. However, why does reading online texts in new formats place students in a more equitable learning situation? Most commonly, discussions centre around access to computers and the Internet, how often and for what purpose students use the devices and their engagement with online text types.

Most students in Scandinavian countries have access to these new forms of reading material. In Norway and Denmark, access to computers and the Internet in schools has been a strong political priority for over two decades. PISA 2009 showed

that access to both PCs and the Internet was at a very high level in Norway and Denmark—the highest in an international context—without this being crucial to how students performed on the online reading test (Frønes & Narvhus, 2011; Mejding, 2011). Even though deploying computers in Swedish schools occurred later, by 2018, the coverage in both homes and schools was reported to be at a comparable level. In the same way, there is no reason to expect huge differences between students' use of computers either at school or at home. Studies have shown that students have similar leisure uses of computers and use computers relatively little for school work (Bundsgaard & Gerick, 2017; Frønes & Narvhus, 2011; Mejding, 2011). Here, the Scandinavian countries can be characterised by very little variance: several studies confirm that 'everyone' has access to the Internet, that 'everyone' performs the same activities and that the background variables have little power to describe the differences between students (Egeberg, Hultin, & Berge, 2016; Rohatgi & Throndsen, 2015). However, the relation between reading online texts and reading activities may be more complex than indicated here because of imprecise measuring instruments. Also, access is not a reliable predictor of teachers' actual implementation of digital technology (Gil-Flores, Rodríguez-Santero, & Torres-Gordillo, 2017).

Previously, we have substantiated that reading online dynamic texts and/or multiple texts is more demanding for readers. How is it, then, that more students perform at a higher level when encountering these texts? We covered the reasons that might explain why many students are experienced in these new text formats: most adolescents in the Scandinavian countries live digital lives. The new text formats might also give opportunities to learn for a broader group of students. Many students report a higher motivation for reading online texts (OECD, 2019b) and boys have been shown to have an advantage over girls in specific aspects of the comprehension of online texts and hypertexts (Rasmusson & Åberg-Bengtsson, 2015). The reason for this may be that boys have developed their visuo-spatial abilities more than girls, a benefit from playing computer games.

On the other hand, this might be a too optimistic position when considering the educational context for online reading. Several studies have shown that online reading comprehension and strategies are seldom taught at school, even though it is a part of the curriculum in language arts and other subjects (Blikstad-Balas & Klette, 2020). Both Norway and Denmark were among the first in the world to integrate digital skills in the national curriculums but did not emphasise online reading when doing so. There is also a larger between-school variation in online reading performance than in traditional reading tests, which might be explained by decentralised and personalised teaching practices and a discrepancy between access and teachers' preparedness to use the technology in teaching (Carlsten, Caspersen, Vibe, & Aamodt, 2014; Gudmundsdottir & Ottestad, 2016; Throndsen, Carlsten, & Björnsson, 2019). It seems as if the development of the students' online reading skills is largely left to their own literacy practices.

Researchers have agreed that there is a need for specialised strategies when reading online (Afflerbach & Cho, 2009; Coiro, 2011), and these strategy areas—text construction, managing working memory and self-regulation—need to be explicitly

taught to students. Reading dynamic or multiple texts online is especially challenging for students with few reading strategies in their repertoire or with fewer effective strategies. According to Cho (2014), expert readers in an online environment conduct several continuous and parallel reading activities when constructing reading paths, comprehending multiple texts and evaluating and judging the relevance, trustworthiness and usefulness of texts. This description mimics the reading challenges in PISA 2018, which tended to be so hard that many students omitted them. For these students, more explicit instruction in reading digital and online texts is required to ensure that the equity potential shown to be linked to digital text formats is realised.

12.6   Closing Remarks

We emphasise that the findings that give cause for concern are the trends among minority readers from underprivileged homes and the large gender differences. Although Scandinavian reading performance is high, there are many signs that reading education is not as equitable as it should be. In all countries, school policies state that the educational system needs to prioritise the compensatory aim with schooling. However, our analyses confirm Lundahl's claim (2016) that it is highly doubtful if one can still speak of a Nordic model of education, both as an idea of equity and fairness and in the lack of unity across countries because of the development of low-SES students and students with a minority background in Sweden.

However, there is reason to believe that new initiatives and reforms may come. In Norway, the dropout rate for boys in many areas has been investigated by the Stoltenberg Committee (NOU, 2019, p. 3), which has led to discussions on boys' underprivileged position, especially those from low-SES homes or minority backgrounds. In Sweden, a public inquiry has proposed a number of measures to revise the free schooling development to ensure more equal schools and reduced school segregation (SOU, 2020, p. 28). In Denmark, Sweden and Norway, we see an increasing awareness in the academic and policy level of the need for informed didactics for reading instruction in new text formats. A necessary alignment of curriculum, teacher training and teaching practices might open up new equitable opportunities for learning and, hopefully, remove a gatekeeper for participation in our text-based, digitised society.

Appendix

Annex A

Majority and minority students' reading performance in 2000, 2009, and 2012 in Denmark, Norway and Sweden.

		Majority		Minority		
		Average	S.E.	Average	S.E.	Ma-Mi
2018	Denmark	509	1.9	444	3.5	65
	Norway	509	2.1	457	5.7	52
	Sweden	525	2.7	443	5.8	82
2009	Denmark	502	2.2	438	3.8	64
	Norway	508	2.6	456	5.9	52
	Sweden	507	2.7	442	6.9	65
2000	Denmark	504	2.2	432	7.0	72
	Norway	510	2.8	460	6.3	50
	Sweden	524	2.1	467	5.1	56

Annex B

Percent minority students among the bottom SES-students in Denmark, Norway and Sweden.

	2000	2009	2018
Denmark	10.95	18.23	23.95
Norway	7.52	13.52	23.13
Sweden	15.29	21.48	36.79

Annex C

Average equated proportion omitted static/dynamic and multiple/single text items in PISA 2018, in Denmark (DNK), Norway (NOR) and Sweden (SWE).

	Denmark		Norway		Sweden	
	Single	Multiple	Single	Multiple	Single	Multiple
Static						
n items	57	29	57	29	57	29
Mean	0.96	1.33	2.07	2.64	2.02	2.2
Std. dev.	1.07	2.05	1.79	2.78	2.15	2.26
Dynamic						
n items	4	66	4	66	4	66
Mean	3.19	1.99	3.24	3.48	2.78	3.46
Std. dev.	4.84	3.2	3.73	5.12	2.86	4.97

Annex D

Majority and minority students' reading performance in 2000, 2009, and 2012.

		Majority		Minority		
		Average	S.E.	Average	S.E.	Ma-Mi
2018	Denmark	509	1.9	444	3.5	65
	Norway	509	2.1	457	5.7	52
	Sweden	525	2.7	443	5.8	82
2009	Denmark	502	2.2	438	3.8	64
	Norway	508	2.6	456	5.9	52
	Sweden	507	2.7	442	6.9	65
2000	Denmark	504	2.2	432	7.0	72
	Norway	510	2.8	460	6.3	50
	Sweden	524	2.1	467	5.1	56

References

Aasen, P., Møller, J., Rye, E., Ottesen, E., Prøitz, T. S., & Hertzberg, F. (2012). *Kunnskapsløftet som styringsreform—et løft eller et løfte? Forvaltningsnivåenes og institusjonenes rolle i implementeringen av reformen* [The knowledge leap as a management reform – A leap or a promise? The role of the institutions and the different levels of management regarding the implementation of the reform] (NIFU-rapport 20/2012). NIFU & ILS, UiO.

Afflerbach, P., & Cho, B.-Y. (2009). Identifying and describing constructively responsive compre-hension strategies in new and traditional forms of reading. In S. E. Israel & G. G. Duffy (Eds.), *Handbook of research on reading comprehension* (pp. 69–90). New York: Routledge.

Andersen, A. M., Egelund, N., Jensen, T. P., Krone, M., Lindenskov, L., & Mejding, J. (2001). *Forventninger og færdigheder—danske unge i en international sammenligning* [Expectations and skills – Danish youngsters in an international comparison]. Socialforskningsinstituttet.

Antunes, F. (2012). 'Tuning' education for the market in Europe? Qualifications, competences and learning outcomes: Reform and action on the shop floor. *European Educational Research Journal, 11*(3), 446–470. https://doi.org/10.2304/eerj.2012.11.3.446

Berge, K. L. (2005). Skriving som grunnleggende ferdighet og som nasjonal prøve—ideologi og strategier [Writing as a basic skill and as national test—Ideology and strategies]. In A. J. Aasen & S. Nome (Eds.), *Det nye norskfaget* (pp. 161–188). Bergen, Norway: Fagbokforlaget.

Beuchert, L., & Christensen, V. (2019). Læsning og indvandrerbaggrund [Reading literacy and migrant background]. In V. T. Christensen (Ed.), *PISA 2018 – Danske unge i en international sammenligning* (pp. 178–207). Copenhagen: VIVE.

Blikstad-Balas, M., & Klette, K. (2020). Still a long way to go. Narrow and transmissive use of technology in the classroom. *Nordic Journal of Digital Literacy, 15*(1), 55–68. https://doi.org/10.18261/issn.1891-943x-2020-01-05

Blossing, U., Imsen, G., & Moos, L. (2014). Nordic schools in a time of change. In U. Blossing, G. Imsen, & L. Moos (Eds.), *The Nordic education model. A "school for all" encounters neo-liberal policy* (pp. 1–14). Dordrecht, The Netherlands: Springer.

Bråten, I., Britt, M. A., Strømsø, H. I. & Rouet, J.-F. (2011). The role of epistemic beliefs in the comprehension of multiple expository texts: Toward an integrated model. *Educational Psychologist, 46*(1), 48–70. https://doi.org/10.1080/00461520.2011.538647

Braun, H., & von Davier, M. (2017). The use of test scores from large-scale assessment surveys: Psychometric and statistical considerations. *Large-scale Assessments in Education, 5*, 17. https://doi.org/10.1186/s40536-017-0050-x

Bremholm, J., & Bundsgaard, J. (2019). Læsning i PISA 2018 [Reading literacy in PISA 2018]. In V. T. Christensen (Ed.), *PISA 2018 – Danske unge i en international sammenligning* (pp. 18–69). Copenhagen: VIVE.

Bundsgaard, J., & Gerick, J. (2017). Patterns of students' computer use and relations to their computer and information literacy: Results of a latent class analysis and implications for teaching and learning. *Large-scale Assessment in Education, 5*, 16. https://doi.org/10.1186/s40536-017-0052-8

Carlbaum, S., Andersson, E., & Hanberger, A. (2016). *Utvärdering av Läslyftet. Delrapport 2: Erfarenheter av Läslyftet läsåret 2015/16* [Evaluation of Läslyftet. Report 2: Experiences with Läslyftet in the year 2015/16]. Umeå Centre for Evaluation Research.

Carlsten, T. C., Caspersen, J., Vibe, N., & Aamodt, P. O. (2014). *Resultater fra TALIS 2013. Norske funn fra ungdomstrinnet i internasjonalt lys, Arbeidsnotat 10/2014* [Results from TALIS 2013. Norwegian results from 8th to 10th grade in international comparison]. NIFU.

Cho, B.-Y. (2014). Competent adolescent readers' use of internet reading strategies: A think-aloud study. *Cognition and Instruction, 32*(3), 253–289. https://doi.org/10.1080/07370008.2014.918133

Coiro, J. (2003). Exploring literacy on the Internet. Reading comprehension on the Internet: Expanding our understanding of reading comprehension to encompass new literacies. *The Reading Teacher, 56*(5), 458–464.

Coiro, J. (2011). Predicting reading comprehension on the Internet. *Journal of Literacy Research, 43*(4), 352–392. https://doi.org/10.1177/1086296X11421979

Egeberg, G., Hultin, H., & Berge, O. (2016). *Monitor skole 2016. Skolens digitale tilstand* [The monitor school 2016. The digital condition of the school]. Senter for IKT i utdanningen.

Egelund, N. (Ed.) (2010). *PISA 2009 – Danske unge i en international sammenligning* [PISA 2009 – Danish youngsters in an international comparison]. Danmarks Pædagogiske Universitetsskole.

Espinoza, O. (2007). Solving the equity—Equality conceptual dilemma: A new model for analysis of the educational process. *Educational Research, 49*(4), 343–363.

ETS. (2019). *Computation of equated proportion correct for PISA MSAT and evaluation of the 2015 and 2018 MS reading results* (Unpublished manuscript).

EVA, Danmarks Evalueringsinstitut. (2009). *Særlige ressourcepersoner i folkeskolen* [Specialized personnel in the public school]. Rapport. Danmarks Evalueringsinstitut.

Fraillon, J., Ainley, J., Schulz, W., Friedman, T., & Gebhardt, E. (2014). The influence of students' personal and home background on computer and information literacy. In J. Fraillon, J. Ainley, W. Schulz, T. Friedman, & E. Gebhardt (Eds.), *Preparing for life in a digital age* (pp. 101–124). Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-319-14222-7_5

Frønes, T. S., & Narvhus, E. K. (2011). *Elever på nett. Digital lesing i PISA 2009* [Students on the web. Digital reading in PISA 2009]. Institutt for lærerutdanning og skoleforskning, Universitetet i Oslo.

Gil-Flores, J., Rodríguez-Santero, J., & Torres-Gordillo, J.-J. (2017). Factors that explain the use of ICT in secondary-education classrooms: The role of teacher characteristics and school infrastructure. *Computers in Human Behavior, 68*, 441–449. https://doi.org/10.1016/j.chb.2016.11.057

Gudmundsdottir, G. B., & Ottestad, G. (2016). På vei mot profesjonsfaglig digital kompetanse [Towards digital competence in the educational professions]. In R. J. Krumsvik (Ed.), *Digital læring i skole og lærerutdanning* (pp. 70–82). Oslo, Norway: Universitetsforlaget.

Hodgson, J., Rønning, W., Skogvold, A. S., & Tomlinson, P. (2010*). På vei fra læreplan til klasserom: Om læreres fortolkning, planlegging og syn paa LK06. Report.* [From curriculum to classroom: Teachers' interpretations, planning and view of LK06]. Nordland Research Institute.

Imsen, G., Blossing, U., & Moos, L. (2017). Reshaping the Nordic education model in an era of efficiency. Changes in the comprehensive school project in Denmark, Norway, and Sweden since the millennium. *Scandinavian Journal of Educational Research, 61*(5), 568–583. https://doi.org/10.1080/00313831.2016.1172502

Jensen, F., Frønes, T. S., Kjærnsli, M., & Roe, A. (2020). Norske elevers leseforståelse i PISA 2000–2018 [Norwegian students' reading comprehension in PISA 2000–2018]. In T. S. Frønes & F. Jensen (Eds.), *Like muligheter til god leseforståelse? Lesing i PISA i et 20-års perspektiv*. Oslo: Universitetsforlaget.

Kiili, C., Laurinen, L., & Marttunen, M. (2008). Students evaluating internet sources—From versatile evaluators to uncritical readers. *Journal of Educational Computing Research, 39*(1), 75–95.

Klette, K. (2018). Individualism and collectivism in Nordic schools. A comparative approach. In N. Witoszek & A. Midttun (Eds.), *Sustainable modernity: The Nordic model and beyond* (pp. 59–79). London: Routledge.

Kuhlman, J., & Rydén, M.M. (2011). *Er der en læsevejleder på denne skole? En kvalitativ undersøgelse af muligheder og modstande i læsevejlederfunktionen. Report* [Is there a reading counselor at this school? A qualitative study of the possibilities and barriers of having reading counselors at schools]. Nationalt Videncenter for Læsning.

Lawless, K. A. & Kulikowich, J. M. (1996). Understanding hypertext navigation through cluster analysis. *Journal of Educational Computing Research, 14*(4), 385–399. https://doi.org/10.2190/DVAPDE23-3XMV-9MXH

Lie, S., Kjærnsli, M., Roe, A., & Turmo, A. (2001). *Godt rustet for framtida? Norske 15–åringers kompetanse i lesing og realfag i et internasjonalt perspektiv* [Well prepared for the future? Norwegian 15-year-old students' competences in reading and science in an international perspective]. *Acta Didactica*. Institutt for lærerutdanning og skoleutvikling, Universitetet i Oslo.

Lundahl, L. (2016). Equality, inclusion and marketization of Nordic education: Introductory notes. *Research in Comparative and International Education, 11*(1), 3–12.

Martin, A. J., & Marsh, H. W. (2006). Academic resilience and its psychological and educational correlates: A construct validity approach. *Psychology in the Schools, 43*(3), 267–281.

Masten, A. S. (2018). Resilience theory and research on children and families: Past, present, and promise. *Journal of Family Theory & Review, 10*, 12–31.

Mejding, J. (2011). *PISA 2009. Læsning af elektroniske tekster* [PISA 2009. Reading of digital texts]. DPU

Mejding, J. (2019). Test i historisk perspektiv. Hvorfor tester vi? [Tests in a historical perspective. Why do we test?]. *Viden om Literacy, 25*.

Molander, B. O., Pettersson, A., Skarlind, A., & Taube, K. (2001). *PISA 2000: Svenska femtonårings läsförmåga och kunnande i matematik och naturvetenskap i ett internationellt perspektiv* [PISA 2000: Swedish 15-year-old students' reading literacy and skills in mathematics and science in international perspective]. Skolverket.

Moos, L. (2014). Educational governance in Denmark. *Leadership and Policy in Schools, 13*(4), 424–443. https://doi.org/10.1080/15700763.2014.945655

NOU. (2019). *NOU 2019: 3. Nye sjanser—bedre læring.* Kjønnsforskjeller i skoleprestasjoner og utdanningsløp [New opportunities—Better learning. Gender differences in school results and educational trajectories]. https://nettsteder.regjeringen.no/stoltenbergutvalget

OECD. (2001). *Knowledge and skills for life. First results from the OECD Programme for International Student Assessment (PISA) 2000*. Paris: OECD.

OECD. (2009). *PISA data analysis manual, SPSS®* (2nd ed.). Paris: OECD Publications.

OECD. (2011). How are school systems adapting to increasing numbers of immigrant students? *PISA in FOCUS, 11*, 1–4.

OECD. (2015). *Improving schools in Sweden: An OECD perspective*. Paris: OECD Publishing.

OECD. (2019a). *PISA 2018 assessment and analytical framework*. Paris: OECD Publishing. https://doi.org/10.1787/19963777

OECD. (2019b). *PISA 2018 results. What students know and can do* (Vol. I). Paris: OECD Publishing. https://doi.org/10.1787/19963777

OECD. (2019c). *PISA 2018 results. Where all students can succeed*. Paris: OECD Publishing. https://doi.org/10.1787/b5fd1b8f-en

OECD. (2020). *Scaling PISA data* (PISA 2018 Technical Report). OECD Publishing. http://www.oecd.org/pisa/data/pisa2018technicalreport/Ch.09-Scaling-PISA-Data.pdf

Olsen, R. V., Hatlevik, O. E., & Loi, M. (2015). Digitale skiller [Digital divisions]. In O. E. Hatlevik & I. Throndsen (Eds.), *Læring av IKT. Elevenes digitale ferdigheter og bruk av IKT i ICILS 2013* (pp. 146–170). Universitetsforlaget.

Piacentini, M., & Pacileo, B. (2019). *How are PISA results related to adult life outcomes? PISA in Focus, No. 102*. Paris: OECD Publishing. https://doi.org/10.1787/7b60595e-en

Rasmusson, M. (2016). A multilevel analysis of Swedish and Norwegian students' overall and digital reading performance with a focus on equity aspects of education. *Large-scale Assessments in Education, 4*(3), 1–25. https://doi.org/10.1186/s40536-016-0021-7

Rasmusson, M., & Åberg-Bengtsson, L. (2015). Does performance in digital reading relate to computer game playing? A study of factor structure and gender patterns in 15-year-olds' reading literacy performance. *Scandinavian Journal of Educational Research, 59*(6), 691–709. https://doi.org/10.1080/00313831.2014.965795

Roe, A. (2012). Ungdom og lesing—hva forteller nasjonale og internasjonale leseprøver? [Youth and reading—What do national and international reading assessments tell?]. In A. M. Bjorvand & E. S. Tønnessen (Eds.), *Den andre leseopplæringa. Utvikling av lesekompetanse hos barn og unge* (pp. 103–131). Oslo, Norway: Universitetsforlaget.

Rohatgi, A., & Throndsen, I. (2015). Elevenes IKT-bruk [The students use of ICT]. In O. E. Hatlevik & I. Throndsen (red.), *Læring av IKT: Elevenes digitale ferdigheter og bruk av IKT i ICILS 2013* (pp. 93–110). Universitetsforlaget.

Rutkowski, L., Gonzalez, E., von Davier, M., & Zhou, Y. (2014). Assessment design for international large-scale assessments. In L. Rutkowski, M. von Davier, & D. Rutkowski (Eds.), *Handbook of international large-scale assessment. Background, technical issues and methods of data analysis* (pp. 75–96). Boca Raton, FL: RC Press.

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., & Rutkowski, D. (2017). Improving the comparability and local usefulness of international assessments: A look back and a way forward. *Scandinavian Journal of Educational Research, 62*(3), 354–367. https://doi.org/10.1080/00313831.2016.1261044.[13]

Salmerón, L., Strømsø, H. I., Kammerer, Y., Stadtler, M., & van den Broek, P. (2018). Comprehension processes in digital reading. In M. Barzillai, J. Thompson, S. Schroeder, & P. van den Broek (Eds.), *Learning to read in a digital world* (pp. 91–120). Amsterdam: John Benjamins Publishing Company.

SFS 2010:800. (2010). *Skollag*. [The Swedish School Law]. Stockholm: Utbildningsdepartementet.

Shapiro, A., & Niederhauser, D. (2004). Learning from hypertext: Research issues and findings. In D. H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 605–620). New York: Lawrence Erlbaum Associates.

Sivesind, K., Akker, J. v. d., & Rosenmund, M. (2012). The European curriculum: Restructuring and renewal. *European Educational Research Journal, 11*, 320–327. https://doi.org/10.2304/eerj.2012.11.3.320

Skolverket. (2020). *Utvärdering av Läslyftet—Slutrapport* [Evaluation of Läslyftet – Final report].

SOU. (2020). *En mer likvärdig skola—minskad skolsegregation och förbättrad resurstilldelning* [A more equal school – reduced segregation of schools and improved allocation of resources]. Regeringskansliet.

Spiro, R. J., Feltovich, P. J., Jacobson, M. J., & Coulson, R. L. (1992). Cognitive flexibility, constructivism and hypertext: Random access instruction for advanced knowledge acquisition in ill-structured domains. In T. Duffy & D. Jonassen (Eds.), *Constructivism and the technology of instruction* (pp. 57–76). Hillsdale, NJ: Erlbaum.

Spiro, R. J., Klautke, H., & Johnson, A. K. (2015). All bets are off: How certain kinds of reading to learn on the web are totally different from what we learned from research on traditional text comprehension and learning from text. In R. J. Spiro, M. DeSchryver, M. S. Hagerman, P. M. Morsink, & P. Thompson (Eds.), *Reading at a crossroads? Disjunctures and continuities in current conceptions and practices* (pp. 45–50). New York: Routledge.

Statistics Denmark/DS. (2020). *Flytninger til og fra utlandet* [Relocations to and from abroad]. https://www.dst.dk/da/Statistik/emner/befolkning-og-valg/flytninger/flytninger-til-og-fra-udlandet

Statistics Norway/SSB. (2020). *Innvandring og innvandrere* [Immigration and immigrants]. https://www.ssb.no/innvandring-og-innvandrere

Støle, H., Mangen, A., Frønes, T. S., & Thomson, J. (2018). Digitisation of reading assessment. In M. Barzillai, J. Thompson, S. Schroeder, & P. van den Broek (Eds.), *Learning to read in a digital world* (pp. 205–224). Amsterdam: John Benjamins Publishing Company.

Swedish Migration Agency. (2020). *Granted residence permits overviews*. https://www.migrationsverket.se/English/About-the-Migration-Agency/Statistics/Granted-permits-overviews.html

Telhaug, A. O., Aasen, P., & Mediås, O. A. (2004). From collectivism to individualism? Education as nation building in a Scandinavian perspective. *Scandinavian Journal of Educational Research, 48*, 141–158. https://doi.org/10.1080/0031383042000198558

Telhaug, A. O., Mediås, O. A., & Aasen, P. (2006). The Nordic model in education: Education as part of the political system in the last 50 years. *Scandinavian Journal of Educational Research, 50*, 245–283. https://doi.org/10.1080/00313830600743274

Throndsen, I., Carlsten, T. C., & Björnsson, J. K. (2019). *TALIS 2018—Første hovedfunn fra ungdomstrinnet* [TALIS 2018—The first major findings from 8th to 10th grade]. ILS/NIFU.

UFD. (2003). *Gi rom for lesing! Strategi for stimulering av leselyst og leseferdighet 2003–2007* [Make room for reading! Strategy for the stimulation of reading competency and the joy of reading]. Utdannings- og forskningsdepartementet.

UNESCO. (2004). *The plurality of literacy and its implications for policies and programmes*. Paris: United Nations Educational, Scientific and Cultural Organization.

UVM [Undervisningsministeriet/Ministry of Education]. (2001). *Klare mål – Dansk* [National Curriculum – Danish as L1]. Undervisningsministeriet.

UVM [Undervisningsministeriet/Ministry of Education]. (2009). *Fælles Mål – Dansk* [National Curriculum – Danish as L1]. Undervisningsministeriet.

UVM [Undervisningsministeriet/Ministry of Education]. (2014). *Fælles Mål – Dansk* [National Curriculum – Danish as L1]. Undervisningsministeriet.

Wylie, J., Thomson, J., Leppänen, P. H. T., Ackerman, R., Kanniainen, L., & Prieler, T. (2018). Cognitive processes and digital reading. In M. Barzillai, J. Thompson, S. Schroeder, & P. van den Broek (Eds.), *Learning to read in a digital world* (pp. 57–90). Amsterdam: John Benjamins Publishing Company.

Chapter 13
Implications of Changing the Delivery Mode on Reading Tests in Norway—A Gender Perspective

Ragnhild Engdal Jensen

Abstract What can be seen as a digital shift in society is also visible in the Norwegian educational system, as the use of digital devices has increased in both teaching and learning activities. Together with some practical and logistical reasons, the former has very much facilitated the change of delivery mode of the Norwegian National Assessment of Reading Literacy. At the same time, a concern arose regarding whether the test will continue to measure the same underlying concept of reading as before. Furthermore, from the equity perspective, it is important that the change of mode is not disfavourable to any particular group of students. As a solution to this, the format of the test is preserved using fixed, as opposed to dynamic, texts, assuming that fixed texts are consumed in the same way regardless of whether they are presented on paper or on screen. Building on this, this chapter reports on a field trial study for the 2016 Norwegian National Assessment in reading. Nine hundred seventy-three eighth graders from nine different schools participated in completing reading tests on either paper or screen. The main aim of the study is to explore to what extent delivery mode seems to influence students' outcomes. In particular, we investigate whether the change in delivery mode affects boys' and girls' results on reading comprehension tests in the same way. For the purpose of analysis, the Rasch model will be used as a measure of student ability and a multiple regression model will be used to investigate gender differences across the modes. Based on the research so far, we assume that the change in mode will not have a significant impact on student performance relative to gender. The results will be discussed in the light of the gender gap in reading achievement present in the Norwegian educational system.

Keywords Reading comprehension · National tests in reading · Norway · Delivery mode · Gender

R. Engdal Jensen (✉)
Department of Teacher Education and School Research, University of Oslo, Oslo, Norway
e-mail: r.e.jensen@ils.uio.no

13.1   Introduction

Over the past decades, digital technologies have changed how we read and manage information. This phenomenon is evident in many aspects of our lives, and in the field of education, digital technologies are transforming teaching and learning, as well as the ways in which schools assess students. The background for this study relates to recent trends in paper-based reading assessments and their replacement with on-screen assessments. In 2015, the Programme for International Student Assessment (PISA) was delivered on computers for the first time, and in 2016, the Norwegian National Assessment of Reading Literacy Skills was too. The change reflects how students and societies now commonly access, use and communicate information (OECD, 2019), and it is advantageous considering the logistical aspects and security issues of administering the assessments. At the same time, there is a concern regarding whether the tests continue to measure the same underlying concept of reading as before. The use of digital devices as reading tools calls into question how these potentially alter perceptions of what it means to read and the comprehension that results from the activity itself (Singer & Alexander, 2017b).

This chapter explores to what extent delivery mode affects students' outcomes in reading, using the 2016 field study data of the Norwegian national test in reading. While the idea of 'school for all' dominates the school system in Norway, the change in the delivery mode may have significant implications for educational justice. From an equity perspective, it is important that the change of mode is not disadvantageous to any particular group of students. As the trend of girls outperforming boys on reading assessments is well known (Jensen et al., 2019; Solheim & Gourvennec, 2017), we investigate if the change in delivery mode affects boys' and girls' results differently and whether this change has implications for boys and girls having equal opportunities in the test situation.

13.2   Theoretical Background

13.2.1   Mode Effect

Dillon's (1992) review of the literature, intended to examine differences that might exist between reading from a print compared to an electronic source, is referred to as a starting point by several researchers (Delgado, Vargas, Ackerman, & Salmerón, 2018; Singer & Alexander, 2017b). In recent years, a large body of research has emerged, and several updated reviews have been published (Clinton, 2019; Delgado et al., 2018; Kong, Seo, & Zhai, 2018; Singer & Alexander, 2017b). The reviews vary in content and scope; still, all reviews find that, overall, readers demonstrate better comprehension when reading on paper compared to when they read on screen or digitally.

Singer and Alexander's (2017b) narrative review includes 36 studies from the period 2001–2017. Their examination of the literature showed that studies were diverse both in how they define reading in the different media, as well as in how text comprehension is measured. One important finding was that there seems to be an association between the length of the text and the medium, and that readers demonstrate significantly better comprehension when reading on paper if the texts are longer than 500 words or one page. If the texts are shorter, there is no significant difference in the reading comprehension of texts presented in different media. This was evidenced by over 90% of the charted studies in which text length was specified. Further, they emphasize that print seems to be the favourable processing medium when individuals are reading for depth of understanding and not solely for gist.

The review of Delgado et al. (2018) includes 54 studies conducted between 2000 and 2017 that compare reading comprehension when reading printed and digital texts. Thirty-eight of these studies had a between-participants design – participants read either on paper or on the screen – whereas 16 were within-participants studies in which participants read texts in both modes. The results of their meta-analysis showed an advantage for printed texts regardless of the design, with effect sizes being significant (Hedge's $g = -.21$, $d_c = -.21$). The reviews of Kong et al. (2018) and Clinton (2019) include a smaller number of studies – 17 and 33, respectively. Still, the meta-analyses show similar effect sizes to those found by Delgado et al. (2018) (Hedge's $g = -.21$ and $-.25$, respectively). Reading texts from a screen had a small but significant adverse effect on comprehension scores compared to reading from paper.

The tree reviews also included analysis on possible moderators. Clinton (2019) found that readers had a significantly better-calibrated judgement of their performance when reading from paper compared to digitally. Both Delgado et al. (2018) and Clinton (2019) found that the advantage of reading in print increased when participants read expository texts as opposed to narrative texts. Further, Delgado et al. (2018) found that the advantage of reading in print was significantly higher in studies with time constraints compared to studies where participants were allowed to self-pace their reading. This finding was not confirmed by Kong et al. (2018) and Clinton (2019). Although the moderating effect of scrolling did not reach significance, Delgado et al. (2018) emphasize this variable, as their analysis showed a substantial advantage for paper-based reading when scrolling was necessary to read texts on the screen. Finally, it is worth mentioning that text length was not found to be a significant moderator by Delgado et al. (2018), as suggested by Singer and Alexander (2017b).

13.2.2 Mode and Text Processing

The fore-mentioned reviews comprise much of the literature on the field of reading comprehension in different modes. In this section, single studies of particular relevance to the focus of this chapter will be highlighted and discussed in more detail. In a study by Wästlund, Reinikka, Norlander, and Archer (2005), two experiments were performed to investigate the influence of video display terminals (VDT) and paper presentation of text on reading comprehension and the production of information. The results from the study showed that participants reading from computers reported higher levels of experienced stress and tiredness compared to those reading on paper. Furthermore, they found that in both experiments, performance in the VDT presentation condition was inferior to that of the paper presentation condition. Hence, they concluded that the dual-task effects of fulfilling the assignment and working with the computer resulted in a higher cognitive workload.

The supposition that working with a computer results in a higher cognitive workload compared to working on paper is supported by several studies. In a comparison of an identical comprehension task presented on paper and on a computer, Mayes, Sims, and Koonce (2001) showed a significant negative relationship between workload and comprehension scores. The comprehension task was to read a text and answer ten multiple-choice questions, and the workload was measured by the Task Load Index (NASA-TLX). The result showed that increased workload was associated with lower scores. This finding was replicated with thirty undergraduate students by Noyes, Garland, and Robbins (2004). The students read an article, presented in a closely matched form either on paper or on a computer, and then answered ten multiple-choice questions to measure comprehension. Finally, the NASA-TLX was administered. The results showed that there was a significant difference in the perceived effort needed for the computer-based test, and further, that those with lower comprehension scores experienced a higher workload. These findings indicate that lower-performing individuals might be disadvantaged when completing computer-based assessments as compared to similar tasks on paper.

In their research, Noyes and Garland (2003) have also paid attention to the potential impact of presentation mode on cognitive processing and, in turn, learning performance. In a study that examined directly comparable text presented on screen and paper, they included a measure of memory awareness in addition to looking at reading time and comprehension. Such memory awareness measures have been widely used in psychology as a means of gauging recall and, hence, learning. They are based on the work of Tulving (1985), who developed the Remember-Know paradigm. The paradigm describes two types of retrieval response, 'Remember' and 'Know'. 'Remembered' knowledge is typically being recalled in association with related information about the learning episode, whereas 'known' knowledge is recalled without being tied to contextual details or associations. Tulving argued that with time, memory of specific events fades or reduces in contextual details. This implies that 'remembered' knowledge gets less accessible with time. Findings by

Conway, Gardiner, Perfect, Anderson, and Cohen (1997) suggest that knowledge that is 'known' is more readily applied and, as such, indicative of better learning.

The results from the study of Noyes and Garland (2003) indicate that when the material is matched adequately across media, reading time and number of correct answers do not differ. However, a significant effect of awareness frequencies was found in the study. The rate of 'remember' responses was approximately twice that of 'know' responses when reading on screen. In contrast, levels of 'remember' and 'know' responses were similar when reading on paper. The results indicate that cognitive processing associated with memory assimilation differs across mode conditions. Noyes and Garland (2003) suggest that characteristics of the computer screen, such as refresh rate and fluctuating luminance, might interfere with cognitive processing for long-term memory. These findings were confirmed in a later study (Garland & Noyes, 2004) that showed that the manner in which the knowledge was retrieved varied between presentation formats – screen and paper. The study was longitudinal, and the results suggest that repeated exposure to and rehearsal of computer-based information is needed to equate knowledge retrieval with that achievable from paper. The knowledge transition when reading from the screen was much more rapid compared to paper, which indicates that knowledge seems to be better adapted and, in turn, more easily applied when presented in paper format. Garland and Noyes conclude that "there still appears to be a benefit attached to learning from paper-based rather than computer-based material" (2004, p. 51).

Another possible explanation of the apparent comprehension differences across modes might be related to metacognitive skills (Ackerman & Goldsmith, 2011; Ackerman & Lauterman, 2012; Lauterman & Ackerman, 2014). When comparing the reading performance of undergraduate students who read identical texts on screen and paper, Ackerman and Goldsmith (2011) found that, under a fixed study time, test performance did not differ between the two media. However, when the study time was self-regulated, students performed poorer on screen than on paper. Further, the results showed that the students were less able to give an accurate prediction of their performance, tending to overestimate their comprehension when reading on screen. This was also accompanied by poorer allocation of study time. Hence, Ackerman and Goldsmith (2011) conclude that the primary difference between the two media is not cognitive but rather metacognitive. The authors conclude that metacognitive processes might be less effective on screen due to higher-order metacognitive beliefs. Previous research shows that people seem to perceive printed paper as the medium best suited for effortful learning, whereas the electronic medium is suited to fast and shallow reading of short texts, such as news and e-mails (Shaikh, 2004; Spencer, 2006). Such a perception might reduce the mobilization of cognitive resources that are needed for effective self-regulation (Ackerman & Goldsmith, 2011).

Also, research shows that people's use of digital media makes them less likely to engage in reflective thought (Annisette & Lafreniere, 2017). This is consistent with what has come to be known as the 'shallowing hypothesis' (Carr, 2010). The hypothesis proposes that the frequent use of ultra-brief social media, such as texts and tweets, characterized by quick, social interactions, promotes rapid, shallow and

non-reflective thought. Further, people typically process digital texts in a shallow or superficial way, and such digital activities might, in turn, prevent success when performing more complex activities that require sustained attention – for instance, processing longer texts.

The assumptions associated with the shallowing hypothesis are in line with findings showing that readers spend less time processing digital texts compared to paper-based texts. A study by Singer Trakhman, Alexander, and Berkowitz (2017) explored the effects of print and digital texts on readers' comprehension and processing time. They predicted that there would be differences in the time spent reading digital compared to printed texts, and further that processing time would serve as a mediator between medium and comprehension performance. This is in keeping with the speed-accuracy trade-off hypothesis (Wickelgren, 1977), which suggests a trade-off between the speed at which a certain task is performed and the quality of the product. The results showed that participants read significantly faster when texts were displayed on a computer than when texts were on paper, and that there was a significant direct effect of the medium on overall comprehension. Further, medium predicted processing time, which in turn predicted comprehension scores. Processing time significantly mediated the effects of the medium on readers' comprehension (Singer Trakhman et al., 2017).

Another topic that has been of interest to Singer Trakhman and colleagues (Singer & Alexander, 2017a; Singer Trakhman et al., 2017) is students' calibration when they read in print or digitally. Calibration can be defined as the distance between perceived performance and demonstrated levels of understanding or competence (Alexander, 2013). Singer and Alexander (2017a) examined whether students' judgments of their reading comprehension abilities under print and digital conditions would match their actual comprehension performance. The results showed that when asked to judge the medium in which they performed best, the majority of the participants indicated the digital medium. However, more students demonstrated stronger comprehension when they were reading on paper. This indicates that participants were generally poorly calibrated. The number of participants that presumed they would be better at performing in the digital medium but, in reality, comprehended better on paper, was significant. This is in line with the research of Ackerman and Goldsmith (2011) and was also confirmed by Singer Trakhman et al. (2017), who found that the participants' calibration was significantly worse when reading on screen compared to paper. They suggest that this may be explained by the potential influence of processing speed, and that calibration, as well as comprehension, might be subject to the speed-accuracy trade-off. They also suggest that there might be an association between the level of effort exerted in a task and the judgement of comprehension, referring to research by Koriat, Ma'ayan, and Nussinson (2006) that showed that the less effort exerted in task performance, the higher the judgement of learning.

Another recent study that confirms poorer calibration when reading on screen was published by Halamish and Elbaz (2019). This study makes an important contribution as it discusses the mode effect on children's comprehension and meta-comprehension judgements. In their meta-analysis, Delgado et al. (2018) did not

find age to be a moderator for the effect of medium on reading comprehension. Halamish and Elbaz (2019) suggest that this finding should be considered with caution as the number of studies on children included in the meta-analysis was small. However, it implies that children also tend to comprehend texts better on paper than on screen.

The study by Halamish and Elbaz (2019) gathered 38 fifth-grade children who read short texts on paper and screen. The students estimated their comprehension of each text and answered a reading comprehension test. The results showed that the children's comprehension was better when reading on paper compared to on the screen. Nevertheless, most children judged their comprehension to be the same on paper and screen, which suggests that they were metacognitively unaware of the effect of medium on their comprehension. Another study with 82 children of 11–12 years of age (Dahan Golan, Barzillai, & Katzir, 2018) also found that performance was better when reading on paper and that the children were more confident and better calibrated than when reading on screen. However, the majority of the children stated that they preferred to read on screens. This preference underpins the suggestion that children are unaware of the effect of medium on their comprehension. A recent study by Støle, Mangen, and Schwippert (2020) also found paper to be advantageous for children's reading comprehension. In this study, 1139 fifth-grade students participated, taking two comparable versions of a reading comprehension test, one on paper, and one digitally. Their results further showed that the negative effect of screen reading was evident for both boys and girls, but most profound among high-performing girls.

The same tendency is visible when looking at studies that concern adolescents. In 2017, Eyre et al. published a report on the digitization of the PAT: Reading Comprehension, a low-stakes, standardized assessment developed for use in New Zealand Schools, grades 4–10. Close to 200,000 assessment records were collected, and results showed that comprehension was lower when texts and items were presented on screen compared to when they were presented on paper. Mangen, Walgermo, and Brønnick (2013) also found that students who read texts in print showed significantly better comprehension than students who read on screen when exploring mode effect on 15-year-old's reading of linear texts. In line with this are the findings from a study by Rasmusson (2015), who investigated differences in performance when 14-year-olds did the same reading test on paper and screen. The results showed a difference in favour of reading in print.

13.2.3   Gender Differences in Reading

The overarching values within the Norwegian education system include social justice, equity, equal opportunities to learn, inclusion and democratic participation for all students, regardless of their social and cultural background and abilities. All these ideas are interwoven within what is known as the Nordic model (Imsen, Blossing, & Moos, 2017). Results from PISA 2018 indicate that the Norwegian

educational system can be seen as equitable with respect to the socioeconomic status (SES) of the students. The influence of SES on achievement is significantly lower in Norway than the average across the OECD countries. Furthermore, only small differences related to students' performance are observed between schools (Jensen et al., 2019). However, the trend of girls outperforming boys on reading assessments is well known, and this trend applies to almost all countries that participate in large-scale assessments, such as PISA and Progress in International Reading Literacy Study (PIRLS) (Roe, 2013). In Norway, the phenomenon has been paid a great deal of attention, as the gender gap is significantly bigger than the OECD average (Jensen et al., 2019). Furthermore, this gap has been stable since the first PISA cycle in 2000, indicating that Norwegian schools fall behind on gender equity in reading.

Looking at the distributions of boys and girls across the levels of reading proficiency gives a more detailed picture of the gender gap. The proportion of Norwegian boys performing below level 2[1] was 26% in PISA 2018. For girls, this was 12%. Correspondingly, more girls were performing at the highest levels compared to boys (Jensen et al., 2019). The results from PIRLS show the same tendency. On the two lowest proficiency levels, the proportion of boys is almost 70%. Correspondingly, the proportion of girls is larger on the high comprehension levels. On the most advanced level, 64% are girls and 36% are boys (Solheim & Gourvennec, 2017).

Girls also outperform boys on the Norwegian National Assessment of Reading Literacy. Roe and Vagle (2012) found that open constructed-responses show a larger difference for boys and girls than multiple-choice items. This is partly due to boys skipping the open constructed-responses and partly due to short or wrong answers. This result was also confirmed when observing the performance of Norwegian students in PISA (Roe & Vagle, 2010). In their review of the national assessments, Roe and Vagle (2012) also found that the gender gap was larger for fictional texts than for factual texts. In particular, if the main character was female, the gender gap was twice as big as when the main character was a boy. This is in line with research showing that boys perform better on texts they like and find interesting than on texts they do not like (Oakhill & Petrides, 2007). Girls' performance, on the other hand, does not seem to be affected by motivational factors to the same extent, as their achievement is largely the same across all text types. Frønes (2016) suggests that this could be related to leisure time reading habits, with girls reading more diverse texts compared to boys.

Concerning reading habits, girls report that they read fictional literature more often than boys, who prefer reading newspapers both on paper and online. Results show that students who often read fictional literature demonstrate better reading comprehension than those who do not (Roe, 2020). Further, boys and girls express different engagement in reading. Girls spend significantly more time reading for pleasure than boys. In Norway, the results from PISA show that boys view reading

---

[1] Level 2 is set by OECD as the baseline where students begin to demonstrate the competencies that will enable them to participate effectively and productively in life as continuing students, workers and citizens.

as a mere necessity and that they, to a larger degree than girls, only read if they have to. The same picture is portrayed across all participating OECD countries. However, Norwegian boys are among the least positive, a tendency that has persisted since the first PISA administration in 2000 (Roe, 2020).

In PISA 2009 and 2018, student's metacognitive reading strategies were also measured. The students were asked to rate the usefulness of different strategies proposed for different reading situations, and their answers were compared to the judgements of expert raters. The results show that Norwegian students score close to average when judging the strategies. However, the score difference between boys and girls was similar to that of the reading comprehension test; that is, favouring girls. The reason for boys' lower scores is that they do not distinguish between good and poorer strategies. Instead, they tend to rate all strategies as fairly good (Hopfenbeck & Roe, 2010; Jensen et al., 2019).

In terms of motivation, it is a common expectation that doing tests in a digital environment will benefit the performance of boys, because computers may motivate them more than paper and pencil tests do (Martin & Binkley, 2009). However, such an assumption should be tied to research on students' digital habits. In PISA 2009, the students reported on their use of computers (Frønes & Narvhus, 2011). The results showed that more than 70% of both boys and girls used computers daily or almost every day for chatting and for surfing on the Internet. For most other activities, such as homework and reading email, gender differences were small as well. Also, in PISA 2018, boys and girls report on their digital habits quite similarly (Roe, 2020). In PISA 2009, close to 50% of Norwegian boys, compared to less than 10% of the girls, reported that they used computers for gaming daily or close to daily. Updated numbers from The Norwegian Media Authority (2020) show that more girls are now interested in gaming, but the gender differences remain large. Ninety-six percent of the boys and 76% of the girls play games. In all age groups, the proportion among boys is larger than among girls, and among the girls, gaming becomes less widespread the older they get.

13.2.4   Digitization of Reading Assessments

Since 2004, The Norwegian National Assessment has been administered annually to students in grades 5 and 8 (10- and 13-year-olds). Students' skills in reading, mathematics and English as a second language are assessed. The tests provide information concerning individual students, student groups and schools, and are used both as an indicator for school improvement at a political level as well as the basis for formative assessments of students learning by teachers. The reading tests are, to a large extent, modelled like the international large-scale assessments PISA and PIRLS and share many similarities in terms of how the reading construct is defined and operationalized. The purpose of the tests is to measure students' reading literacy skills in terms of text comprehension as a basic skill (The Norwegian Directorate for Education and Training, 2017a). Thus, reading literacy is broadly defined as

being able to understand, use, reflect on and engage with texts. The definition is consistent with the definition used for the reading assessment in PISA: "Reading literacy is understanding, using, evaluating, reflecting on and engaging with texts in order to achieve one's goals, to develop one's knowledge and potential and to participate in society" (OECD, 2019, p. 28).

Following the digitization of PISA in 2015, The Norwegian National Assessment was administered on screen for the first time in 2016. The digitization of assessments has some advantages. In many cases, costs can be reduced; data collection is automatized and does not necessarily need to be supervised by researchers. Further, for some item types, scoring can be done by computers, which also eliminates error from manual scoring. Another advantage, pointed out by Støle et al. (2020), is the greater flexibility in text presentation when tests are computer-based, for instance, by using hyperlinks and dynamic elements. This allows for displaying texts that resemble the online texts children and adolescents meet in different types of electronic platforms. However, this also sheds light on some of the challenges with digitizing reading assessments. As new opportunities arise; consideration must be paid to ensure continuity with previous paper-based reading assessments. In addition, it is important to ensure that the change in the test conditions does not hinder students' opportunities to succeed, regardless of the possible constraints related to some underlying factors (Espinoza, 2007).

In many cases, paper-based assessments have simply been replaced by digital assessments because mode equivalence has been assumed (Noyes & Garland, 2008). This could, however, be considered a break with traditional ways of categorizing reading activity and texts, as a distinction has often been made between paper-based and digital texts. The framework for PISA 2009 uses the terminology 'print-medium texts' and 'electronic-medium texts' (OECD, 2009). Print-medium texts have a static existence – the amount of text is immediately visible and the physical status of the text encourages the reader to approach the content in a certain order. Electronic-medium texts, on the other hand, are hypertext featuring navigation tools that make non-sequential reading possible, and often necessary. The reader chooses his or her reading path, and since the text is undefined and dynamic, it can be customized during the reading, often by the reader himself. On screen, only a fraction of the available text can be seen at any one time, and the extent of the text is unknown.

The distinction between print-medium and electronic-medium texts used in the framework for PISA 2009 (OECD, 2009) underlines the importance of medium for the categorization of texts. However, in the PISA 2015 framework, this distinction is no longer made due to the digitization of the test, meaning that the texts that were previously presented on paper were now delivered on screen. Although it is emphasized that the change of mode implies a break with previous assessments, it is argued that "both 'print-medium' and 'electronic-medium' texts can be consumed onscreen" (OECD, 2013, p. 15). Hence a new distinction is made between fixed and dynamic texts, moderating the link between text and medium (OECD, 2017). It is, however, a concern whether fixed texts are processed in the same way regardless of presentation mode. This concern pertains to the assumption that the medium itself

might set the premises for how texts are read. When processing dynamic texts, the use of navigation tools is essential for constructing meaning through the non-sequential reading path. Such navigation tools might be scrollbars, tabs, and various displays of hyperlinks. Although they have not been paid as much attention, navigation tools are also available when reading fixed texts, for instance, tables of contents, chapters, headlines and page numbers. Transferring fixed texts to the screen implies that the navigation tools of the print condition need to be accompanied by tools that are unique to the electronic medium. On this basis, one might question whether delivery mode can be disregarded when categorizing texts. Mangen and Kristiansen (2013) argue that texts read on screen, in essence, are volatile, dynamic and changeable, even if they are not multimodal or hypertext, but linear and in most ways look as if they are printed on paper. Even if the text is the same, the different affordances of the print and electronic media might affect the reading processing in different ways (Mangen, 2010).

In the framework for The Norwegian National Assessment, it is emphasized that the texts included in the test are meant to reflect the diversity of texts that the students typically encounter in the different subjects – not only verbal text but also illustrations, graphic representations, symbols and other possible ways of expression. Knowledge about different types of texts and text functions is therefore considered a crucial part of students' reading literacy skills (The Norwegian Directorate for Education and Training, 2017a). Regarding the digitization of the test in 2016, it is essential to point out that it resembles the way it was carried out for PISA in 2015. Despite the quite broad definition of text in the framework of The Norwegian National Assessment, computers are used for assessing fixed texts and not dynamic texts.

In the framework for The Norwegian National Assessment, it is further stated that the description of the tests might be revised as more results are obtained on how the digitized version of the test is working (The Norwegian Directorate for Education and Training, 2017a). On this notion, it is recognized that mode equivalence cannot easily be assumed. Moreover, it is crucial to obtain knowledge on how or if delivery mode affects students' reading comprehension and whether it affects everybody in the same way. From an equity perspective, one would want to assure that the change of mode is not disadvantageous to any particular group of students. Seen in the context of 'the equality-equity model' of Espinoza (2007), this can be linked to the output stage of the educational process and the importance of securing equity for equal achievement. When implementing new test conditions, it is essential for fairness that students who have achieved the same in the past continue to achieve similarly irrespective of the mode change. More specifically, if mode change is beneficial to some students and not to others, knowledge needs to be obtained so that fairness can be assured. This could, for instance, have implications for teacher practice, requiring change and customization of the reading instruction.

13.2.5 The Present Study

The present study presents results from the double mode assessment of reading comprehension, which was part of the preparations for changing the delivery mode of the Norwegian National Assessment in reading. The study uniquely contributes to the understanding of how delivery mode may affect the reading of Norwegian adolescents. An essential purpose of the study was to establish empirical evidence refuting or supporting the assumption of mode equivalence. Against the background of the research that has been shown so far, we can see that, overall, readers demonstrate better comprehension when reading on paper compared to when they read on screen or digitally (Clinton, 2019; Delgado et al., 2018; Kong et al., 2018; Singer & Alexander, 2017b). The first research question aims at further investigating this, based on the Norwegian context:

1. To what extent does overall comprehension performance differ when students process texts and solve items on paper and screen?

Further, in terms of equity, the change of mode should not be disadvantageous to any particular group of students (Espinoza, 2007). Research shows that girls outperform boys on reading tests (Jensen et al., 2019; Solheim & Gourvennec, 2017). However, little attention has been paid to see how delivery mode may affect gender differences, and the findings so far are inconclusive (Støle et al., 2020). The second research question, therefore, aims at exploring if the gender gap seen on paper-based reading assessments will translate to digitally delivered assessments:

2. Does change in delivery mode affect boys' and girls' results on reading comprehension tests in the same way?

13.3 Methods

13.3.1 Participants, Test Design and Administration

The study was administered in February 2016. Nine hundred seventy-three students from eighth grade (age 13–14) participated (48.7% female). The students came from nine different lower secondary schools. The schools were randomly picked from a list provided by the Norwegian Directorate for Education and Training and were spread geographically across the country. Both urban and rural schools were represented. The number of students from each school was distributed evenly. In conclusion, the number of participating schools makes the sample non-representative. However, the process has resulted in a sample of schools covering a relevant variation of contextual factors in Norway.

The study entailed two reading comprehension tests, Test 1 and Test 2. Each test consisted of seven texts that were similar concerning text length, text types and

formats. Both short and long texts were included, ranging in length from 228 to 1022 words. As the purpose of the assessment was to measure students' reading literacy skills in terms of text comprehension as a basic skill, the tests were designed from a wide selection of texts within different subjects (The Norwegian Directorate for Education and Training, 2017b). Both tests included expository texts, continuous and non-continuous, representing diverse subjects, such as history, natural science, social science and language arts. In each test, there was also one narrative text. To avoid gender effects as a result of text topic, the texts included in the tests were assumed to appeal to both boys and girls. The National Assessment typically contains 40 reading items. In all, 92 items were piloted. Test 1 contained 47 items (35 multiple-choice and 12 short-answer constructed-response items). Test 2 contained 45 items (only multiple-choice items). All multiple-choice items had four alternatives – one correct answer and three distractors. All items were scored dichotomously.

Both tests were administered digitally and on paper, and for each test, the screen version and paper version were made as close to identical as possible. The digital tests were completed on computers, and the students read from standard computer screens, typically 20 inches or a little smaller if using a laptop. Mouse and keyboard were used for navigation, selecting multiple-choice responses, and to type answers to constructed-response items. No training was provided in advance as most students were likely to previously have used the computers in classrooms or computer labs at the schools. Furthermore, no training was provided in using the digital platform, as it was familiar to students from the national assessments of mathematics and English that were digitized in 2014. The paper versions of the tests were formatted in A4 size and were handed out as booklets. Most texts filled about two pages, including tables, illustrations and graphics. Some of the texts were presented double-paged, while in other cases students had to turn a page. The comprehension items were displayed after each text, and students had to turn up to two pages to see the items connected to the text. Due to the length of most texts, students had to scroll when taking the digital version of the tests. After reading or scrolling through a text, items would appear on the left side in the platform window, while the text continued to be visible to the reader. In both conditions, students had the opportunity to move back and forth between texts and items, and they could revise their responses. The time limit of the test was 90 min.

All students conducted the tests at school administered by their teachers. The teachers had been told to give instructions according to the guidelines provided by the researchers. As the national assessments are administered annually for the full cohort of fifth-, eighth- and ninth-grade students, it is reasonable to assume that many of the teachers would be experienced test administrators. However, the digitized version was new for the reading assessment. The study had a between-participant design, and each student was assigned to one of the two tests, taking it on either paper or screen. The students were assigned randomly to the two tests; 470 students completed Test 1, and 503 students completed Test 2. However, for delivery mode, the students were assigned class-wise, as we wanted to avoid a design too sophisticated for the teachers to handle. In support of this decision is the argument that randomness was ensured by randomly assigning students to the two different

tests. Further, we know that Norwegian classrooms tend to be quite heterogeneous considering that students with different home background and abilities are mixed. After completing the tests, data from the digital version was generated automatically, and the paper booklets were returned to the researchers, who scored the responses. For all short-answer constructed responses, both digital and on paper, at least two experienced raters scored each response to secure inter-rater reliability.

13.3.2   Data Analysis

The data collected from the study provided the following information for each student: school, mode, gender and score (item format and frequency on multiple-choice items). Probabilistic test theory was employed to give a measure for student achievement. To be more specific, the Rasch model was applied, which allows for characterization of students' proficiency and difficulty of items as locations on the same continuous scale. The origin of the scale is identified by the mean item difficulty. Students' proficiency corresponds to the point on the scale where they will have a 50% probability of responding correctly to an item. Given that the two tests were unique and non-linked test forms, the scaling was done separately for each of the two test forms. However, the two versions of each test (paper and on-screen) were calibrated together. The software package RUMM2030 was used for the scaling, while statistical analyses were conducted in SPSS 26.

Gender differences across the modes were investigated through multiple regression analysis, using dummy-coded variables for mode (1 = screen) and gender (1 = girl) as predictors. Also, an interaction term for mode and gender was included in the model as the product of the variables for mode and screen. As exemplified by Aiken and West (1991), this allows for the exploration of conditions under which causal relationships are moderated or strengthened. In the case of this study, the interaction term makes it possible to see if the effect of the mode change is the same for boys and girls.

Several assumptions (i.e., homoscedasticity, normality and independence of residuals, multicollinearity, as well as variables tolerance) were checked to ensure that the regression model fit the data (Cohen, Cohen, West, & Aiken, 2003). All assumptions were met. Furthermore, as extreme scores potentially have a great impact on regression models (Osborne & Overbay, 2004), the dataset was screened to detect outliers. As a result, 3 and 9 students were removed from the dataset in Test 1 and 2, respectively.

13.4   Results

The first research question guiding this study focused on the role of the medium in students' overall reading comprehension when processing texts and solving items in print and on screen. Descriptive data on the two comprehension tests overall and by medium are displayed in Table 13.1. The data show that students scored a little lower on both tests when they were administered digitally. However, the differences between the mean scores of the two modes are low. On average, students score 0.05 higher on Test 1 and 0.07 higher on Test 2 when reading on paper compared to screen. This difference by medium on overall comprehension is not significant, and the prediction of students having higher comprehension scores when reading on paper is not confirmed.

As a first insight into the second research question for this study – whether the change in delivery mode affects boys' and girls' results on reading comprehension tests in the same way – descriptive data for reading comprehension on both tests, split on medium and gender, are provided in Table 13.2. On both tests, girls performed better when the test was administered on screen as compared to on paper. For Test 1 the mean score for girls' comprehension was .53 (SD = 1.06) on screen and .42 (SD = .97) on paper, a mean difference of .11 between modes. On Test 2, the mean difference for girls' comprehension was .03, favouring the digital condition. Boys, on the other hand, perform better on paper than screen, the mean difference being .25 on Test 1 and .17 on Test 2. The difference in mean scores between modes is larger for boys than it is for girls.

Table 13.1 Means and standard deviations for reading comprehension by medium on Test 1 and Test 2

	Test 1			Test 2		
	N	*M*	SD	*N*	*M*	SD
Paper	247	.31	1.02	250	.28	.95
Screen	223	.26	1.10	253	.21	.93
Total	470	.28	1.06	503	.25	.94

Table 13.2 Means and standard deviations for reading comprehension by medium and gender

Test	Mode	Gender	*N*	*M*	SD	*p*-Value
Test 1	Paper	Boy	131	.22	1.06	.139
		Girl	110	.42	.97	
	Screen	Boy	109	−.03	1.07	.000
		Girl	114	.53	1.06	
Test 2	Paper	Boy	126	.24	.98	.421
		Girl	115	.34	.94	
	Screen	Boy	133	.07	.91	.010
		Girl	120	.37	.94	

Note: p-Value from t-test for individual samples, comparing scores of boys and girls by mode on the two tests. *p*-value is considered statistically significant at *p* < 0.05

Table 13.3 Regression model with reading comprehension scores on Test 1 as an outcome (Outliers removed, $N = 3$)

Model	Variable	R^2	R^2 change	F change	b	p-Value
1		.032	.032	7,569		.001
	Mode (0 = paper, 1 = screen)				−.091	.330
	Gender (0 = boy, 1 = girl)				**.357**	.000
2		.041	.009	4.375		.037
	(Constant)				**.247**	.005
	Mode (0 = paper, 1 = screen)				**−.281**	.032
	Gender (0 = boy, 1 = girl)				.169	.193
	Interaction mode-gender				**.391**	.037

Note: p-Value is considered statistically significant at $p < 0.05$

Table 13.4 Regression model with reading comprehension scores on Test 2 as an outcome (Outliers removed, $N = 9$)

Model	Variable	R^2	R^2 change	F change	b	p-Value
1		.014	.014	3.359		.036
	Mode (0 = paper, 1 = screen)				−.053	.499
	Gender (0 = boy, 1 = girl)				**.194**	.013
2		.023	.009	4.464		.035
	(Constant)				**.260**	.001
	Mode (0 = paper, 1 = screen)				−.208	.052
	Gender (0 = boy, 1 = girl)				.026	.815
	Interaction mode-gender				**.328**	.035

Note: p-Value is considered statistically significant at $p < 0.05$

Even if the girls performed better than the boys overall, when breaking down the scores by test form, medium and gender, as shown in Table 13.2, it is evident that girls outperformed boys in the on-screen condition, with a difference of .56 and .44 for the two tests, respectively. As shown by the column to the far right in Table 13.2, listing the *p*-values from a t-test for individual samples, comparing scores of boys and girls by mode on the two tests, both differences are statistically significant ($p = $ .000 and $p = $ .010, respectively). However, the gender differences for the paper-based tests are trivial and non-significant. Before turning to the regression analysis, it is worth noting that these results indicate the existence of an interaction effect.

Tables 13.3 and 13.4 show the results of the regression analysis for Test 1 and Test 2, respectively, and overall, the results show the same tendencies for the two tests. The values of R^2 for the steps in both analyses show that very little variance in the criterion variable (the score) is explained by the models. This amounts to about 3–4% for Test 1 and about 1.5–2.5% for Test 2. It is, however, significant at the .05 level (Test 1, $p = $ .001, Test 2, $p = $ .036). Explanatory power was improved in the second model, as can be seen from the positive change in R^2, the change being significant in both cases (Test 1, $p = $ .037, Test 2, $p = $ .035).

In line with the results from the descriptive analysis, the first model of the regression analysis shows no significant difference for mode, neither for Test 1 nor Test 2. However, there is a strong and significant difference for gender (Test 1, b = .357, *p*

= .000, Test 2, b = .194, p = .013), most prominent in Test 1. The fact that R^2 is low indicates that the variance in reading comprehension is much larger among each gender, respectively, compared to the difference between boys and girls. Turning to the second model of the regression analysis, with the interaction term included, this also confirms the descriptive analysis, as the interaction effect is significant on both tests (Test 1, b = .391, p = .037, Test 2, b = .328, p = .035). Further, with the interaction term included, the gender difference is no longer significant on any of the tests. However, mode turns out to have a significant effect on reading comprehension in the second model for Test 1(b = −.281, p = .032), and it is close to significant for Test 2 (b = −.208, p = .052).

As a further illustration of the interaction effect documented by both the descriptive analyses and the regression models, Figs. 13.1 and 13.2 show predicted values for boys and girls across modes for Test 1 and Test 2, respectively. The predicted values are calculated from the regression coefficients by using the equation:

$$\hat{Y} = b_1 X_1 + b_2 X_2 + b_{12} X_1 X_2 + b_0$$

In the equation, X_1 pertains to 'Mode' and X_2 to 'Gender'. Given that we, for example, want to know the predicted scores on Test 1 for boys reading on screen, the calculation will be: (1*(−.281)) + (0*.169) + (0*.391) + (.247) = −.034. The same is done for the other conditions; the results are given in the plots. The fact that the lines for the two groups go in separate directions illustrates very well the interaction effect with a widening of the gender gap from paper to screen.
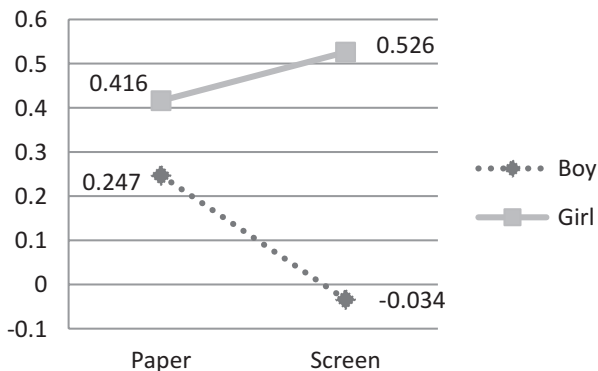

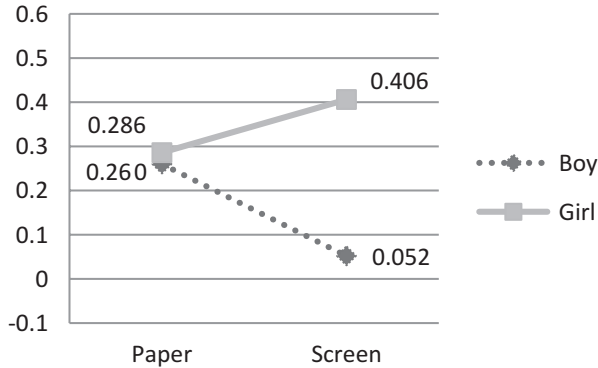**Fig. 13.1** Predicted values for Test 1, model 2

Fig. 13.2 Predicted values for Test 2, model 2

13.5 Discussion

This study was motivated by recent trends in the field of large-scale assessments, as paper-based reading tests are being replaced with digitally delivered, on-screen assessments. The preparation of digitizing the Norwegian National Assessment in reading in 2016 offered a unique opportunity to perform a mode effect study among adolescents. In particular, two questions were addressed: first, to what extent overall comprehension performance differs when students process texts and solve items on paper and screen, and second, if change in delivery mode affects boys' and girls' results on reading comprehension tests in the same way. Investigating these questions is relevant for understanding how delivery mode may affect students' reading and, in turn, how changes in test conditions may have implications for fairness in student assessment. From an equity perspective, it is important that students have the same opportunity to succeed as they have had in the past (Espinoza, 2007).

The results of this study did not reveal significant differences in overall reading performance among 13–14-year-olds as an effect of delivery mode. This is contrary to what could be expected, reviewing the literature in the field (Clinton, 2019; Delgado et al., 2018; Kong et al., 2018; Singer & Alexander, 2017b). At the same time, 13–14-year-olds are often labelled as digital natives (Prensky, 2001), and most of them are likely to possess extensive digital skills and experience. Within the Norwegian educational policy, children's digital skills have been prioritized (The Norwegian Directorate for Education and Training, 2017b), and computers and tablets are widely used as learning tools in primary and secondary schools. Norwegian children and adolescents also have experience with digital devices at home. Ninety-nine percent of 17–18-year-olds have their own mobile phone, and more than 98% have their own computer (The Norwegian Media Authority, 2020). Also, younger students have wide access to digital devices. In PIRLS 2016, Norway ranked highest concerning children's access to digital devices; high access was reported for 58% of the children, whereas low access was not reported (Mullis, Martin, Foy, & Hooper, 2017).

Furthermore, when preparing the test for each mode, efforts were made to secure a low-threshold digital solution. In order to keep the digitized test in line with the previous paper-based version, it was important to make the screen and print version as similar as possible. Equal attention was paid to ensuring that the technical requirements of using computers were not higher than what could be expected for the age group. More specifically, the students had to use the mouse and keyboard for responding to items and for navigation, the navigation tools mainly being scrollbars for displaying longer texts and tabs for moving back and forth between texts. Considering the navigation skills that were anticipated among these students, the requirements were not expected to be too challenging.

Turning to the second question addressed by this study, the reviewed literature did not propose a clear hypothesized answer, as little research has been done on mode effect and gender differences (Clinton, 2019; Delgado et al., 2018; Kong et al., 2018). The results showed a widening of the gender gap, with boys clearly not benefitting from completing the tests on screen. This is a matter of concern for educational justice, as possible constraints related to underlying factors should not hinder students' opportunity to succeed (Espinoza, 2007). Knowing that research indicates that Norwegian schools fall behind on gender equity in reading (Jensen et al., 2019) makes it particularly important to further understand what might affect the gender gap to increase when changing the test conditions. This could be of guidance to policy makers and teachers.

Computers have been assumed to motivate boys, and it is a common expectation that boys will benefit from tests being digitized (Martin & Binkley, 2009). The results of this study show that this assumption might not hold. The gender gap increased in the on-screen version of both tests, boys' comprehension scores being negatively affected by the screen condition. Several factors may have contributed to the results. First, it is uncertain to what degree boys' motivation for using computers is transferrable to completing reading comprehension tests on screen. As can be seen from the report on children and media (The Norwegian Media Authority, 2020), for instance, boys are motivated to use computers for gaming. Whether this would translate into motivation for digitized reading assessments is not clear. However, as children and adolescents today are considered to be digital natives, the use of computers as such has likely been de-mystified.

The use of screens for reading, both in and out of school, steadily increases. The activities children and adolescents most frequently use computers and digital devices for at home are watching video clips and listening to music, visiting social network profiles, socializing and communication, playing games, and searching for information to satisfy curiosity (Mascheroni & Cuman, 2014). Both boys and girls also report that they use computers for activities which, to a greater extent, are related to reading in particular, such as chatting, surfing on the Internet, searching for information and doing homework (Frønes & Narvhus, 2011; Roe, 2020). Most of these texts that are encountered on screen share the features of being dynamic, undefined and interactive. Considering the 'shallowing hypothesis' (Annisette & Lafreniere, 2017; Carr, 2010), digital texts are often processed in a shallow or superficial way, as digital texts may promote a way of reading that typically involves

skimming and scanning. In turn, this could contribute to some children and adolescents developing a screen reading behaviour that is not beneficial for deep reading and processing of longer texts. If the student's screen reading is modelled on strategies efficient for quick and superficial reading, this might explain why the scores of students taking the tests on screen were poorer than the scores of those who took the tests on paper. However, as boys and girls do not report very differently about their digital habits, further explanations are needed to understand why boys' comprehension scores on the reading tests are more negatively affected than those of girls.

One possible explanation may relate to metacognitive comprehension. Several studies (e.g. Ackerman & Goldsmith, 2011; Dahan Golan et al., 2018; Halamish & Elbaz, 2019; Singer & Alexander, 2017a; Singer Trakhman et al., 2017) show that readers are poorly calibrated when asked to judge the medium in which they perform best. This may imply that they are unaware of whether they comprehend better when reading from paper or screen. Many presume that they are better at reading in the digital medium, but, in reality, they comprehend better when reading on paper. The miscalibration is likely to be underpinned by the reading activity itself. As digital reading is perceived to be easy and fast, the reader's sense of achievement is likely to rise, even if this is not the case. For this reason, awareness of expedient reading strategies seems even more important when reading on screen. Considering that boys generally demonstrate lower metacognitive skills in reading, as shown by the Norwegian PISA results (Hopfenbeck & Roe, 2010; Jensen et al., 2019), this may have had a negative effect on their scores on the screen version of the tests. However, as collection of data on students' calibration was not within the scope of this study, such an explanation may not be ascertained. A future study should address more specifically the metacognitive comprehension of boys and girls reading across different modes.

Another possible factor that could have contributed to the widening of the gender gap from paper to screen could be that girls' reading habits are also beneficial for on-screen reading. Several studies confirm that reading traditional extended texts, especially fictional books, is a strong predictor of reading comprehension, even if many reading activities are digitized (Duncan, McGeown, Griffiths, Stothard, & Dobai, 2016; Pfost, Dörfler, & Artelt, 2013). This is in line with the Norwegian PISA results as well, showing that students who report that they read for enjoyment comprehend significantly better than those who do not read. Furthermore, students who report that they prefer reading books on paper outperform students who read books more often on digital devices and students who read books equally often in paper format and on digital devices (Roe, 2020). This indicates that book reading and reading for enjoyment have a positive effect on reading comprehension regardless of presentation mode. As girls spend significantly more time reading for pleasure than boys, who for a large part report that reading is seen as a mere necessity, girls are more likely to develop reading skills that are beneficial to reading across all text presentation media.

Results of several large-scale assessments administered to Norwegian children and adolescents consistently show that the proportion of boys on the lowest comprehension levels is significantly larger than the proportion of girls (Jensen et al., 2019;

Solheim & Gourvennec, 2017). This finding should also be considered when trying to understand the widening of the gender gap in on-screen testing. Research shows that reading from a screen involves a higher cognitive workload and can be more tiring than reading from paper (Wästlund et al., 2005). Especially low-performing students experience a higher workload when reading from a screen (Noyes et al., 2004). Hence, they may be additionally disadvantaged when completing computer-based assessments as compared to similar tasks on paper.

The higher cognitive load associated with the screen condition may be especially true for assessments that involve sophisticated tasks that require sustained attention (Eyre, Berg, Mazengarb, & Lawes, 2017). Further, it may also be related to issues of navigation. Bridgeman, Lennon, and Jackenthal (2003) suggest that the resolution of the monitor and amount of scrolling required by the test-taker could affect performance. They found that students who could see the whole passage of a text without scrolling comprehended better on reading assessments than those who had to scroll to see the full passage. This is in line with the result of Delgado et al. (2018), showing that the advantage of paper-based reading is significant when scrolling is necessary to read texts on screen.

As pointed out by Sanchez and Wiley (2009), scrolling is likely to draw on the limited capacity of the working memory needed for reading. Furthermore, Kingston (2008) argues that reading while scrolling is cognitively different from reading a page. While reading a page, the reader can use spatial memory clues to remember the location (e.g. toward the upper right portion of a page) of information that is pertinent – for instance, when answering a particular question. Parallel clues are not available when scrolling is needed for reading texts on screen. Scrolling constantly changes the spatial frame of reference, which may have a negative effect on the readers' mental reconstruction of the text. By implication, this also has a negative effect on comprehension, as having a good spatial mental representation of the physical layout of a text supports comprehension. Cataldo and Oakhill (2000) found that good comprehenders were more efficient than poor comprehenders at remembering and relocating the order of information in texts. This suggests that there is a relationship between mental reconstruction of text structure and reading comprehension.

The present study did not control for factors that seem to increase the demands of reading on screen, such as the resolution of the monitor and amount of scrolling. Consequently, the extra cognitive load of creating mental representations of texts the spatial frames of which constantly changed may have contributed to the fact that low-performing students participating in this study comprehended worse in the screen condition compared to the paper condition. As the proportion of low-performing students is higher among the boys than among girls, this may have further contributed to the increase of the gender gap from paper to screen. However, a future study should explore this assumption more closely, as comprehension differences across modes for high- and low-performing students were not in the scope of this study.

13.6 Concluding Remarks

As the empirical evidence of children's and adolescent's reading comprehension on paper compared to on screen remains rather sparse, this study adds valuable information about the way delivery mode affects reading comprehension. The study particularly broadens the field by exploring how the gender gap seen in reading is affected. Both the large sample size and the fact that both reading tests showed the same results are strengths of the study. The results of the study have several pedagogical implications. Showing that the gender gap increases when reading on screen, the results confirm that equivalence cannot easily be assumed. This has implications for policy makers, as consideration should be paid to the increasing use of digital technologies in education. Furthermore, care must be taken to ensure the fairness of student assessment. Even though the transition from paper to screen is the same for all students, this study exemplifies how equality in some cases does not contribute to equity (Espinoza, 2007). However, awareness of this matter may promote measures that can be levelling in an educational system aiming to be a 'School for All'.

Although no differences in overall reading performance among 13–14-year-olds as an effect of the delivery mode were found, the results of the present study indicate that different media may affect the reading of students differently. Moreover, students are likely to exhibit different reading behaviour and apply diverse strategies for different reading purposes. However, attention must be paid to what reading behaviour is useful. It is evident that the skimming and scanning strategies readily applied for online information-seeking and entertainment do not benefit all reading situations on screen. On the contrary, in-depth reading strategies are more beneficial for completing digitized reading assessments.

Garland and Noyes (2004) point out that repeated exposure and rehearsal of computer-based information is needed to equate knowledge retrieval with that achievable from paper, and research by Lauterman and Ackerman (2014) shows that encouragement of in-depth processing on screen may reduce the inferiority of screen reading. This has implications for teachers and educators. Children and adolescents need to develop awareness of useful reading behaviour and should be taught effective and expedient strategies for reading on screens. This may contribute to the overall fairness in the assessment situation and eliminate some of the adverse effects of the screen for students susceptible to these. Moreover, both boys and girls, as well as both high- and low-performing students, would benefit from this.

References

Ackerman, R., & Goldsmith, M. (2011). Metacognitive regulation of text learning: On screen versus on paper. *Journal of Experimental Psychology: Applied, 17*, 18–32.

Ackerman, R., & Lauterman, T. (2012). Taking reading comprehension exams on screen or on paper? A metacognitive analysis of learning texts under time pressure. *Computers in Human Behavior, 28*, 1816–1828.

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA and London: SAGE Publications, Inc..

Alexander, P. A. (2013). Calibration: What is it and why it matters? An introduction to the special issue on calibrating calibration. *Learning and Instruction, 24*, 1–3.

Annisette, L. E., & Lafreniere, K. D. (2017). Social media, texting, and personality: A test of the shallowing hypothesis. *Personality and Individual Differences, 115*, 154–158.

Bridgeman, B., Lennon, M., & Jackenthal, A. (2003). Effects of screen size, screen resolution, and display rate on computer- based test performance. *Applied Measurement in Education, 26*, 191–205.

Carr, N. G. (2010). *The shallows: What the internet is doing to our brains*. New York: W. W. Norton.

Cataldo, M. G., & Oakhill, J. (2000). Why are poor comprehenders inefficient searchers? An investigation into the effects of text. *Journal of Educational Psychology, 92*(4), 791–799.

Clinton, V. (2019). Reading from paper compared to screens: A systematic review and meta-analysis. *Journal of Research in Reading, 42*(2), 288–325.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Conway, M. A., Gardiner, J. M., Perfect, T. J., Anderson, S. J., & Cohen, G. (1997). Changes in memory awareness during learning: The acquisition of knowledge by psychology undergraduates. *Journal of Experimental Psychology: General, 126*(4), 393–413.

Dahan Golan, D., Barzillai, M., & Katzir, T. (2018). The effect of presentation mode on children's reading preferences, performance, and self-evaluations. *Computers & Education, 126*, 346–358.

Delgado, P., Vargas, C., Ackerman, R., & Salmerón, L. (2018). Don't throw away your printed books: A meta-analysis on the effects of reading media on reading comprehension. *Educational Research Review, 25*, 23–38.

Dillon, A. (1992). Reading from paper versus screens: A critical review of the empirical literature. *Ergonomics, 35*, 1297–1326.

Duncan, L. G., McGeown, S. P., Griffiths, Y. M., Stothard, S. E., & Dobai, A. (2016). Adolescent reading skills and engagement with digital literacies as predictors of reading comprehension. *British Journal of Psychology, 107*, 209–238.

Espinoza, O. (2007). Solving the equity-equality conceptual dilemma: A new model for analysis of the educational process. *Educational Research, 49*(4), 343–363.

Eyre, J., Berg, M., Mazengarb, J., & Lawes, E. (2017). *Mode equivalency in PAT: Reading comprehension*. Wellington, NZ: New Zealand Council of Educational Research.

Frønes, T. S. (2016). Resultater i lesing [Results in reading]. In M. Kjærnsli & F. Jensen (Eds.), *Stø kurs. Norske elevers kompetanse i naturfag, matematikk og lesing I PISA 2015* (pp. 136–171). Oslo, Norway: Universitetsforlaget.

Frønes, T. S. & Narvhus, E. K. (2011). *Elever på nett. Digital lesing i PISA 2009* [Students online. Digital reading in PISA 2009]. Oslo, Norway: ILS, UiO.

Garland, K. J., & Noyes, J. (2004). CRT monitors: Do they interfere with learning? *Behavior and Information Technology, 23*(1), 43–52.

Halamish, V., & Elbaz, E. (2019). Children's reading comprehension and metacomprehension on screen versus paper. *Computers & Education, 145*, 103737.

Hopfenbeck, T. N., & Roe, A. (2010). Lese- og læringsstrategier [Reading and learning strategies]. In M. Kjærnsli & A. Roe (Eds.), *På rett spor. Norske elevers kompetanse i lesing, matematikk og naturfag I PISA 2009* (pp. 118–137). Oslo, Norway: Universitetsforlaget.

Imsen, G., Blossing, U., & Moos, L. (2017). Reshaping the Nordic education model in an era of efficiency. Changes in the comprehensive school project in Denmark, Norway, and Sweden since the millennium. *Scandinavian Journal of Educational Research, 61*(5), 568–583.

Jensen, F., Pettersen, A., Frønes, T.S., Kjærnsli, M., Rohatgi, A., & Narvhus, E. K. (2019). *PISA 2018. Norske elevers kompetanse i lesing, matematikk og naturfag* [PISA 2018. What Norwegian students know in reading, mathematics and science]. Oslo, Norway: Universitetsforlaget.

Kingston, N. M. (2008). Comparability of computer- and paper-administered multiple-choice tests for K-12 populations: A synthesis. *Applied Measurement in Education, 22*(1), 22–37.

Kong, Y., Seo, Y., & Zhai, L. (2018). Comparison of reading performance on screen and on paper: A meta-analysis. *Computers & Education, 123*, 138–149.

Koriat, A., Ma'ayan, H., & Nussinson, R. (2006). The intricate relationship between monitoring and control in metacognition: Lessons for the cause-and-effect relation between subjective experience and behavior. *Journal of Experimental Psychology: General, 135*(1), 36–69.

Lauterman, T., & Ackerman, R. (2014). Overcoming screen inferiority in learning and calibration. *Computers in Human Behavior, 35*, 455–463.

Mangen, A. (2010). Lesing – på skjerm eller papir; er det så nøye, da? [Reading – on screen or paper; does it really matter?]. *Norsklæreren*, 3–10, 16–20.

Mangen, A., & Kristiansen, M. (2013). Tekstlesing på skjerm: Noen implikasjoner av et digitalt grensesnitt for lesing og forståelse [Reading texts on screen: Implications of a digital interface for reading and comprehension]. *Norsk Pedagogisk Tidsskrift, 97*, 52–62.

Mangen, A., Walgermo, B. R., & Brønnick, K. (2013). Reading linear texts on paper versus computer screen: Effects on reading comprehension. *International Journal of Educational Research, 58*, 61–68.

Martin, R., & Binkley, M. (2009). Gender differences in cognitive tests: A consequence of gender-dependent preferences for specific information presentation formats? In F. Scheuermann & J. Björnsson (Eds.), *The Transition to computer-based assessment: New approaches to skills and implications for large-scale testing* (pp. 75–82). Luxembourg, Luxembourg: OPOCE.

Mascheroni, G., & Cuman, A. (2014). *Net children go mobile: Final report. Deliverables D6.4 & D5.2*. Milano: Educatt.

Mayes, D. K., Sims, V. K., & Koonce, J. M. (2001). Comprehension and workload differences for VDT and paper-based reading. *International Journal of Industrial Ergonomics, 28*, 367–378.

Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2017). *ePIRLS 2016 international results in online informational reading.* Retrieved from Boston College, TIMSS & PIRLS International Study Center website: https://timssandpirls.bc.edu/pirls2016/international-results/

Noyes, J., & Garland, K. J. (2003). VDT versus paper-based text: Reply to Mayes, Sims and Koonce. *International Journal of Industrial Ergonomics, 31*, 411–423.

Noyes, J. M., & Garland, K. J. (2008). Computer- vs. paper-based tasks: Are they equivalent? *Ergonomics, 51*(9), 1352–1375.

Noyes, J., Garland, K., & Robbins, L. (2004). Paper-based versus computer-based assessment: Is workload another test mode effect? *British Journal of Educational Technology, 35*, 111–113.

Oakhill, J. V., & Petrides, A. (2007). Sex differences in the effects of interest on boys' and girls' reading comprehension. *British Journal of Psychology, 98*, 223–235.

OECD. (2009). *PISA 2009 assessment framework. Key competencies in reading, mathematics and science.* Retrieved from https://www.oecd.org/pisa/pisaproducts/44455820.pdf

OECD. (2013). *PISA 2015 draft reading literacy framework*. Retrieved from https://www.oecd.org/pisa/pisaproducts/Draft%20PISA%202015%20Reading%20Framework%20.pdf

OECD. (2017). *PISA 2015 assessment and analytical framework: Science, reading, mathematic, financial literacy and collaborative problem solving, revised edition*. Paris: PISA, OECD Publishing.

OECD. (2019). PISA 2018 reading framework. In *PISA 2018 assessment and analytical framework*. Paris: PISA, OECD Publishing.

Osborne, J. W., & Overbay, A. (2004). The power of outliers (and why researchers should always check for them). *Practical Assessment, Research & Evaluation, 9*(6), 1–12.

Pfost, M., Dörfler, T., & Artelt, C. (2013). Students' extracurricular reading behavior and the development of vocabulary and reading comprehension. *Learning and Individual Differences, 26*, 89–102.

Prensky, M. (2001). Digital natives, digital immigrants. Part 1. *On the Horizon, 9*(5), 1–6.

Rasmusson, M. (2015). Reading paper – Reading screen. *Nordic Studies in Education, 35*, 3–19.

Roe, A. (2013). Norske gutters resultater på nasjonale og internasjonale leseprøver [Norwegian boys' results on national and international reading assessments]. In K. Kverndokken (Ed.), *Gutter og lesing* (pp. 13–31). Bergen, Norway: Fagbokforlaget.

Roe, A. (2020). Elevenes lesevaner og holdninger til lesing [Students' reading habits and attitudes towards reading]. In T. S. Frønes & F. Jensen (Eds.), *Like muligheter til god leseforståelse? 20 år med lesing i PISA* (pp. 107 –134). Oslo, Norway: Universitetsforlaget.

Roe, A., & Vagle, W. (2010). Resultater i lesing [Results in reading]. In M. Kjærnsli & A. Roe (Eds.), *På rett spor. Norske elevers kompetanse i lesing, matematikk og naturfag I PISA 2009* (pp. 59–93). Oslo, Norway: Universitetsforlaget.

Roe, A., & Vagle, W. (2012). Kjønnsforskjeller i lesing – et dybdedykk i nasjonale prøver på åttende trinn fra 2007 til 2011 [Gender differences in reading – A closer look at the national assessments for grade eight from 2007 to 2011]. *Norsk Pedagogisk Tidsskrift, 96*, 425–440.

Sanchez, C. A., & Wiley, J. (2009). To scroll or not to scroll: Scrolling, working memory capacity, and comprehending complex texts. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 51*(5), 730–738.

Shaikh, D. (2004). Paper or pixels: What are people reading online? *Usability News, 6*.

Singer, L. M., & Alexander, P. A. (2017a). Reading across mediums: Effects of reading print and digital texts on comprehension and calibration. *Journal of Experimental Education, 85*, 155–172.

Singer, L. M., & Alexander, P. A. (2017b). Reading on paper and digitally: What the past decades of empirical research reveal. *Review of Educational Research, 87*, 1007–1041.

Singer Trakhman, L. M., Alexander, P. A., & Berkowitz, L. E. (2017). Effects of processing time on comprehension and calibration in print and digital mediums. *The Journal of Experimental Education, 87*, 1–15.

Solheim, O. J., & Gourvennec, A. F. (2017). Gutte- og jentelesere i topp og bunn. En utforsking av leserprofiler [High and low performing readers among boys and girls. Examining reading profiles]. In E. Gabrielsen (Ed.), *Klar framgang! Leseferdighet på 4. og 5. trinn i et femtenårsperspektiv* (pp. 186–203). Oslo, Norway: Universitetsforlaget.

Spencer, C. (2006). Research on learners' preferences for reading from a printed text or from a computer screen. *Journal of Distance Education, 21*, 33–50.

Støle, H., Mangen, A., & Schwippert, K. (2020). Assessing children´s reading comprehension on paper and screen: A mode-effect study. *Computers & Education, 151*, 1–13.

The Norwegian Directorate for Education and Training. (2017a). *Rammeverk for nasjonale prøver* [Framework for the national assessments]. Retrieved from https://www.udir.no/eksamen-og-prover/prover/rammeverk-for-nasjonale-prover/

The Norwegian Directorate for Education and Training. (2017b). *Rammeverk for grunnleggende ferdigheter* [Framework for basic skills]. Retrieved from https://www.udir.no/laring-og-trivsel/lareplanverket/grunnleggende-ferdigheter/rammeverk-for-grunnleggende-ferdigheter/

The Norwegian Media Authority. (2020). *Barn og medier 2020* [Children and Media 2020]. Retrieved from https://medietilsynet.no/globalassets/publikasjoner/barn-og-medier-undersokelser/2020/200402-delrapport-3-gaming-og-pengebruk-i-dataspill-barn-og-medier-2020.pdf

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology, 26*, 1–12.

Wästlund, E., Reinikka, H., Norlander, T., & Archer, T. (2005). Effects of VDT and paper presentation on consumption and production of information: Psychological and physiological factors. *Computers in Human Behavior, 21*, 377–394.

Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica, 41*(1), 67–85.

Chapter 14
The Importance of Parents' Own Reading for 10-Year Old Students' Reading Achievement in the Nordic Countries

Hildegunn Støle, Åse Kari H. Wagner, and Knut Schwippert

Abstract   The Nordic education model of an inclusive school for all aims at giving children equal, and excellent, opportunities for acquiring high levels of reading ability. It is well documented that both students' and their parents' reading interest is closely and positively associated with students' reading achievement. There is therefore cause for concern when reading interests seem to be in decline both among parents and among today's students. Family socio-economic background is also well known to relate strongly to students' reading achievement. Especially children of parents with low education are likely to be deprived of opportunities of beneficial reading activities, such as seeing their parents read, being read to by family members, and learning to enjoy reading for themselves in the early years of school. On the other hand, it is possible that parents who enjoy reading and/or read much at home, provide their children with a basis for acquiring good reading skills, regardless of their educational background. Our article analyses data from four cycles (2001–2016) of the Progress in International Reading Literacy Study (PIRLS), and several Nordic countries, in order to establish whether parental reading can compensate for low parental education levels. We find that parents' reading enjoyment, but not their frequent reading in their spare time, to some degree does compensate for lack of tertiary (high) education. However, if increasingly fewer parents like to read, more children will go without the opportunity to develop reading enjoyment themselves, and this will likely affect more children from low-SES backgrounds than from higher SES-backgrounds.

Keywords   PIRLS · Parents' reading enjoyment · Nordic education · Literacy

H. Støle (✉) · Å. K. H. Wagner
The National Reading Centre, University of Stavanger, Stavanger, Norway
e-mail: hildegunn.stole@uis.no

K. Schwippert
Universität Hamburg, Hamburg, Germany

14.1   Introduction and Background

Reading literacy is vital for the individual's success in education, and for equal participation in school, on the work market, and in society at large. The post-war Nordic model of education, an inclusive School for All (Antikainen, 2006; Telhaug, Mediås, & Aasen, 2006), aimed at giving all students equal opportunity to achieve the skills and knowledge required to enter the workforce (see Chap. 2). The new national curriculum in Norway illustrates the typically high ambitions that the Nordic countries still have for how their school systems should provide all students with "a good basics for participation in every area of education, work and social life" (Norwegian Department of Education, 2017). Schools thereby have a special responsibility to ensure that all children have equal opportunities to learn to read well. As described in Chap. 2, equity in the Nordic educational systems in the twenty-first century is anchored both in main aims of schooling and in students' legal rights to adapted education in free, public schools. This is in line with Espinoza's (2007, p. 354) idea of Equity for equal achievement: "that individuals with similar academic achievement will obtain similar job statuses, incomes and political power". Reaching this goal depends on a school system that does not segregate children of different backgrounds (intentionally or unintentionally).

The degree of success of reading education has been monitored by national and international surveys in many countries during the last decades. The Progress in International Reading Literacy Study (henceforth: PIRLS) is one such large-scale survey, measuring reading literacy among 10-year olds around the world every 5 years. Some of the Nordic countries have participated in all PIRLS cycles since 2001, whereas others have joined in later (see Table 14.1). Norwegian results from PIRLS 2001 revealed a large spread in student reading achievement (Mullis, Martin, Gonzalez, & Kennedy, 2003; Solheim & Tønnessen, 2003), meaning that early reading education during the late 90s had failed in providing equity in Norwegian 4th graders' reading ability. In Norway, the PIRLS 2001 results as well as the Programme for International Student Assessment (PISA) 2000 results, gave rise to an educational policy debate that in turn led to a new national curriculum, implemented in 2006 ("Kunnskapsløftet", often translated to the "Knowledge Promotion").

Exploring 15 EU countries participating in PISA 2000, Gorard and Smith (2004) found that Denmark, Finland and Sweden (Norway is not part of the EU), had less

Table 14.1 Nordic countries participating in PIRLS since 2001 through 2016

	2001	2006	2011	2016
Denmark		X	X	X
Finland			X	X
Iceland	X	X		
Norway	X	X	X	X
Sweden	X	X	X	X

Note: For overviews of all countries participating in each PIRLS cycle, see the respective PIRLS publications, e.g. online at https://timssandpirls.bc.edu

segregation on most indicators than the EU average. These indicators were parental occupation, family wealth, reading performance, students' sex and students' (and parents') country of origin. As mentioned, social fairness in an inclusive school for all has been a political goal in the Nordic countries since the Second World War (Telhaug et al., 2006). Around 2000, it was still hoped that a comprehensive and free education system providing equal opportunities regardless of children's social background (OECD, 2018) would yield equitable outcomes. However, as the PIRLS and PISA results documented relatively large gender and achievement gaps, at least in Norway, it appeared that equitable outcomes were not achieved. Further, Nordic education systems no longer only aim at giving students the same opportunities to acquire basic skills, but focus increasingly on performing better than average in e.g. OECD and other large-scale international skills assessments.

The Nordic countries have relatively small and homogenous populations, ranging from 360,000 in Iceland to 10.1 million in Sweden. The Nordic countries are characterised by high prosperity (Grunfelder, Rispling, & Norlén, 2018; Legatum Prosperity Index Report, 2018), and high levels of parental education (OECD, 2018). This is reflected in the PIRLS 2016 study (Mullis, Martin, Foy, & Hooper, 2017), where the Nordic countries have among the highest scores on students' home resources for learning. The composite variable "Home resources for learning" consisted in 2016 of parents' education, parents' occupation, the number of books in the home, the number of children's books, and "home study support", i.e. Internet connection and/or the child having its own room.[1] These variables are associated with high levels of reading achievement in PIRLS, as they are in most studies of the relationship between student background and reading literacy (Buckingham, Beaman, & Wheldall, 2014). The composite PIRLS home resources for learning variable represents both cultural and economic resources, and is often used as a proxy for socio-economic background in analyses of PIRLS results.

Parents in the Nordic countries report more positive attitudes towards reading than the international average (Mullis et al., 2017). Positive parental attitudes to reading is also associated with higher average reading achievement in PIRLS. The current study aims to investigate whether parents' own reading matters for students' reading proficiency independently of parents' educational level. The study contributes to the research on the relations between home factors and students' reading achievement by exploring Nordic PIRLS results across four cycles, i.e. 15 years. This approach enables conclusions both about trends as well as about consistency (or non-consistency) of our findings.

---

[1] Graphics of "Home Resources for Learning" from the latest PIRLS report (2017) are supplied in Appendix 14.1. The home background questionnaire addresses the parents or guardians of the child. For ease of reading, only the term "parent" is used in this article.

14.1.1   Parental Reading

Parents play an important role in preparing children for learning, not only as providers of resources but also as role models for reading engagement. Parents who enjoy reading may foster the same interest in their children and nurture an emergent positive reader self-concept, associated with high reading achievement in school (Walgermo, Foldnes, Uppstad, & Solheim, 2018). Parental attitudes to reading can thus be an important factor for equity in learning. Rowe (1991, p. 19) expressed it as follows: "regardless of family socio-economic status, age and gender, 'Reading Activity at Home' had significant positive influences on measures of students' reading achievement, attitudes towards reading and attentiveness in the classroom." Since Rowe's findings (1991), however, many things have changed regarding reading activities in the homes. Mullis et al. (2017) found a decline in parental interest in reading from PIRLS 2011 to 2016. Similarly, Norwegian findings from the PISA 2018 show that 15-year olds read less than before in their spare time (Jensen et al., 2019). Especially many boys report that they "never or almost never" choose to read for pleasure.

Adolescents who reported reading fiction performed significantly better on PISA 2009 than those who read other kinds of reading material (magazines, non-fiction, fiction, newspapers and comics) (Jerrim & Moss, 2019). Norwegian children who enjoyed reading and read in their spare time, performed better than those who did not, both on the paper-based PIRLS 2016, as well as on the online informational reading assessment (ePIRLS) in 2016 (Støle & Schwippert, 2017). Mol and Jolles (2014) found that students' enjoyment of reading was socially stratified and related to gender. Children of parents who enjoy reading do better on the PIRLS reading test than their counterparts with parents who are less interested in reading (Mullis et al., 2017, p. 156). Even though Nordic parents in average report positive attitudes to reading, the general decline in parental reading also affected students in the Nordic countries (ibid.).

14.1.2   Parents' Education and Socio-economic Status

Parents' education, their occupation, and family income constitutes a child's socio-economic status (SES) (Buckingham et al., 2014), but according to a meta-analytic review by Sirin (2005), it varies how much each of these factors contributes in predicting a child's academic success. Several studies conclude that parents' education matters substantially, and sometimes is the most salient factor in analyses of the effect of socio-economic status on children's achievement in school (Buckingham et al., 2014; Caro, Sandoval-Hernández, & Lüdtke, 2014; Yang & Gustafsson, 2004), and on reading achievement in particular (Myrberg & Rosén, 2006, 2009). In a Norwegian study of associations between a child's home language, home resources for learning to read, and reading achievement in PIRLS 2016, Strand and Schwippert

(2019) found that parents' education mattered more than books in the home, a factor well known to be associated with economic as well as cultural background, and more than the disadvantage of coming from a non-native language family.

Myrberg and Rosén (2009) explored the indirect, direct, and total effects of parents' education on Swedish 4th graders' reading achievement in PIRLS 2001. They found that the "total effect of parents' education is substantial, but that almost half of this effect is mediated through other variables, i.e. the number of books at home, early literacy activities and emergent literacy abilities…" (Myrberg & Rosén, 2009, p. 695). Myrberg and Rosén (2009) found that even though the direct effect (standardised regression coefficient) of parental education on children's reading achievement was modest, at 0.17, the total effect reached 0.34. This is because well educated parents tend to offer children more books and preschool literacy activities than do parents with only little education (Hemmerechts, Agirdag, & Kavadias, 2017). However, home literacy environments may vary considerably in low SES families (Buckingham et al., 2014; van Steensel, 2006). Positive reading attitudes among parents with low education levels may compensate for a situation of sparse resources and provide children with sufficiently good emergent literacy skills for them to develop into good readers and successful learners.

14.1.3   Books in the Home

Evans, Kelley, Sikora and Treiman (2010) found families' book ownership to matter for students' reading achievement consistently across diverse cultures and at different times in the twentieth century. They found that students from low socio-economic backgrounds gain especially from having access to books at home. Inspecting PISA data (15-year old students) from 42 nations, Evans, Kelley and Sikora (2014) again found book ownership to matter regardless of student background across the national ideologies. Similarly to Rowe (1991), Bus, van Ijzendoorn and Pellegrini (1995) found in their meta-study that children from low SES families gained as much as their wealthier peers from their parents' engaging them in joint book reading prior to school entry. They found significant associations on outcome measures of language growth, emergent literacy, and reading achievement (Bus et al., 1995). There is plentiful evidence that children's book and/or fiction reading is a strong predictor of reading achievement (Cunningham & Stanovich, 1997), also in a twenty-first century, longitudinal study which included children's reading in digital environments (Pfost, Dörfler, & Artelt, 2013), as well as in recent PISA studies (Jerrim & Moss, 2019).

14.1.4   What If Fewer Parents Like to Read?

Many factors, such as parents' educational levels, positive attitudes towards reading and home library, work together in providing children with rich opportunities for developing literacy skills needed for academic success and meaningful societal participation. However, as argued, it is conceivable that parents' engagement in reading is not always related to socio-economic background or their level of education, and thus, that even children of relatively poor backgrounds may have parents who provide them with positive attitudes towards reading. Reversely, it is likely that children adopt negative attitudes towards reading from parents who do not like to read in spite of having long educations. Further, if the decline in parental spare time reading continues, more children will grow up in families in which only little reading occurs, even if their parents actually like reading. Fewer children may benefit from a rich "family scholarly culture" (Evans et al., 2010), regardless of whether their parents are well educated or not.

14.2   This Study

The present study analyses Nordic PIRLS data from all four cycles (2001–2016) to explore associations of parents' educational level, their reading habits, and number of books at home, and students' reading achievement. Cross-sectional studies like PIRLS dip into one cohort of students at a certain point in time, making it difficult to draw conclusions with certainty. Comparing trends and countries, on the other hand, controls for spurious correlations and yield more robust findings than observations from just one survey. When similar results occur across different cohorts over time, it enables researchers to conclude more solidly about the relationship between outcome and explanatory variables. However, the variables explored across cycles need be the same. Therefore, we apply variables of e.g. home resources and parents' attitudes to reading that consist of questions that reoccur in all cycles, rather than applying the PIRLS composite variables which vary somewhat from 2001 to 2016.

   We hypothesise that there is an association between parents' interest for spare time reading, including book ownership, and children's reading achievement, regardless of parents' level of education. As we explore PIRLS results across four cycles, we use parents' education as a proxy for socio-economic status (SES), in accordance with the literature presented in Sect. 14.1.2. (e.g. Caro et al., 2014; Yang & Gustafsson, 2004). Of the three most used SES-factors, i.e. parental income, occupation and education, the latter is the only variable that has been consistently probed throughout PIRLS cycles.

   The composite variable "Home resources for learning" has also varied in content since 2001, which is why we let the single variable of number of books in the home represent home literacy resources in our analyses.

14.3   Methods

We address our research question through a sequence of secondary analyses using data from the Nordic cohorts participating in PIRLS 2001, 2006, 2011 and 2016. The Nordic countries are Denmark, Finland, Iceland, Norway and Sweden, albeit not all participating in every cycle (see Table 14.1). Therefore, the results presented in tables in Sect. 14.4, vary in terms of which Nordic countries appear in each calculation. Below, we describe PIRLS, the variables, and the analytical procedures.

14.3.1   The PIRLS Survey

PIRLS measures 10-year old students' reading literacy much like the better known PISA study does, through a reading test consisting of texts (literary and informational) with questions of comprehension in the form of multiple choice items and constructed response items for which students write a response based on what they have read. As in PISA, some items are repeated across cycles, thereby functioning as anchors for trends analyses. For further descriptions of the design, see PIRLS 2016 assessment framework (Mullis & Martin, 2015).

The PIRLS survey also includes background questionnaires to the school (principal or other school leader), to the teacher of the test language (i.e. English teacher in English-speaking countries, Norwegian language teacher in Norway etc.), to the home (parents or guardian), as well as to the students themselves. Together, the reading test and the background questionnaires give plentiful information about reading achievement and its associations to background factors in and across the participating countries.

In collaboration with each country's National Research Coordinator, Statistics Canada draws a representative sample of the targeted grade 4.[2] In general, the number of children who participate in PIRLS varies little, and around 4000 per country has been quite common.[3] Norway, for example, had 3211 students participating in 2011 and 4354 in 2016 (Gabrielsen & Strand, 2017).

As a general description, PIRLS uses a stratified two-stage cluster sample design. Schools are selected at a first stage, and then, at a second stage, one or more whole classes of students are selected from each of the sampled schools. All students, with very few exceptions, are expected to participate. Strict rules apply for school-level and within-school level exclusions. Methods and procedures concerning sampling, instrument development, data collection and reporting are described in detail in

---

[2] In addition, Norway has included a cohort of 5th grade children since 2006, because these are around 10 years of age, i.e. the same age as 4th graders in Denmark, Finland and Sweden.

[3] Occasionally larger samples are drawn. For example, Sweden had a sample of more than 10,000 in PIRLS 2001 (Myrberg & Rosén, 2009), in order to compare to the 1991 Reading Literacy Study.

separate publications from the various cycles (e.g. Martin & Mullis, 2013; Martin, Mullis, & Hooper, 2017; Mullis et al., 2003).

14.3.2   Variables

14.3.2.1   Parents' Reading

In the international PIRLS reports (e.g. Mullis et al., 2017), parental attitudes are analysed as a composite "Parents Like Reading" scale (see Appendix 14.2, from Mullis et al., 2017, p. 15). Rather than using this composite variable, we inspected "reading frequency" and "reading enjoyment" as separate phenomena (question 10 and 12 in Appendix 14.2). One reason for treating the scales separately, is that the 2016 composite "Parents Like Reading" scale has not been used consistently across the PIRLS cycles (see e.g. the composite PATR variable from PIRLS 2001, in Mullis et al., 2003).

In 2016, the PIRLS Question 10 to parents explores how much time they spend reading for themselves at home any kind of reading material, such as "books, magazines, newspapers, and materials for work (in print or digital media)". We used this frequency scale for parents' reading to group parents dichotomously: parents reading 5 h or less a typical week at home, and parents reading more than 5 h weekly. Group 1 includes parents who read little at home ("1 to 5 hours a week") and those who do not read at all ("less than one hour a week"). Group 2 includes parents who read more than 5 h but also "more than 10 hours a week" at home.

Whereas question no. 10 about reading frequency includes e.g. work documents that are read digitally, there is reason to believe that the next two questions to parents (nos. 11 and 12) about reading enjoyment, are associated by many respondents with fiction reading. Both probe parents about their enjoyment of reading: no. 11 about frequency of reading for enjoyment, and no. 12 about attitudes towards reading. Neither question indicates anything about text type or medium for reading, but it seems likely that the respondent when filling in the Home Questionnaire, will consider question 11 about reading enjoyment as something different from the previous question (no. 10) about general reading frequency of any type of material. We included only questions 10 and (parts of) 12 in our analyses.

From scale 12, we selected the three most salient variables on how much (or little) parents enjoy reading: (12a) "I read only if I have to" (reversed), (12 c) "I like to spend my spare time reading", and (12 h) "Reading is an important activity in my home". The Likert scale contains four categories from "agree a lot" via "agree", "disagree a little" to "disagree a lot". These three variables were combined and a mean was calculated if at least two questions of the three have been answered. The internal consistency for this reading enjoyment scale exceeds the value r tt >0.700 for all cycles and countries, with one exception only (Iceland 2001 r tt = 0.642). Finally, the score was z-transformed into a scale expressing parents' enjoyment of reading with a mean of zero and standard deviation of one for the regression

analyses. The distribution is skewed, since reading is well liked among Nordic parents compared to the PIRLS average.

14.3.2.2 Number of Books in the Home

PIRLS probes the number of books families have at home, both in the student questionnaire and by asking parents in the home questionnaire. We used data from the latter, and we split the scale into a dichotomous variable consisting of group 1, who have 100 books or fewer at home, and Group 2, who have more than 100 books. The question only probes print books and does not include reading material such as magazines, e-books, or children's books.

14.3.2.3 Parents' Educational Level

The PIRLS questionnaire to parents surveys their level of education by asking them to select among nine alternatives ranging from no education to doctorate degrees ("not applicable" is a tenth alternative, see Appendix 14.3). In the Nordic countries, it is common that parents have comparatively high levels of education. Parents having really low levels, i.e. no education or none after primary school, is rather uncommon. Our goal is to find out whether parents' reading can compensate for little parental education, but when exploring low education yet plentiful reading, we found this group too small for conclusions. Therefore, we made a dichotomous variable of education level by combining the two lower levels, primary school or secondary school only, as one, low parental education group. Parents who completed some tertiary education made up the other, high parental education group. For the analyses, we used the highest reported level of education of one parent.

14.3.2.4 Analytical Procedures

For the multivariable analysis, we used multiple linear regression (ordinary least square). In multivariable regression analyses, we included variables known to matter for children's reading achievement: number of books in the home and parents' level of education. These are often associated with social background or SES. We also included two variables less commonly studied: parents' enjoyment of reading and parents' frequency of reading at home. We decided to apply the regression model for the whole population rather than considering the class or school structure, since we are interested in the overall effects in a country and not in average effects in schools or classes. For the calculation of the regression models – and later also the cross-tables and mean differences – we used the IDB-Analyser of the IEA Hamburg. This tool offers the possibility to calculate the appropriate standard errors of the statistics by taking the special structure of the data into account (weighting

and jack-knifing). For all analyses that included reading achievement scores, all five plausible values have been taken into account.

14.4 Results

14.4.1 Hypothesis: Parents' Reading Matters Independently of Their Education Level

Is there an association between parents' reading at home and children's reading achievement, regardless of the educational level the parents have reached? Tables 14.2, 14.3, 14.4, and 14.5 show our findings from the Nordic countries participating in all PIRLS cycles from 2001 through 2016. The dependent variable is the PIRLS student achievement score in overall reading achievement. Please note that all calculations are based on data from 4th grade Norwegian students, who are 1 year younger than 4th graders in the other Nordic countries and whose reading achievement scores therefore are lower than those of the others.

Table 14.2 shows expected student achievement (Intercept) when controlling for number of books in the home and parents' education. It reveals that only one of the two parental reading variables contributed significantly to the reading achievement of the Nordic children who participated in PIRLS 2001. When parents reported that they enjoyed reading, it predicted a significant gain in student score. In Iceland, the expected gained score was approximately 3.6 point, in Norway it was 5.8, and in Sweden 8 points. To illustrate, a gain of 8 points equalled the differences between eight countries (Latvia's average score 545, Canada Quebec, Lithuania, Hungary, the USA, Italy, Germany and the Check republic's average score 537) in 2001 (Mullis et al., 2003, p. 36). Parents' reading frequency, on the other hand, did not

Table 14.2 Multivariable regression Nordic PIRLS 2001

2001	Intercept	Reading enjoyment[b] Coeff.		S.E.	Reading >5 h/ week[c] Coeff.		S.E.	> 100 books[c] Coeff.		S.E.	Min. tertiary education[c] Coeff.		S.E.
Iceland	490.57	3.59	*	1.33	1.41	n.s.	3.29	21.85	*	2.99	33.63	*	2.98
Norway	469.60	5.84	*	2.15	7.81	n.s.	4.14	19.99	*	4.49	29.02	*	4.39
Sweden	546.81	8.04	*	1.02	1.43	n.s.	2.62	14.28	*	3.64	21.09	*	2.85

Expected student achievement (Intercept) in relation to parents' reading enjoyment (high), reading frequency (more than 5 h per week), the number of books in the home (101 or more) and parents' level of education (minimum tertiary)

Notes: [a]Significance is marked with an asterisk*; non-significance as "n.s." Significance level is 5%
[b]For the variable "reading enjoyment", the coefficient indicates the change in the Intercept once parental reading enjoyment increases by one standard deviation
[c]The regression coefficients for the dichotomous variables "reading >5 h/w", ">100 books" and "min. tertiary education" indicate the mean differences in the Intercept compared with the reference group

Table 14.3 Multivariable regression Nordic PIRLS 2006

2006	Intercept	Reading enjoyment[b] Coeff.		S.E.	Reading >5 h/ week[c] Coeff.		S.E.	> 100 books[c] Coeff.		S.E.	Min. tertiary education[c] Coeff.		S.E.
Denmark	526.20	10.04	*	2.27	−5.86	n.s.	3.49	20.13	*	3.79	22.04	*	3.39
Iceland	492.78	7.21	*	1.74	4.52	n.s.	3.03	10.88	*	3.34	28.08	*	2.83
Norway	473.23	5.03	*	1.69	1.91	n.s.	3.15	16.15	*	3.75	29.28	*	3.00
Sweden	529.39	7.39	*	1.75	1.39	n.s.	3.27	18.81	*	3.23	23.22	*	3.03

Expected student achievement (Intercept) in relation to parents' reading enjoyment (high), reading frequency (more than 5 h per week), the number of books in the home (101 or more) and parents' level of education (minimum tertiary)

Notes: [a]Significance is marked with an asterisk*; non-significance as "n.s." Significance level is 5%
[b]For the variable "reading enjoyment", the coefficient indicates the change in the Intercept once parental reading enjoyment increases by one standard deviation
[c]The regression coefficients for the dichotomous variables "reading >5 h/w", ">100 books" and "min. tertiary education" indicate the mean differences in the Intercept compared with the reference group

Table 14.4 Multivariable regression Nordic PIRLS 2011

2011	Intercept	Reading enjoyment [b] Coeff.		S.E.	Reading >5 h/ week[c] Coeff.		S.E.	> 100 books[c] Coeff.		S.E.	Min. tertiary education[c] Coeff.		S.E.
Denmark	531.08	7.37	*	1.57	5.83	*	2.74	21.83	*	2.73	16.70	*	2.78
Finland	546.71	8.43	*	1.40	2.84	n.s.	3.10	15.74	*	3.18	21.61	*	2.98
Norway	484.86	7.92	*	1.71	6.59	n.s.	3.54	15.56	*	3.77	19.36	*	3.65
Sweden	522.83	9.48	*	1.65	−2.09	n.s.	2.90	21.46	*	2.60	23.34	*	2.82

Expected student achievement (Intercept) in relation to parents' reading enjoyment (high), reading frequency (more than 5 h per week), the number of books in the home (101 or more) and parents' level of education (minimum tertiary)

Notes: [a]Significance is marked with an asterisk*; non-significance as "n.s." Significance level is 5%
[b]For the variable "reading enjoyment", the coefficient indicates the change in the Intercept once parental reading enjoyment increases by one standard deviation
[c]The regression coefficients for the dichotomous variables "reading >5 h/w", ">100 books" and "min. tertiary education" indicate the mean differences in the Intercept compared with the reference group

contribute significantly to reading achievement in any of the Nordic countries in 2001.

As expected, we found that books in the home contribute strongly to how well students perform on reading achievement. This is in accordance with previous research, e.g. Evans et al. (2010, 2014) concerning the importance of a home library, i.e. a "scholarly culture" providing children with learning resources, regardless which social class they belong to. Table 14.2 shows that owning more than 100 books yielded an expected gain in student achievement of around 21.9 points in Iceland, 20 points in Norway, and 14.3 points in Sweden (relative to families owning 100 books or fewer). 20 points can be interpreted as approximately half a year of schooling.

Table 14.5 Multivariable regression Nordic PIRLS 2016

2016	Intercept	Reading enjoyment[b]			Reading >5 h/week[c]			> 100 books[c]			Min. tertiary education[c]		
		Coeff.		S.E.	Coeff.		S.E.	Coeff.		S.E.	Coeff.		S.E.
Denmark	519.04	7.44	*	1.46	−2.86	n.s.	3.23	20.10	*	3.38	28.45	*	3.87
Finland	544.87	10.56	*	1.48	3.32	n.s.	2.82	14.68	*	2.68	23.25	*	3.07
Norway (4)	476.21	6.64	n.s.	4.14	0.52	n.s.	9.69	31.63	*	10.68	31.81	*	8.10
Sweden	533.00	8.59	*	1.89	2.85	n.s.	3.66	17.38	*	3.82	26.26	*	3.74

Expected student achievement (Intercept) in relation to parents' reading enjoyment (high), reading frequency (more than 5 h per week), the number of books in the home (101 or more) and parents' level of education (minimum tertiary)

Notes: [a]Significance is marked with an asterisk*; non-significance as "n.s." Significance level is 5%
[b]For the variable "reading enjoyment", the coefficient indicates the change in the Intercept once parental reading enjoyment increases by one standard deviation
[c]The regression coefficients for the dichotomous variables "reading >5 h/w", ">100 books" and "min. tertiary education" indicate the mean differences in the Intercept compared with the reference group

The great effect of parents' level of education was obvious in the Nordic countries participating in PIRLS 2001. This finding is also as expected from research such as that by Myrberg and Rosén (2009), or Strand and Schwippert (2019) analysing PIRLS 2001 data in Sweden and PIRLS 2016 data for Norway, respectively. Table 14.2 shows that Nordic children whose parents had tertiary education, i.e. university level, performed much better than those who did not have highly educated parents in PIRLS 2001. In Iceland, the expected achievement gain was 33.6 points, in Norway 29 points, and in Sweden the gain was 21.1 points.

The finding that parents' reading enjoyment matters for students' reading achievement was true of Iceland, Norway and Sweden in 2001, but is it also in the later cycles, and is it true in the other Nordic countries? Further, is it consistent that it does not matter how often parents read? We followed the same procedure with PIRLS data from Nordic countries in later cycles.

In 2006, Denmark entered the PIRLS assessment. Table 14.3 shows similar results as the calculations of the 2001 data: Also in 2006, parents' reading enjoyment mattered significantly for student achievement when accounting for both the number of books in the home and parents' education. Like earlier, parents' reading frequency did not contribute significantly to student results. Parents who reported to enjoy reading contributed 10 score points on student achievement in Denmark, 7.2 in Iceland, 5.0 in Norway and 7.4 in Sweden. To illustrate, a 10 point gain in average reading achievement would have lifted Denmark's international ranking seven places (Mullis, Martin, Kennedy, & Foy, 2007, p. 37).

Again, of course, the number of books in the home and parents having high educational levels contributed substantially to student achievement. In Denmark, access to a rich home library (101 books or more) was almost as important as having parents with high levels of education, yielding 20.1 points gain in student achievement score (books) and a 22.0 points gain (education) respectively.

PIRLS 2011 again witnessed some changes in the Nordic country participation: Iceland withdrew, while Finland participated for the first time. Still, the calculations based on PIRLS 2011 data confirm the patterns from 2001 to 2006. Parents' enjoyment of reading contributed significantly to student achievement in all four countries, whereas their reading frequency at home did not. This holds true independently of the number of books in the home and parents' level of education. The latter factors contributed more than the reading variables. This is in accordance with the literature on the strong associations of SES-related factors and student achievement.

Interestingly, having plenty of books mattered more to Danish children than having highly educated parents in 2011 (21.8 and 16.7 expected score points respectively). In Denmark 5 years earlier, in 2006, books mattered almost as much as parents' education (20.1 and 22.0 respectively; Table 14.3). Similarly for Sweden in 2011: Plenty of books yielded an expected gain of 21.5 points and high parental education 23.3 points, i.e. a mere couple of points more. This pattern occurs again 5 years later, in 2016, but this time for Norway: Many books gave an expected gain of 31.6 student achievement points; high parental education gave the same-size expected gain of 31.8 points (Table 14.5).

As earlier, the positive outcome of parents' reading enjoyment was far from ignorable in 2011. In Denmark, the expected gain from having parents who enjoyed reading was 7.4 student score points, in Finland it was 8.4, in Norway it was 7.9 and in Sweden 9.5.

Bearing in mind the decline in parents' interest in reading (Mullis et al., 2017) from PIRLS 2011 to 2016, we performed an identical regression analysis also of data from the latest PIRLS cycle in 2016.

In PIRLS 2016 the patterns observed previously, appear again, with one exception: Parents' reading enjoyment ceased to be significant in Norway's grade 4 sample. This might simply be caused by the large errors of measurement (S.E.).[4] In the other three countries, parents' reading enjoyment contributes significantly to student achievement, with an expected gain of 7.4 student score in Denmark, 10.6 in Finland, and 8.6 in Sweden. As before, the amount of time parents spent reading "books, magazines, newspapers, and materials for work (in print or digital media)" did not contribute significantly to student reading achievement.

The number of books in the home (more than 100) and parents' education (high) yield substantial contributions to student achievement; in Norway these variables are equally important with 31.6 score points for books and 31.8 for high education. In Denmark, Finland and Sweden, parents' educational level mattered more than a rich home library in 2016.

---

[4] We checked and found that the large S.E.s are not due to a small sample size or low participation rates. However, the jack-knifing procedure entails that the standard errors are of less importance than in some other calculations.

14.4.2 Does Parents' Reading Enjoyment Matter for Children of Parents with Low Education Levels As Well as for Children of Parents with High Education Levels?

Having established that the variable "reading enjoyment" contributes significantly to Nordic students' reading achievement throughout the PIRLS cycle, we proceeded to explore parental reading enjoyment further. We dichotomised the variable in order to compare across groups of parents with low (maximum secondary) versus high (minimum tertiary) education. The mean scale was thus split into: (1) those parents scoring below the maximum of possible reading enjoyment (low reading enjoyment), and (2) those parents whose scores indicate a maximum mean of possible reading enjoyment (high reading enjoyment). The bar chart in Fig. 14.1 illustrates the relationship between parental reading enjoyment in two education level groups, and student reading achievement in PIRLS 2016 in four countries.

Figure 14.1 illustrates, as expected, that children of parents with high education levels score better on the PIRLS assessment than children of parents with low education (maximum secondary school). However, in all four Nordic countries, parents' reading enjoyment plays a significant role regardless of their education level. The confidence interval bars (confidence level at 95%) in Fig. 14.1 show that there are significant differences in students' reading achievement (y-axis) between children of parents who do not enjoy reading (less than maximum on our reading enjoyment variable) and children of parents who do enjoy reading in their spare time (maximum on reading enjoyment). Notably, the achievement gap is eradicated in both Denmark and Finland between children of parents with low education yet high
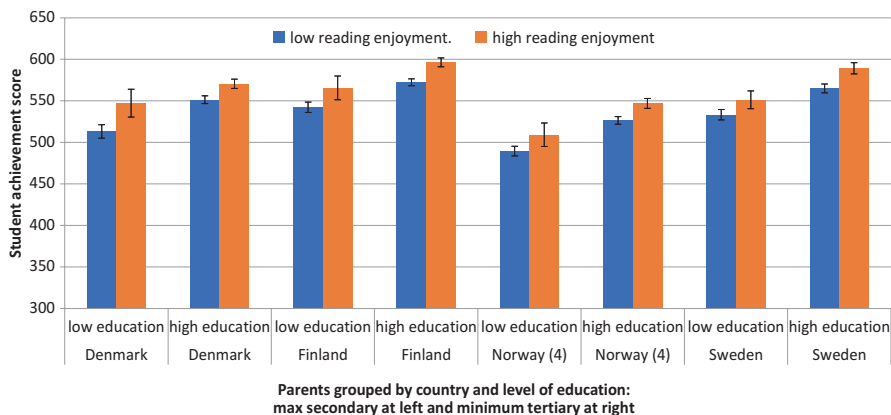
Fig. 14.1 Parental education and reading enjoyment in four Nordic countries in PIRLS 2016. (Note: light grey columns represent parents scoring less than maximum (of mean) on the variable reading enjoyment. Dark grey columns represent parents who score maximum reading enjoyment. For each country, parents with low education levels (maximum secondary) appear at left and parents with high education levels (university level) to the right. The bracketed (4) after "Norway" serves as a reminder that this is the 4th grade sample only

reading enjoyment and children of well-educated parents with low reading enjoyment. In other words, it appears that parental reading enjoyment did compensate for little education, or SES, in these countries in 2016.

14.5   Discussion

In sum, our findings regarding parental education, books, parents' reading and children's reading achievement turn out to be stable across countries and time. Thus, the correlations appear solid and results can be discussed more generally.

Our study of the relationship between home factors and student reading achievement in PIRLS showed that parents' education level matters much for how well students read in the Nordic countries across the assessment cycles. This is well-known from the literature about the contribution of social background on student learning (Myrberg & Rosén, 2009; Strand & Schwippert, 2019). Our analyses also show that having plenty of books in the home (>100) has contributed substantially to student achievement in all Nordic countries that have participated in PIRLS since 2001. This is no surprise, either. Books can be seen as cultural capital, and the home library factor has been found to be consistently associated with reading achievement, regardless of social background (Evans et al., 2010, 2014).

Controlling for the number of books in the home and parents' level of education, we found that parents' reading enjoyment contributes significantly to children's reading achievement as measured in all four PIRLS cycles in all Nordic countries. In contrast, parents' reading frequency, that is how much parents read (e.g. newspapers, work documents, journals, on screen or paper) was not significant for student results in PIRLS. Whereas the questions about reading enjoyment in the questionnaires (both student and home questionnaires) will most likely be associated with reading books for pleasure in the spare time, the question about how much parents read in a typical week includes various genres and both print and electronic media. It thus seems likely that reading enjoyment is associated with the cultural capital of a family, also reflected in the number of books in the home. Further, reading for pleasure is usually associated with long form fiction reading, most typically novels. This is the kind of reading known to be beneficial for children's development of reading ability, in contrast to their reading of other kinds of texts (e.g. Jerrim & Moss, 2019; Pfost et al., 2013).

Decades of research has provided evidence of the strong link between extracurricular reading and reading comprehension. In a longitudinal study, Cunningham and Stanovich (1997) found that children's book reading predicted reading ability 10 years later. Pfost et al. (2013) analysed spare time reading habits in both print and electronic media (also in a longitudinal study), finding that book reading affected reading ability positively, whereas e.g. online chatting had a negative effect on reading achievement. Through regression analyses controlling for a great number of variables, Jerrim and Moss (2019) found a strong link between teenagers' voluntary fiction reading and their reading achievement in the PISA Reading survey from

2009. The same was not true of other types of texts, i.e. magazines, non-fiction, newspapers and comics.

Mol and Jolles (2014) documented that among Dutch secondary school children (*n* = 1071), leisure reading frequency was especially low among students in the pre-vocational track compared to the higher, pre-academic track, and in general, boys reported to read less than girls. However, Mol and Jolles (2014, p. 1) also found that "Non-leisure readers who reported that they enjoyed reading got higher school grades in the higher educational [pre-academic] track", as was also true for girls (but not boys) in the lower educational track. This finding indicates that adolescents who cease to read for pleasure in their teens, may still experience a positive effect of already established positive attitudes towards reading. It resembles our finding that parents' enjoyment of reading matters more for their children's reading performance than does the actual parental reading frequency.

The PIRLS composite scale "Parents like reading" documents a decline in parents' positive attitudes toward reading between 2011 and 2016 in all four Nordic countries. Conversely, more parents report to "not like reading" in 2016 than in 2011 (Mullis et al., 2017, p. 157). However, our study shows that even parents with little education, may contribute positively to their children's reading development if these parents enjoy reading. Therefore, in terms of equity, i.e. overcoming social background (OECD, 2009), it is especially important for children of parents with low education levels that their parents enjoy reading and provide a home library (Evans et al., 2010, 2014; Pfost et al., 2013; Rowe, 1991). These parents may not be able to provide the same support as highly educated parents when it comes to their children's education and/or homework, but if they like reading, they may pass positive attitudes towards leisure reading on to their children and thus help them develop high reading literacy. It is also likely that parents who like reading engage their young children in shared reading and other literacy activities that contribute to vocabulary development and print-knowledge, factors known to benefit early literacy development as well as later reading achievement (Buckingham et al., 2014).

Recently, a decline in spare-time voluntary reading was documented among Norwegian PISA students (15 year-olds) (Jensen et al., 2019). Significantly more teenagers than before reported that they "never or almost never" read in their spare time. This was particularly true of boys. Analysing PISA results, Jerrim and Moss (2019) found that the positive effect of spare-time reading on reading achievement stems from fiction reading only. Other genres and types of reading material (e.g. comics and newspapers) do not contribute to reading development. Therefore it is important that children learn to appreciate fiction early. Pleasure reading can be stimulated by e.g. shared book reading in the home and/or in kindergarten.

Children and adolescents will most likely only read books if they find it pleasurable (Guthrie, Wigfield, Metsala, & Cox, 1999), and children's motivation for reading is developed early (Schiefele, Stutz, & Schaffner, 2016). When parents do not enjoy reading, their children will likely never see them read books, and thus miss out on the opportunity to discover leisure reading as a pleasurable experience for themselves. For the purpose of equity through education, it may be that schools have to take on more of the responsibility of teaching children to enjoy reading long

form texts. This is particularly true of boys from disadvantaged families, since more boys than girls report to not like reading, and more boys than girls perform at the lower end of reading achievement scales in studies like PIRLS and PISA. Kindergarten teachers, school teachers and school librarians can act as adult role models for pleasure reading, giving all children, regardless of their home background or gender, equal chances of obtaining high levels of reading literacy. The foundations for positive attitudes toward reading should be laid already in kindergarten and the early grades of school (Bus et al., 1995), bearing in mind especially those children who do not come from a family culture with positive attitudes towards pleasure reading. In the long term, reading books will benefit both the children themselves and their own children in the future.

14.5.1 Implications

The Nordic ideal of a "School for All" must also cater for those children who are not rich in home resources for learning, be it financial riches or well-educated parents who provide their young ones with early literacy activities at home, help with homework after school entry and the latest in digital devices. A school for students who do not have such resources to draw on (and they are not only immigrant children, neither only poor children), needs to provide these students not with the "same-size" opportunities but with compensating didactics to ensure equitable outcomes of education.

14.5.2 Limitations

Our study only includes the Nordic countries, and the significant associations we have found between parents' reading enjoyment and student achievement might be different in other countries, where e.g. parents report less interest in reading than the very positive attitudes reported among Nordic parents (Mullis et al., 2017).

We would ideally have liked to inspect three groups of parental education levels: low (primary school only or less, i.e. Groups 1–3 in Appendix 14.3), middle (completed secondary education), and high (parents with tertiary education). Few Nordic parents have only primary school or less, and we found that in Scandinavia combined (Denmark, Norway and Sweden), there were fewer than 10 parents in 2016 with only primary school yet reporting high levels of reading enjoyment. This number was too small for analysis, but we encourage researchers to explore the effect of parents' reading enjoyment on student reading achievement in other countries.

Appendices

Appendix 14.1: Composite Variable "Home Resources for Learning" in PIRLS 2016 (Mullis et al., 2017) Comprises Five Items

Item	Options
Number of books at home (students)	0–10 11–25 26–100 101–200 More than 200
Number of home study supports (students)	None Internet connection or own room Both
Number of children's books at home (parents)	0–10 11–25 26–50 51–100 More than 100
Highest level of education of either parent (parents)	Finished some primary or lower secondary Finished lower secondary Finished upper secondary Finished post-secondary education Finished university or higher
Highest level of occupation of either parent (parents)	Has never worked outside home for pay, general labourer or semi-professional (skilled agricultural or fishery worker, craft or trade worker, plant or machine operator) Clerical (clerk or service or sales worker) Small business owner Professional (corporate manager or senior official or professional, technician or associate professional)

Appendix 14.2: Composite Variable "Parents Like Reading" in PIRLS 2016 (Mullis et al., 2017)

In question no. 12, variables (a) and (d) are reversely coded. In our analyses for the present study, we used scales 10 and 12, but not 11. We employed those variables from question 12 that have the strongest association with student reading achievement, i.e. variables a), (c) and (h) (see Sect. 14.3.2).

Question wording for "Literacy at home" section	Options
Question 10 – In a typical week, how much time do you usually spend reading <u>for yourself</u> at home, including books, magazines, newspapers, and materials for work (in print or digital media)?	Less than 1 h a week 1–5 h a week 6–10 h a week More than 10 h a week
Question 11 – When you are at home, how often do you read for your own enjoyment?	Every day or almost every day Once or twice a week Once or twice a month Never or almost never
Question 12 – Please indicate how much you agree with the following statements about reading.	(a) I read only if I have to (b) I like talking about what I read with other people (c) I like to spend my spare time reading (d) I read only if I need information (e) Reading is an important activity in my home (f) I would like to have more time for reading (g) I enjoy reading (h) Reading is one of my favorite hobbies

Appendix 14.3: Parents' Level of Education, from PIRLS 2016 (Mullis et al., 2017)

What is the highest level of education completed by the child's father (or stepfather or male guardian) and mother (or stepmother or female guardian)?

1. Did not go to school
2. Some primary education (ISCED Level 1) or lower secondary (ISCED Level 2)
3. Lower secondary education (ISCED Level 2)
4. Upper secondary education (ISCED Level 3)
5. Post-secondary, non-tertiary (ISCED Level 4)
6. Short-cycle tertiary (ISCED Level 5)
7. Bachelor's or equivalent (ISCED Level 6)
8. Master's or equivalent (ISCED Level 7)
9. Doctor or equivalent (ISCED Level 8)
10. Not applicable

References

Antikainen, A. (2006). In search of the Nordic model in education. *Scandinavian Journal of Educational Research, 50*(3), 229–243.

Buckingham, J., Beaman, R., & Wheldall, K. (2014). Why poor children are more likely to become poor readers: The early years. *Educational Review, 66*(4), 428–446.

Bus, A. G., van Ijzendoorn, M. H., & Pellegrini, A. D. (1995). Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy. *Review of Educational Research, 65*(1), 1–21.

Caro, D. H., Sandoval-Hernández, A., & Lüdtke, O. (2014). Cultural, social, and economic capital constructs in international assessments: An evaluation using exploratory structural equation modelling. *School Effectiveness and School Improvement, 25*(3), 433–450.

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology, 33*(6), 934–945.

Espinoza, O. (2007). Solving the equity-equality conceptual dilemma: A new model for analysis of the educational process. *Educational Research, 49*(4), 343–363.

Evans, M. D. R., Kelley, J., & Sikora, J. (2014). Scholarly culture and academic performance in 42 nations. *Social Forces, 92*(4), 1573–1605.

Evans, M. D. R., Kelley, J., Sikora, J., & Treiman, D. J. (2010). Family scholarly culture and educational success: Books and schooling in 27 nations. *Research in Social Stratification and Mobility, 28*, 171–197.

Gabrielsen, E., & Strand, O. (2017). Rammer og metoder for PIRLS 2016 [Framework and methods in PIRLS 2016]. In E. Gabrielsen (Ed.), *Klar framgang! Leseferdighet på 4. og 5. trinn i et femtenårsperspektiv* [Progress! Reading achievement in the 4th and 5th grades through 15 years] (pp. 13–31). Oslo, Norway: Universitetsforlaget.

Gorard, S., & Smith, E. (2004). An international comparison of equity in education systems. *Comparative Education, 40*(1), 15–28.

Grunfelder, J., Rispling, L., & Norlén, G. (Eds.). (2018). *State of the Nordic region. Nordic Council of Ministers 2018*. Copenhagen, Denmark: Nordisk ministerråd/Nordic Council of Ministers.

Guthrie, J. T., Wigfield, A., Metsala, J. L., Kathleen, E., & Cox, K. E. (1999). Motivational and cognitive predictors of text comprehension and reading amount. *Scientific Studies of Reading, 3*(3), 231–256.

Hemmerechts, K., Agirdag, O., & Kavadias, D. (2017). The relationship between parental literacy involvement, socio-economic status and reading literacy. *Educational Review, 69*(1), 85–101.

Jensen, F., Pettersen, A., Frønes, T.S., Kjærnsli, M., Rohatgi, A., Eriksen, A., & Narvhus, E. K. (2019). *PISA 2018: Norske elevers Kompetanse i lesing, matematikk og naturfag* [PISA 2018: Norwegian students' competencies in reading, mathematics, and science]. Oslo, Norway: Universitetsforlaget.

Jerrim, J., & Moss, G. (2019). The link between fiction and teenagers' reading skills: International evidence from the OECD study. *British Educational Research Journal, 45*(1), 181–200.

Martin, M. O., & Mullis, I. V. S. (Eds.). (2013). *Methods and procedures in TIMSS and PIRLS 2011*. Boston: TIMSS & PIRLS International Study Center Lynch School of Education, Boston College.

Martin, M. O., Mullis, I. V. S., & Hooper, M. (Eds.). (2017). *Methods and procedures in PIRLS 2016*. Boston: TIMSS & PIRLS International Study Center Lynch School of Education, Boston College.

Mol, S. E., & Jolles, J. (2014). Reading enjoyment amongst non-leisure readers can affect achievement in secondary school. *Frontiers in Psychology, 5*(1214), 1–10.

Mullis, I. V. S., & Martin, M. O. (Eds.). (2015). *PIRLS 2016 assessment framework* (2nd ed.). Retrieved from Boston College, TIMSS & PIRLS International Study Center. Website: http://timssandpirls.bc.edu/pirls2016/framework.html

Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2017). *PIRLS 2016 international results in reading*. Boston College, TIMSS & PIRLS International Study Center. Website: http://timssandpirls.bc.edu/pirls2016/international-results/

Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., & Kennedy, A. M. (2003). *PIRLS 2001 International Report*. Boston: International Study Center, Lynch School of Education, Boston College.

Mullis, I. V. S., Martin, M. O., Kennedy, A. M., & Foy, P. (2007). *PIRLS 2006 international report*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Myrberg, E., & Rosén, M. (2006). Reading achievement and social selection into independent schools in Sweden – Results from IEA PIRLS 2001. *Scandinavian Journal of Educational Research, 50*(2), 185–205.

Myrberg, E., & Rosén, M. (2009). Direct and indirect effects of parents' education on reading achievement among third graders in Sweden. *British Journal of Educational Psychology, 79*, 695–711.

Norwegian Department of Education. (2017, to be implemented 2020). *Fagfornyelsen: Overordnet del: Prinsipper for læring, utvikling og danning* [Curriculum Principles for learning and development]. https://www.udir.no/lk20/overordnet-del/prinsipper-for-laring-utvikling-og-danning/

OECD. (2009). *PISA 2009 results: Overcoming social background: Equity in learning opportunities and outcomes (volume II)*. Paris: OECD Publishing.

OECD. (2018). *Equity in education: Breaking down barriers to social mobility*. Paris: PISA, OECD Publishing.

Pfost, M., Dörfler, T., & Artelt, C. (2013). Students' extracurricular reading behaviour and the development of vocabulary and reading comprehension. *Learning and Individual Differences, 26*, 89–102. https://doi.org/10.1016/j.lindif.2013.04.008

Rowe, K. (1991). The influence of reading activity at home on students' attitudes towards reading, classroom attentiveness and reading achievement: An application of structural equation modelling. *British Journal of Educational Psychology, 61*(1), 19–35.

Schiefele, U., Stutz, F., & Schaffner, E. (2016). Longitudinal relations between reading motivation and reading comprehension in the early elementary grades. *Learning and Individual Differences, 51*, 49–58.

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytical review of research. *Review of Educational Research, 75*(3), 417–453.

Solheim, R.G., & Tønnessen, F. E. (2003). *Slik leser 10-åringer i Norge: En kartlegging av leseferdigheten blant 10-åringer i Norge 2001* [How 10-year olds in Norway read: An assessment of reading literacy among 10-year olds in Norway 2001]. UiS: Senter for leseforskning [University of Stavanger: Centre for reading research].

Støle, H., & Schwippert, K. (2017). Norske resultater fra ePIRLS – online informational reading [Norwegian results from ePIRLS – Online informational reading]. In Gabrielsen (Ed.), *Klar framgang! Leseferdighet på 4. og 5. trinn i et femtenårsperspektiv* [Progress! Reading achievement in the 4th and 5th grades through 15 years] (pp. 50–74). Oslo, Norway: Universitetsforlaget.

Strand, O., & Schwippert, K. (2019). The impact of home language and home resources on reading achievement in ten-year-olds in Norway; PIRLS 2016. *Nordic Journal of Literacy Research, 5*(1), 1–17.

Telhaug, A. O., Mediås, O. A., & Aasen, P. (2006). The Nordic model in education: Education as part of the political system in the last 50 years. *Scandinavian Journal of Educational Research, 50*(3), 245–283.

The Legatum Prosperity Index Report. 2018 (12th edition). *Creating the pathways from poverty to prosperity*. London: Legatum Institute. https://prosperitysite.s3-accelerate.amazonaws.com/2515/4321/8072/2018_Prosperity_Index.pdf

Van Steensel, R. (2006). Relations between socio-cultural factors, the home literacy environment and children's literacy development in the first years of primary education. *Journal of Research in Reading, 29*(4), 367–382.

Walgermo, B. R., Foldnes, N., Uppstad, P. H., & Solheim, O. J. (2018). Developmental dynamics of early reading skill, literacy interest and readers' self-concept within the first year of formal schooling. *Reading and Writing, 31*(6), 1379–1399. https://doi.org/10.1007/s11145-018-9843-8

Yang, Y., & Gustafsson, J.-E. (2004). Measuring socioeconomic status at individual and collective levels. *Educational Research and Evaluation, 10*(3), 259–288.

Part IV
Critical Observations and Looking Ahead

Chapter 15
The Black Box of Nordic Education Held Against the Light of Large-Scale International Assessment Resources—A Critical Commentary

Fritjof Sahlström

Keywords Black box · Nordic education · Large-Scale International Assessment · Classroom interaction

This book answers the following general question: when it comes to the impact of socio-economic status (SES) on student results in the context of the so-called Nordic model, what can we learn from large-scale international student assessments? The findings presented are not only new and valuable, but they also raise critical questions, some of which I will discuss below.

Being both insightful and problematic is a feature that the chapters in this volume share with much other education research. Whichever way an educational problem, such as equity, is approached, it seems that one runs the risk of getting stumped by the overwhelming complexity of educational processes. However, being insightful and problematic at the same time is what serious research in the field of teaching and learning looks like. As David Berliner (2002) puts it succinctly in *Educational Researcher*, 'Education is the hardest science of all' (p. 18). Berliner's text was written in response to government expectations of education science to deliver so-called hard results. In the article, Berliner (2002) argues vehemently for a more reflexive view of educational science.

Almost 20 years later, I think Berliner still has something to say in relation to the chapters in this volume. 'The remarkable findings, concepts, principles, technology, and theories we have come up with in educational research are a triumph of doing our damndest with our minds. We have conquered enormous complexity', Berliner

F. Sahlström (✉)
Åbo Akademi University, Turku, Finland
e-mail: Fritjof.Sahlstrom@abo.fi

(2002) writes (p. 20). I think he is correct in this observation and that it also pertains to this book and to the many readings I hope it will get.

It is in this context that the volume should be read and appreciated—as a serious, empirical effort to deepen our knowledge of how schooling in Nordic countries contributes to equity and equality. It has its challenges, of which I think the limited attention paid to Nordic contextualisation is the most substantial, but it is still very much a worthwhile read, particularly for anyone interested in how education for all turns out to be more for others and not always the ones we are aiming for.

The volume fits into the category of black box literature; the basic idea here is that some kind of input goes into a system and gets chewed within the inner workings of this system, and then some kind of output comes out at the other end. The system is the black box. The research interest is in the relation between the input and the output, rather than in what goes on under the hood of the system.

The term 'black box' might initially bring to mind the crash investigation method, in which technical log data from the black box of an airplane or some other mode of transport is investigated in order to understand why an accident occurred. Reading the volume in this way would be to intentionally misunderstand it for several reasons. One of them is that black box investigations of this kind are carried out when things fall to the ground or sink. In general, Nordic education has not crashed. In fact, it is doing quite well, as it continues to fly the post-modern skies of a rapidly changing world whilst encountering some turbulence on the way.

This brings us to the related but much more difficult approach to understanding processes, namely, trying to determine not how and why things crash but how they actually work. In the case of this volume, how does education for all in Nordic countries stay in the air, where is it heading, and why? This is the pursuit of the chapters of this volume, and their shared chosen approach is to dig deeper into the findings of international large-scale assessment studies.

Understanding how things work is much more difficult than understanding why they fail. Despite doing their best with what is available and succeeding well in doing so (more about this later), large-scale assessments cannot really claim to access any inner workings of education. Rather, what we have at hand are studies which, in essence, are refined and sophisticated input–output models put to work. The chapters are black box studies in the sense of systems engineering and economic production functions, in which inputs (e.g. money spent per pupil, facilities, teacher qualifications) go into a box called *schools*, and outputs emerge (e.g. test scores, skilled and knowledgeable high school graduates, humane and community engaged adults). Still, the matter of how inputs are converted into outputs within the black box continues to be unknown (Cuban, 2016).

The research is focused on illuminating the effectiveness of Nordic education in greater detail than what has been done in prior research in relation to matters of equity and equality. This is a fine ambition. In Chap. 2, the authors present this ambition quite explicitly in response to criticism of large-scale assessment:

> Only the use of standardised and internationally comparable instruments makes it possible as objectively as possible to assess the performance of school systems and thereby give a non-biased indication of the level of achievement of equity and equality in the educational

system, at least in part. The SES of students as a psychometric construct is defined in various large-scale studies using a conglomerate of different variables and is linked to students' performance in order to obtain scientifically justified statements. (Buchholtz, Stuart & Frønes, 2020, p. 35)

In addition to being ambitious, it is also quite bold to state at the outset that the chosen instruments in the chapters are the **only** way forward. In principle, there are some quite considerable doubts in relation to any approach being the **only** one within educational sciences. Many researchers, also within the quantitative paradigm, would disagree in principle. So do I, but there is no point in extending that argument here; the arguments are well known. No educational research approach can claim an epistemic monopoly, and even without monopoly claims, no educational research approach can claim to be objective. However, overseeing these principal problems is worthwhile to get to other and more interesting empirical problems. Of these, there are many. Here, I will focus on two aspects: (1) the understanding of the Nordic model displayed in the chapters and issues pertaining to this matter and (2) the understanding of the inner workings of the systems that are claimed to produce the results compared in the anthology.

15.1   The Nordic Model in the Chapters

In my reading, the book is interested in the following two questions: Does participation in education reduce or widen socio-economic differences? Is there a difference between Nordic countries in the amount of this widening or weakening of difference?

The basic conceptual model is that SES-related differences are the dependent variables, and the variations between countries are the independent variables. For this kind of comparison to be meaningful, the underlying premise is that the Nordic model is stable enough and shared enough to be considered similar in all Nordic countries. Furthermore, within this model, it is presumed that the aim in relation to equity and equality is the same, which would be to reduce SES-related differences. The question to consider, therefore, is whether there is country variation.

This is simply put, but to be as clear as possible.

Reflecting their shared methodological paradigm, the authors, in general, are interested in and skilled at data analysis, and the chapters reflect their depth of knowledge of quantitative analysis. Understandably, issues of contextual and conceptual framing are not developed at the same level of sophistication. The chapters all have an initial section on equality and equity, but fairly little is said in these sections that would substantiate the so-called Nordic model. The not-so-extensive analysis of the Nordic model is also evident in the limited amount of contextualised reasoning of the differences found. What are the consequences of the Nordic model, what are the country consequences, and what is something else? How can one discern the first two from the third? For a volume focusing on the Nordic model, there is surprisingly little discussion of the actual model.

Table 15.1 Chapters and included Nordic countries

Chapter	Author/s	Countries
4	Julius K. Björnsson	Denmark, Finland, Iceland, Norway, Sweden
5	Kajsa Yang Hansen, Jelena Radišić, Xin Liu and Leah Natasha Glassow	Denmark, Finland, Norway, Sweden
6	Anubha Rohatgi, Jeppe Bundsgaard and Ove E. Hatlevik	Denmark, Norway
7	Trude Nilsen, Ronny Scherer, Jan-Eric Gustafsson, Nani Teig and Hege Kaarstein	Sweden, Norway
8	Ronny Scherer	Denmark, Finland, Iceland, Norway, Sweden
9	Guri Anne Nortvedt, Karianne Berg Bratting, Oksana Kovpanets, Andreas Pettersen and Anubha Rohatgi	Norway
10	Ole Kristian Bergem, Trude Nilsen, Oleksandra Mittal and Henrik Galligani Ræder	Norway
11	Jelena Radišić and Andreas Pettersen	Sweden, Norway
12	Tove Stjern Frønes, Maria Rasmusson and Jesper Bremholm	Sweden, Norway, Denmark
13	Ragnhild Engdal Jensen	Norway
14	Hildegunn Støle, Åse Kari H. Wagner and Knut Schwippert	Denmark, Finland, Iceland, Norway, Sweden

The limited attention given to the Nordic model as such could have been balanced by an empirical analysis at the Nordic level, in which breadth in Nordic empirical work would be a counterweight to limits in model-level contextualisation. Despite some chapters succeeding in doing so, this is not fully the case. As shown in Table 15.1, three of the eleven empirical chapters (Chaps. 4, 8 and 14) contain all Nordic countries, i.e. Denmark, Finland, Iceland, Norway and Sweden. These three chapters are the only ones that include Iceland. Chapter 5 includes the remaining four countries. One chapter, Chap. 12, includes the three remaining countries, three chapters (Chaps. 6, 7 and 11) are two-country comparisons and three chapters (Chaps. 9, 10, 13) include Norway only. Norway is the only country included in all chapters. The tally for the countries is as follows: Iceland: 3 chapters, Finland: 4 chapters, Denmark: 6 chapters, Sweden: 7 chapters and Norway: 11 chapters.

Against the background of the focus on Nordic equity, the distribution of countries is a little surprising, particularly when considering the highly empirical character of the research presented.

The skewered inclusion of empirical cases would not necessarily have to be problematic. One could argue, in principle, that the Nordic model is robust enough to be present in all countries and that because of the availability of Norwegian data to Norwegian researchers, it makes sense to use Norway more than what is arithmetically plausible. Furthermore, one could argue that the variation in comparative combinations is warranted against the background of the subject matter to be analysed, and that the arguments for this could have been presented in each of the chapters and in the introduction. As mentioned, the authors of the volume have decided

not to fully pursue this possibility of extending the discussion of the Nordic model. It is also the case that the empirical comparative possibilities have not been fully explored. There are no chapters in which the differences and similarities found are discussed at a more analytical level than the mention of and possible short discussion of intra-country differences.

In combination, the scarcity of both theoretical and empirical discussion of the Nordic model sometimes hampers the scope of the claims made on the basis of the studies carried out, leading to chapter conclusions that are not always as focused on either the Nordic model or country differences as would have been possible. As a consequence, the book is not as succinct as it could have been in relation to the Nordic model and whether its continued existence is an empirical or conditional question.

In the interesting first chapter of the book, written by Nils Buchholtz, Amelie Stuart and Tove Stjern Frønes, there is a thorough discussion of the concepts of equity, equality and diversity from conceptual and philosophical points of view. This analysis is reflexive and widely read, and it contributes to the field with valuable insights and points of view. The conceptual discussion is followed by a discussion per country of what is presented as 'educational policy measures and historical developments in the individual Nordic countries in connection with equity, equality and diversity when dealing with marginalised groups' (Chap. 2, p. 25).

Unfortunately, this ambition falls somewhat short, partly as a consequence of an informed choice, in which lack of space is argued to warrant a focus on diversity primarily. Partly, it is also a consequence of what could be argued to be an under-theorising and under-contextualising of the similarities and differences within the Nordic model in relation to reducing socio-economic differences.

There is substantially more to Nordic SES variation than cultural diversity, and both the introductory chapter and the volume as a whole would have benefitted from a more careful discussion of these matters. As an example, in an extensive review article, Dovemark et al. (2018) from the Nordforsk Excellence Center Justice through Education (JustEd). write in *Education Inquiry* about the changes in Nordic comprehensive education, in a special issue dedicated to Nordic education and social justice, In this article, Dovermark and her colleagues argue that deregulation, marketisation and privatisation have implied many changes in Nordic education, in which the idea of a knowledge economy is replacing the previous welfare narrative but in which degrees of change vary in relation to political and historical contexts. Dovemark and colleagues found that the emerging differences within the Nordic model were found mainly in relation to marketisation and privatisation. Norway and Finland continue to have public education markets, whereas Sweden and Denmark provide a wide range of options. The questions of privatisation were dealt with differently in all five countries because in Sweden and the private educational providers for comprehensive schooling have a far more central role in the educational system than in the other countries. The authors argue that the role of profit-making, which has been enabled in Sweden, may be considered one of the biggest changes in relation to the original model of a uniform comprehensive school, contributing to emerging patterns of social differentiation.

I would have appreciated more of this kind of reasoning in this volume. I have a full understanding of the difficulties in trying to achieve this, but to me, it is precisely this dimension that I find interesting. As the volume stands now, the care paid to the quantitative analyses is not fully rewarded with a similarly insightful contextualisation, neither in relation to the Nordic dimension nor in relation to socioeconomic differences. Yes, there is variation between the chapters, and there are chapters in which the differences between the countries found in the results section are also followed up by a discussion of both the Nordic dimension and socioeconomic aspects (e.g. Chaps. 4 and 14). Overall, however, there is an imbalance. This imbalance takes away some of the value of the volume and puts the burden of making conclusions quite heavily on the shoulders of readers.

15.2   How Understanding the Inside Might Add to Input–Output Studies

As already mentioned above, another possible angle of approach to this volume for a critical reader is to reflect on whether the black box approach taken, i.e. the input–output model discussed above, can be a reasonable way of understanding how a system works; knowing the inner workings of the system also matters for how equity and equality are achieved. A 2018 article by Kirsti Klette and her Nordic colleagues sets out to analyse issues of education and justice from the inside, with a review of empirical analyses drawing on video recordings of Nordic secondary classrooms. Being one of the authors, I will repeat the gist of the argumentation in that article here.

Nordic classrooms share a societal expectation that equal opportunities will be provided for all within the framework of comprehensive schooling. Opportunities to engage in meaningful discursive practices and learning activities are considered key factors in high-quality schooling and education around the world. This interactive and discursive view of learning underscores the power of recurring face-to-face interaction and communication amongst peers, in which language, conceptual familiarity and understanding are seen as critical tools for learning (Sfard, 2008).

As for opportunities for student participation in Nordic classrooms, previous research presents a mixed picture. Several studies show that Nordic classrooms provide ample opportunities for students to speak out and influence classroom discourse, more so than in other countries. However, whilst student engagement and student-active ways of working might be key features of Nordic classrooms, there are also differences within and across Nordic countries. Simola, Kauko, Varjo, Kalalahti and Sahlström (2017) describe Finnish classrooms as places where a substantial amount of time is used on individual tasks, with few opportunities for students to talk. Klette and Ødegaard (2015) argue that Norwegian classrooms support student questioning and engagement; however, student utterances are often used for practical and procedural purposes rather than for cognitively demanding enquiries.

Analysing Swedish mathematics classrooms, Emanuelsson and Sahlström (2008) use the term 'the price of participation' (p. 205) to discuss the relation between the cognitive and communicative aspects of classroom learning that include a high degree of student involvement.

Classrooms today have been connected through extensive digitalisation via laptops, tablets and smartphones, which also bring new and multifaceted possibilities for gaining access to different kinds of content. The rapid and massive connection of classrooms has, to a large extent, been achieved through students bringing their own devices into classrooms. Mobile gadgets, particularly mobile phones, have become a crucial part of the everyday life of young people. Using social media applications, such as Snapchat, WhatsApp, Instagram and Facebook, is a common activity for most of today's students, both in the classroom and outside (Paakari, Rautio, & Valasmo, 2019).

In terms of student engagement, the phone screen enables interactive participation (e.g. written and visual messaging, liking, sharing, browsing) parallel to the teaching, without directly interfering with the teacher's presentation and without violating or threatening the overall participation expectations for students in whole-class teaching segments. Students can and do interact with people outside their classrooms. Compared with talking during the lesson, which previously represented the opportunity for peer-to-peer interaction in whole-class teaching segments, the phone presents a significantly smaller disturbance to the teaching. However, phone use also means that students are less accessible for peer talk, and classrooms become less inclusive as interactional spaces (Sahlström, Tanner, & Olin-Scheller, 2019).

The continued use of whole-class teaching at a time of rapid digitalisation makes sense in relation to the participation constraints of whole-class interaction. The individual access for all students to their devices provides them with the opportunity to participate in interactions in parallel with whole-class teaching, in which their opportunities for participation have always been and continue to be limited, without directly disturbing the teaching. For this reason, the *de facto* digitalisation of secondary-school classrooms via students' mobile phones and laptops seems to conserve rather than change whole-class teaching as a general pattern. Somewhat ironically, the digitalisation of classrooms, which was expected to drive pedagogical change and lead to increased inclusion, seems to have had the opposite effect. There is a considerable increase in student communicative acts within the context of classrooms, but the interactions do not seem to be the kind that straightforwardly supports either learning or equality (Sahlström, Tanner, & Valasmo, 2019).

In general, whole-class teaching as it is currently being practiced in Nordic classrooms does not seem to be particularly conducive to creating equal opportunities for participation. The interactional logic of basic participation frameworks and the turn allocation practices in classroom interactions seem to promote difference rather than equality; this makes it difficult for whole-class teaching to provide what it is aiming for, namely, equal opportunities for all students to develop their communicative and discursive skills and capacities within and beyond their school subjects. Interestingly, there seems to be a dissonance between Swedish and Norwegian teachers' active encouragement of student engagement and participation, on the one

hand, and the persistence of rather stable and teacher-dominated interaction patterns, on the other hand.

In Nordic classrooms, individualised teaching and the presence of digital technology-based 1:1 solutions seem to weaken rather than strengthen social justice and equity. The increased access students have to content unrelated to the classroom has reduced classroom equality, as students have chosen to remove themselves from the learning context. As other studies suggest, individual choice by students in whole-class teaching tends to increase rather than decrease difference (Dalland & Klette, 2016; Österlind, 1998; Sahlström, 1999). Furthermore, new technological devices paired with traditional teaching may actually limit access both to learning-relevant content and to learning-relevant discourse. As students navigate their personal pathways through the Internet in the classroom, the institutional boundaries between the classroom, school and everything else have become blurred.

Within their limits, findings such as these pose some serious challenges for the Nordic welfare society vision of classrooms as core societal hubs for justice and equality. The inner workings of classrooms seem to contain inherent constraints that do not support equitable student engagement. Åse Hansson (2011) makes the argument that teaching seems, in certain ways, to facilitate pedagogical segregation rather than pedagogical inclusion. Furthermore, the way Nordic classrooms have responded so far to the massive digitalisation of society seems to pose further questions rather than provide the needed answers—questions such as what do we need classrooms for after all, and what should teachers and students be doing in them?

15.3   The Most Difficult Science of All

The reason for spending as much space as above on the organisation of classroom interaction is to point out that if one is interested in what happens in the black box of Nordic education as this volume indeed is, then it seems to make sense to recognise that the teaching that is supposed to mediate equity and equality seems to be rigged towards mediating the opposite. The chapters in the volume do not and should not be expected to try to account for the inner workings of the black box in the sense of paying attention to teaching and learning processes as such because the chosen focus is elsewhere. This is understandable, as the focus is on what can be learned from large-scale international assessment studies. As with any choice, however, something is also lost when a certain method and focus have been chosen.

To be fair, from a scientific point of view, the choices made by researchers approaching the classroom from within too often implies the exclusion of the possibility of considering the relevance of analysis from other perspectives, such as international large-scale assessment studies. At a general level, results at the level of PISA rankings and other easily available findings are quite often used for contextualisation, but engaging in analysis and dialogue beyond one's own paradigm is uncommon. As one of the many possible examples, in the field of policy analysis, certain empirical matters are sometimes addressed with a quite light hand, resulting

in excellent historical and structural contextualisation but with rather sweeping treatments of empirical findings. At the other end of the spectrum, within interaction-focused classroom research, it is quite common to have lengthy discussions of inter-actional structures inside schools but with very limited or no discussion of how certain structural features come into play at the individual level.

To a certain extent, these challenges are caused by education indeed being the most difficult science of all, as mentioned in the introduction (Berliner, 2002). Despite the bold statement made by the authors of this book, seeing international large-scale assessment as the only and objective instruments (page 35) of approaching equity in education research, I think that the inclusion of a commentary chapter from such a different paradigm as my own tells a different story. To me, the volume benefits from being read in line with Berliner (2002), who writes that '… ethnographic research is crucial, as are case studies, survey research, time series, design experiments, action research, and other means to collect reliable evidence for engaging in unfettered argument about education issues. A single method is not what the government should be promoting for educational researchers. It would do better by promoting argument, discourse, and discussion' (page 20).

I think all the chapters in this book contribute to precisely the kind of argument and discussion Berliner asks for. More work is needed quickly. Research and learning are slow processes, whereas reality is fast. One of the aspects not discussed much in the volume is digitalisation and how digitalisation is related to the Nordic model, education and SES. In the last 10 years, human sociality has been transformed by the opportunities provided by screen interaction. The COVID-19 pandemic has further underscored that the four walls of the equality-constructing classroom we used to take for granted need to be seen in a different light. The walls have not crumbled but have become permeable, allowing for a literal flow of content in and out of what quite recently used to be closed spaces. If and how these new dimensions of education matter are for us in educational research to find out. When doing so, we will be in good company with, amongst others, the authors of this volume. This, I believe, is something to look forward to and has been made much easier by the work done in this volume.

References

Berliner, D. (2002). Comment: Education: The hardest science of all. *Educational Researcher, 31*(18), 18–20. https://doi.org/10.3102/0013189X031008018

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* (pp.13–41). Cham, Switzerland: Springer.

Cuban, L. (2016, September 2). *Inside the black box of the classroom*. https://larrycuban.wordpress.com/2011/10/16/inside-the-black-box-of-the-classroom/

Dalland, C. P., & Klette, K. (2016). Individual teaching methods: Work plans as a tool for promoting self-regulated learning in lower secondary classrooms? *Education Inquiry, 7*(4), 381–404.

Dovemark, M., Kosunen, S., Kauko, J., Magnúsdóttir, B., Hansen, P., & Rasmussen, P. (2018). Deregulation, privatisation and marketisation of Nordic comprehensive education: Social changes reflected in schooling. *Education Inquiry, 9*(1), 122–141. https://doi.org/10.1080/20004508.2018.1429768

Emanuelsson, J., & Sahlström, F. (2008). The price of participation. Teacher control versus student participation in classroom interaction. *Scandinavian Journal of Educational Research, 52*(2), 205–223.

Hansson, Å. (2011). *Ansvar för matematiklärande. Effekter av undervisningsansvar i det flerspråkiga klassrummet*. [Responsibility for mathematics learning. Effects of teaching responsibility in the multilingual classroom]. University of Gothenburg, Faculty of Education.

Klette, K., & Ødegaard, M. (2015). Instructional activities and discourse features in science classrooms: Teachers talking and students listening or…? In K. Klette, O. K. Bergem, & A. Roe (Eds.), *Teaching and learning in lower secondary schools in the era of PISA and TIMSS* (pp. 17–31). Cham, Switzerland: Springer.

Klette, K., Sahlström, F., Blikstad-Balas, M., Luoto, J., Tanner, M., Tengberg, M., et al. (2018). Justice through participation: Student engagement in Nordic classrooms. *Education Inquiry, 9*(1), 57–77.

Österlind, E. (1998). *Disciplinering via frihet: elevers planering av sitt eget arbete* [Discipline through freedom: Students' planning of their own work]. Unpublished doctoral dissertation. Acta Universitatis Upsaliensis.

Paakari, A., Rautio, P., & Valasmo, V. (2019). Digital labour in school: Smartphones and their consequences in classrooms. *Learning, Culture and Social Interaction, 21*, 161–169.

Sahlström, F. (1999). *Up the hill backwards. On interactional constraints and affordances for equity-constitution in the classrooms of the Swedish comprehensive school* (Publication No. 85). Doctoral thesis, Uppsala University. Acta Universitatis Upsaliensis.

Sahlström, F., Tanner, M., & Olin-Scheller, C. (2019). Smartphones in classrooms: Reading, writing and talking in rapidly changing educational spaces. *Learning, Culture and Social Interaction, 22*, 311–331. https://doi.org/10.1016/j.lcsi.2019.100319

Sahlström, F., Tanner, M., & Valasmo, V. (2019). Connected youth, connected classrooms. Smartphone use and student and teacher participation during plenary teaching. *Learning, Culture and Social Interaction, 21*, 311–331. https://doi.org/10.1016/j.lcsi.2019.03.008

Sfard, A. (2008). *Thinking as communicating: Human development, the growth of discourses, and mathematizing*. Cambridge, UK: Cambridge University Press.

Simola, H., Kauko, J., Varjo, J., Kalalahti, M., & Sahlström, F. (2017). *Dynamics in education politics understanding and explaining the Finnish case*. London: Routledge.

Chapter 16
Equity, Equality and Diversity in the Nordic Countries—Final Thoughts and Looking Ahead

Tove Stjern Frønes, Andreas Pettersen, Jelena Radišić, and Nils Buchholtz

In the process of preparing this volume, especially in our review of previous scientific work on the Nordic model of education, it appeared that different researchers approached the topic primarily in the form of historical–political policy analyses (Telhaug, Mediås, & Aasen, 2006) and through the qualitative description of individual country portraits or the differences between these (e.g., Antikainen, 2006; Blossing, Imsen, & Moos, 2014; Lundahl, 2016). In these previous analyses, the question was raised whether a common Nordic model of education can be identified at all and to what extent neoliberal policies and broader globalisation trends affect the further development of education systems in the Nordic countries. The latter has especially been discussed in light of the increased competition between these systems emerging currently, here running against the common thread that was adopted shortly after World War II. In contrast to the works mentioned above, this book explicitly chose a quantitative empirical approach to the topic, linked with the attempt to indicate, measure and evaluate educational equity across the Nordic countries using data from large-scale assessment studies. Thus, the approach of this book was more data driven and descriptive than oriented on the political question of whether a common model exists.

The chapters in this volume mostly comprise analyses of well-established international large-scale assessments (ILSAs), such as PISA, TIMSS, TALIS, ICILS and PIRLS, all of which are assessments generating data with an especially high level of quality (Gustafsson, 2018). Their impact is mostly recognised in the context of evaluating the efficacy of educational systems, yet they remain an essential source of information in connection to monitoring student learning outcomes. The latter is especially found in the observation of the within-country achievement trendlines in areas such as reading, mathematics and science. In addition, the studies gather a large body of data in connection to the students, schools and the classrooms, while

T. S. Frønes (✉) · A. Pettersen · J. Radišić · N. Buchholtz
Department of Teacher Education and School Research, University of Oslo, Oslo, Norway
e-mail: t.s.frones@ils.uio.no

the scales used have been subjected to rigorous quality assurance processes (Rutkowski, von Davier, & Rutkowski, 2014), allowing for a comparison of the data across different country-specific contexts. Of course, the use of large-scale studies is not free of controversy. Despite—or precisely because of—their supranational orientation and vast use within the context of policy and decision making, they have generated a large number of discussions. In that regard, Hopfenbeck et al. (2018) talk about the massive amount of literature that criticises ILSAs and/or their impact, especially because of the way the results have sometimes been interpreted and used to introduce educational reforms (e.g., Ercikan, Roth, & Asil, 2015; Leung, 2014).

Also, an argument can be raised that such studies greatly oversimplify how educational processes are represented and depicted, showing only a narrow cut-out of the education landscape. Here, the criticism especially focuses on the lack of theoretical foundation in some of the measures being used, how student achievement is measured only for a restricted number of areas and how both these components do not allow for generalising the statements about education processes as a whole (Feuer, 2013; Hopmann, Brinek, & Retzl, 2007). Although the criticism of the quality of the measures subsides in some cases, because of their continuous improvements (e.g., TIMSS, TALIS constructs), a critique that statistical analyses allow for examining only the relationship between the input and output variables is still quite strong (e.g., the commentary chapter Sahlström, this volume). However, if one looks at the indicators typically used to measure educational effectiveness in individual countries, one quickly realises that these indicators are often not comparable across countries and that large-scale studies provide comparable indicators for certain subareas of educational processes, thus providing an important tool for diagnosing the performance and effectiveness of the educational system in individual countries. Although further discussion on these topics can be found elsewhere (e.g., Hopmann et al., 2007; Nasser-Abu Alhija, 2007; Torney-Purta & Amadeo, 2013), it is important to state that this volume also has had no ambition in claiming that large-scale assessments are the only way to provide the sources of knowledge of educational systems and its effectiveness. Rather, it promotes the idea of large-scale assessments as vital points of departure for research on these topics, here coupled with those assessments covering subject didactics in reading, science and mathematics. The choice of editors of this volume and some of its contributing authors further substantiate this argument. Furthermore, the way the idea of educational justice has been examined in this volume showcases that even in ILSAs, multiple lenses and viewpoints can be used to depict a particular phenomenon. Thus, this volume should be read as one source of contemporary knowledge about equity, equality and diversity in Nordic educational systems. In this final chapter, we provide a brief conclusion and summary of the findings presented throughout the book. Furthermore, a perspective for further research is developed following the work in this volume. The text that follows underlines implications on equity, equality and diversity, as well as the Nordic model, here against the background of the empirical work presented, along with the possible pathways for upcoming investigations.

16.1   Important Implications on Equity, Equality and Diversity

The Nordic educational systems are widely considered as rather successful in providing equal access to education and learning opportunities for all, and the Nordic countries tend to rank high on comparative measures on equity in education (Blossing et al., 2014; OECD, 2018). At the same time, socio-economic status (SES) and other background factors still influence the academic achievement of students in the Nordic countries. Furthermore, increasing diversities and social inequalities, globalisation and other changing conditions question the extent to which the Nordic countries are able to maintain a 'School for All' (Lundahl, 2016; Telhaug et al., 2006). In this book, it was possible to identify factors at different levels that influence the achievement of educational equity against this background.

16.2   Teachers and Instructional Quality Play a Key Role in Promoting Equity

Across several chapters, the important role teachers play in promoting equity, equality and safeguarding the diversity of the classrooms has been put to the fore. The different empirical analyses and findings illustrate that any of these three aspects are greatly influenced by what takes place at the junction of teacher professional skills and their instructional practices.

In Chap. 10, Bergem, Nilsen, Mittal and Ræder (2020) found that high instructional quality is especially important for the intrinsic motivation to learn mathematics for low-SES students in both grades 5 and 9. Because there is a strong association between SES and intrinsic motivation in mathematics, instructional quality plays a key role in compensating for this association. Further, in Chap. 9, Nortvedt et al. (2020) argued that teachers' assessment literacy is vital for assessments to function as a means to improve equity in school. This 'assessment literacy' is related to both teacher beliefs and their knowledge and skills. To be able to use assessments and tests as a tool to support teaching and learning, teachers need a positive attitude towards these. Also, teachers need knowledge about the assessments and the assessment data, along with how assessments can be used to support the learning process of students. For the case of the mapping tests discussed in Chap. 9, it was argued that the appropriate use of these tests could help identify students at risk of falling behind and support these students in succeeding.

There are concerns that the focus on attribution and group membership in research means that differences and diversity are associated with problems rather than potential resources and possibilities. As an example, diversity related to language barriers and ethnic and cultural differences need to be taken into account when planning and conducting classroom activities (Robak, Sievers, & Hauenschild, 2013). In Chap. 4, Björnsson (2020) found that across the Nordic countries teachers

general self-efficacy in teaching has the most significant impact on the teachers' experienced ability to handle a multicultural setting. It might not be surprising that teachers who are confident in their general teacher competence also are confident in handling multicultural settings. Yet exposure to multicultural classes seemingly leads to more positive teacher attitudes towards diverse ethnic groups, which again is associated with a better class climate and learning environment. What is perhaps more surprising is that Björnsson (this volume) also found that more experienced teachers are less confident in handling multicultural classrooms; this suggests that for experienced teachers, professional development and in-service training related to diversity and multicultural classrooms may not have successfully targeted their actual needs and that further developments are very much needed.

Finally, in Chap. 5, Yang Hansen, Radišić, Liu and Glassow (2020) found that teacher job satisfaction is connected to certain facets of teacher quality. Although teacher self-efficacy and teacher-students relation were once the common denominators across the Nordic countries, country-specific patterns (e.g., adverse classroom composition in Sweden or teacher effective professional development in Finland) together with teacher-students relation are more decisive to teachers' job satisfaction nowadays. Taken together, both chapters strengthen the perspective that ensuring teacher quality in the Nordic schools as a tool that promotes equity, equality and diversity requires measures and teacher support programmes to be adapted to local needs, not implemented as generic models that are easy to introduce elsewhere.

16.3 The Importance of Teacher Education and Professional Development

It is first and foremost through teacher education and professional development that teachers are provided with the knowledge and skills they need to meet the high demands of a 'School for All', providing equal opportunities to all their students (Imsen & Volckmar, 2014). We highlight two chapters whose findings support this claim.

The importance of professional development is demonstrated in Chap. 7, where Nilsen, Scherer, Gustafsson, Teig and Kaarstein (2020) found that teachers' professional development enhances equity in Sweden by moderating the relation between student SES and science achievement. Both the content and number of hours of teachers' professional development seemed to reduce the performance gap between high- and low-SES students. Furthermore, Nilsen et al. (2020) suggested that the difference in quality and length of professional development in Norway and Sweden could explain why the same results are not found in Norway; they argued that enhancing teachers' qualifications through teacher professional development and specialisation may reduce the effect of students' home background on their achievement, thus enhancing equity.

Advocating for the use of national assessments in promoting equity, in Chap. 9, Nortvedt et al. (2020) argued that teachers' assessment literacy plays a crucial role and that many teachers and schools require support in developing this aspect of their competence and how knowledge of these could be further implemented in classroom practices, as well as school improvement action plans within the schools.

16.4 Ensuring Equity in Digitalised School Settings

With the ongoing digitalisation of society, ICT plays an increasingly important role in schools and classrooms. The integration of ICT in the teaching and learning process introduces another possible source of inequity and inequality in terms of teaching and learning opportunities. In Chap. 6, Rohatgi, Bundsgaard and Hatlevik (2020) found that in Denmark and Norway, there is a lack of variation between schools in teachers' access, use and attitude towards ICT. This indicates institutional or structural equality and a step towards achieving digital equality, thus reducing the overall differences between schools and giving students access to these resources, regardless of where they go to school. At the same time, the authors stressed the importance that the same can be achieved within schools. Therefore, its crucial not only that teachers have access to the relevant resources, but also that they are able to familiarise themselves with the use of ICT.

16.5 When Reading Is Moved Online

Large-scale assessments such as PIRLS and PISA show that there is a relationship between SES and reading achievement for 10-year-olds and 15-year-olds (Mullis, Martin, Foy, & Hooper, 2017; OECD, 2019); however, in many of the Nordic countries, this relationship is weaker than for most other countries (Chaps. 12 and 14). Frønes, Rasmusson and Bremholm (2020) found that the effect of home background is similar since the first round of PISA in 2000, similar to conclusions on Norwegian trend development across reading, science and maths (Olsen & Björnsson, 2018).

What seems to be more problematic from an equity perspective is the large gap in reading achievement between girls and boys and between majority and minority students in the Nordic countries (OECD, 2019). Several PISA cycles have found that girls outperform boys on reading tests and that the gender difference is especially large in Norway and Finland. In her study of Norwegian ninth-grade students, Engdal Jensen (2020; Chap. 13) found that the gender gap increased when reading on a screen compared with reading on paper. This shows that the shift from paper to digital assessments could further influence the reading performance of different students taking the test. At the same time, Frønes et al. (this volume) found that when reading multiple texts—a text format incorporated in the more recent PISA cycles— the gender difference was reduced after accounting for SES. This shows that the

digital reading genres pose even more reading challenges for groups of students but not necessarily because of higher performances for all.

With this challenging digital transition in mind, it is key to gain as much knowledge as possible of how parental reading habits seem to be of significance for students' reading achievement. Adding to the discussion on equity, Støle, Wagner and Schwippert (2020; Chap. 14) found that parents could play an important role in their children's reading development beyond the contribution of SES. Even after controlling for the number of books at home and parents' level of education, Støle et al. (2020) found that parents' reading enjoyment contributes significantly to children's reading achievement across all four PIRLS cycles in all Nordic countries. Considering the gender gap in reading achievement in the Nordic countries, these findings could shed light on particular intervention programmes aiming to reduce this gap. At the same time, this finding corroborates the importance of parental practices and a wider home learning environment in developing children's literacy (Sénéchal & LeFevre, 2002; Skwarchuk, Sowinski, & LeFevre, 2014) while observing the Nordic setting more closely.

16.6   Encompassing Equity—A Wicked Scientific Problem?

Throughout the chapters, equity has been addressed in several ways and with different perspectives and approaches. As described by Buchholtz, Stuart and Frønes (2020) in Chap. 2, different understandings, definitions and measurement instruments could lead to different conclusions related to educational equity and equality, especially when it comes to educational decision making. Buchholtz et al. (2020) further argued that the increasing diversification of the educational landscape in the Nordic countries increases the individual need for compensatory measures, and policy makers are today faced with the ever more difficult challenge of finding fair distributions in the provision of educational resources under these new conditions (e.g., transnationalisation, multicultural ascriptions) that are not at the expense of specific groups. Finding the right balance between the demands of equality (sameness in treatment) and the demands of equity (educational justice) does not only mean guaranteeing formal rights (e.g., in minority language education) but instead taking the requirements of marginalised groups seriously. Educational systems need to ensure that all individuals have the capability to realise their rights and have the material resources to do so.

Another field of knowledge that needs to be further developed relates to the complex mechanisms behind the factors related to equity and equality. In Chap. 3, Mittal, Nilsen and Björnsson (2020) showed that different operationalisations of socio-economic status (SES) lead to different rankings among the Nordic countries in terms of the importance of SES on achievement. The idea is further explored in the work of Scherer (2020), Chap. 8, who found that the disciplinary climate in the classroom may compensate for educational inequalities because of SES across the Nordic countries, but that the disciplinary climate could not mediate or moderate

these inequalities. This implies that the researchers' theoretical perspectives on equity and equality will mainly determine the evaluation of the specific mechanism. At the same time, the diversity that researchers often investigate further contributes to this complexity. This is illustrated in the work of Radišić and Pettersen (2020) in Chap. 11, who found distinctive student profiles that might be more prone to risk after accounting not merely for students' SES, but also individual strengths and disadvantages in the classroom and school setting. In the context of the equality–inequality paradigm, recognition of these potentially at risk profiles strengthens the possibility of reducing the gap in battling the different aspects of inequality across social groups.

16.7 Possible Conclusions About the Nordic Model

This book has not been a quest for evidence of the Nordic model, but rather an empirically based suggestion of the status of the Nordic model and descriptive snapshots of how it is enacted in practice. Seen from the outside, a number of results from large-scale assessments may seem to have testified to a large Nordic tie—with strong similarities between the countries. From PIRLS and PISA, we know that there is a relationship between SES and achievement although in many of the Nordic countries, this relationship is weaker than for most other countries (OECD, 2019). However, there is reason to repeat that there has always been a large variation between the countries' performance and characteristics. One approach to see across this natural variation is to compare trends over time—do countries vary at the same 'pace', or is there any variation? Several chapters have depicted results concerning the trend development, and these can also be seen as determining the status of the Nordic model over time.

In Chap. 12, Frønes et al. (2020) compared subgroup reading literacy trends over time in Denmark, Sweden and Norway, finding that there is no trace of a Nordic unification in the period 2000–2018. The trend is rather the opposite. In Chap. 14, Støle et al. (2020) found that the relationship between parents' reading habits and attitudes on children's reading achievement are fairly stable across countries and time, suggesting that societal family structures are not more or less compensated for over the period 2001–2016.

When it comes to studies that can be said to take the temperature of the contemporary educational systems in the late 2010s, several chapters have indicated that there are similarities that might be attributed more to the system level than to shared cultural communalities. On the school and teacher levels, there are some unifying patterns. In Chap. 4, Björnsson (2020) found that the attitude towards a multicultural classroom is quite similar among the teachers in the Nordic countries. In Chap. 5, Yang Hansen et al. (2020) found evidence in support of student–teacher relations or self-efficacy regarding the importance of teacher job satisfaction for all Nordic countries. In Chap. 6, Rohatgi et al. (2020) discussed how the lack of variation

between schools in teachers' access, use and attitude towards ICT is an indicator of digital equality at the institutional level.

At the same time, there is work pointing in another direction, doubting whether we can still talk about a joint Nordic model regarding policies, school choice and school competition (Klette, 2018) and that the differences in these areas—starting in the early 2000s—threaten the Nordic model (Lundahl, 2016). Lundahl (2016, p. 9) concluded that it is highly doubtful whether one can still speak of a Nordic model of education when considering the development in Sweden from the perspective of extensive marketisation and privatisation practices. Consistent with this, some chapters in this volume also have pointed to differences between the countries, arguing against a strengthening of the Nordic model. In Chap. 2, the country policy review by Buchholtz et al. (2020) regarding minority language students highlighted that the Swedish school system has a distinct different political approach than the other countries (based on Blossing & Söderström, 2014; Gustafson & Yang Hansen, 2017). In Chap. 3, Mittal et al. (2020) noted that Sweden has a different profile than the other countries. In Chap. 5, Yang Hansen et al. (2020) found similar results in the data from the latest TALIS cycle. Finally, in Chap. 12, Frønes et al. (2020) also found support for a clustering trend of reading literacy development in Norway and Denmark on one side and Sweden on the other. While providing evidence both in favour and against the existence of the common Nordic model may seem contradictory and, to an extent, only 'mudding the waters', such results can also indicate that the Nordic countries are more equipped to maintain the common trends in some areas compared with others. Although common cultural attitudes and beliefs are generally relatively stable over time, educational policy and economic relationships can change rapidly, for example, in relation to performance data. In addition, such results may also indicate that nowadays, maintaining the principles embedded in the Nordic model is grounded on different demands than before.

16.8   Future Prospects for Equity, Equality and Diversity

This volume has shown that even if ILSAs offer a rich pool of data that can be used to examine the aspects of equity, equality and diversity across educational systems and time, at the same time, these assessments offer opportunities to observe particular nuances, allowing us to detect students' differences (and similarities) beyond macro-categories. Such fine-grain analyses (e.g., Chaps. 8 and 11) can foster a deeper understanding of the particular relationships within the equity–equality paradigm, allowing us to rediscover some old patterns in a new light.

Examining both the advantages and limitations of that ILSA data may offer, several impending research strands can be identified. Although some focus on improvements in the measures and theoretical foundations of these studies, others concern enrichment in the use of data from large-scale assessment studies and how

these can be used in combination with other data sets and studies to investigate research questions that cannot be resolved with data from a single study only.

Maybe, we state the obvious when we stress that as in all other studies, there is a need for construct development in the ILSAs, especially for more fine-grained measures. Because most large-scale assessments are developed for international and cross-country comparisons, an objection could be made that the measures tend to be too coarse-grained and decontextualised to capture the differences between groups of students when there is relatively small variations in the student population. Stressing again that the relatively egalitarian Nordic societies pose challenges for measuring SES and with the ongoing developments and digitalisation of everyday life, there is an acute need for research to improve the measure to e.g. replace measures as number of books at homes. To some extent, this means integrating measurement differences that stem from cultural, linguistic or other differences in a research design where interesting local differences are appreciated and pursued (Rutkowsi & Rutkowski, 2018). This development on research design, methods, assessment instruments and constructs may lead to more fine-grained measures that provide locally relevant information and results that point in clear directions.

Our last point centres on the relevance of ILSA research. There is a need for the research to be situated in a distinct theoretical field—not only ILSA reporting—to be relevant for both researchers and practitioners and to address research questions noted as important in the theoretical field. In the same way, all educational research should strive to integrate didactic considerations in their own empirical analyses as a guiding principle when interpreting the results and to ensure that the studies are relevant for practitioners and policy makers.

Another main recommendation from this volume is related to the use of data from ILSAs. In Chap. 2, Buchholtz et al. calls for more creative use of ILSA data in research, as a unique source of information encompassing not only student achievement and attitudes, but also including teachers, schools and parents. Further more upcoming trend in using ILSAs data through the lens of person-centred approach, like in Chap. 11, contributes diversity in the use of data within large assessment studies. Moreover, there lies great potential in combining data from different studies more systematically. Strietholt and Scherer (2018) argued that unlike most other datasets in educational research, ILSA data may be combined across studies, cycles and grade levels in numerous ways. These data can also be combined with data from other national and international sources. Combining data from different ILSA projects, as well as combining ILSA data with official statistics and register data, could allow for powerful approaches to investigate research questions that cannot be addressed with the data from a single study (Strietholt & Scherer, 2018). The strength of such a combined approach lies in the possibilities for the cross-validation of findings or a contextual specification and consolidation of research results. This includes, for example, when the results of international comparisons are replicated and reviewed on the national level or when the corresponding findings on the national level are analysed for different temporal cohorts (e.g., different cohorts of students), thus investigating trends. In this volume, we see examples of combining datasets in several chapters. In Chap. 5, Yang Hansen et al. (2020) investigated the

connection between different aspects of teacher quality and job satisfaction using data from two consecutive TALIS cycles: 2013 and 2018. Their findings suggest that the similarities between the Nordic countries found in the 2013 data did not hold true in the 2018 datasets, showcasing diverged country-specific patterns and the possible dissolving of the Nordic model. Furthermore, in Chap. 12, Frønes et al. investigated the reading achievements of students in Denmark, Norway and Sweden over time. To generate trend findings for different groups of students, they used the PISA surveys in reading from 2000, 2009 and 2018. Overall, the trend findings gave the impression that equity related to language background has not improved in any of the three countries.

Even if clear advances can be seen in the combination of different studies and the acquisition of complementary research findings, this is also associated with both specific theoretical and methodological challenges. On the one hand, different large-scale studies use different theoretical frameworks, some of which have been developed and specified over many years (e.g., Stacey & Turner, 2014). As a result, data from studies with different study designs, such as TIMSS and PISA, should only be compared with caution, even if both studies measure mathematics achievements (Wu, 2009). Not only are students of different age groups the target, but also conceptual differences exist: although the data from TIMSS are more curriculum based, PISA tends to test the application of mathematics for solving real-life problems. Another challenge is that empirical data are only comparable to a limited extent over time because statistical indices are collected at different times using different variables (see Chap. 3 and Broer, Bai, & Fonseca. 2019). In addition, there is also the problem of linking existing data from studies and administrative data (e.g. general statistics) with each other because specific methodological and ethical challenges arise, for example, in dealing with personal data, missing values or incorrect linking (including data preparation and deterministic and probabilistic linkage methods; see Harron et al., 2017). However, to reap the benefits of merging data in future research, educational science can learn from health sciences (Bradley, Penberthy, Devers, & Holden, 2010), which are already more advanced in this process, although there are specific differences due to the dynamics of education, for example in the usability of data over a longer period. The potential of ILSAs seems far from fulfilled. Still, the growth of these assessments has been possible due to major advances in technology, measurement theory and statistical modelling the last few decades. Hopefully, additional advances in large-scale assessments, and the methodologies (and theories) on which these are based, in combination with other methods and approaches, can help us provide further insights into the complex nature of concepts such as equity, equality and diversity and the Nordic model.

16.9   Joint Ventures for 'The Hardest Science of All'—A Reply to the Commentary by Sahlström

The author of the commentary Chap. 15 belongs to an epistemological paradigm that is different from the one applied by most of the authors in this volume. Although the authors of this book are largely familiar with the work of ILSAs and apply methods within the quantitative empirical domain (although many of the authors are originally qualitative researchers), as editors, we have deliberately invited a commentator as a critical friend who comes from the field of qualitative research in education and who holds expertise in the area of research on the quality of teaching. This means that the author of the commentary is more interested in observing current events in the teaching–learning process than in less tangible statistical correlations of inputs and outputs. In terms of methodology, qualitative research on teaching tends to rely on small case studies to understand the inner workings of the educational system. Accordingly, the basic message of the commentary is the criticism that descriptive input–output models, which are primarily chosen to identify causal relationships and that are supported by the data ILSA provide, can only provide a reduced view of the findings on the Nordic model of education and that the choice of methods in this volume (and in this field of research) must be critically assessed.

As members of the same criticised field of research, we editors were challenged by the commentary because the idea that the research approach chosen by many of the authors represented in the book—with its corresponding sophisticated scientific methods of analysis—would not be sufficient was in fact only explicitly developed in this form in the commentary chapter. Nevertheless, we have taken the commentary as an occasion for methodological reflection and, therefore, would like to put its inclusion in the book into perspective with some further considerations that we outline below.

In his commentary, Sahlström advocated for methodological approaches to research that examine the inner system of the Nordic model of education based on a deeper 'understanding', especially when it comes to equity and equality. He also cited examples of factors that are less well considered in the book and that can be regarded as influencing factors in the creation of equality of opportunity, such as student activities in class or using digital tools for increased student participation. To a certain extent, Sahlström contrasted the research represented in the book with research stemming from a different empirical standpoint: this and similar work is usually research based on classroom observations and that is mostly qualitative in nature. The results from these observations are undoubtedly an important source of knowledge in the investigation of educational justice in the classroom and, to a greater extent, in the Nordic countries. Thus, we agree with the commentary and must admit that this perspective is almost not represented in the book.

At the same time, one must remember that quantitative and qualitative approaches are opposed to each other, and the strengths and weaknesses of both approaches are well-known (Johnson & Onwuegbuzie, 2004). Precisely because of this, they also

complement each other. Accordingly, Sahlström's criticism of the lack of an answer to the question of a Nordic model must also be classified in this respect.

However, the argument on the 'opposing natures' can also lead us to the insight that the two research perspectives have different expectations of what a Nordic model might be. Quantitative empirical researchers would define a 'model' of education as shared patterns in the data when comparing countries. This is less a weakness in contextual interpretation, as how Sahlström bemoaned, and is instead a more epistemological stance. As Gustafsson (2008, p. 15) pointed out, ILSAs on student achievement have a somewhat deceptive appearance; although they involve students in tasks similar to their everyday schoolwork, the primary purpose is not to provide knowledge about everyday classroom activities but to make generalised descriptions of achievement outcomes at the school system level. So when it comes to identifying a pattern or a 'model', it depends on the interpretation of how big the differences or similarities between countries are as to either call it a Nordic model or not. Therefore, the main focus here is more on describing the available data and identifying empirically validated findings that can serve as a starting point for an interpretation in terms of a common Nordic model. Correspondingly, most chapters in this volume question whether a unifying Nordic model exists.

A qualitative educational researcher will most likely have other criteria to identify some differences and similarities as a 'model' (e.g., shared beliefs and measures based on policy analyses). Both perspectives have advantages and weaknesses (Johnson & Onwuegbuzie, 2004), but they might come to different opinions whether a unifying model exists. Therefore, criticising a research strand to approach the problem from their respective epistemological base has its justification but needs to be interpreted.

However, what the commentary does not develop sufficiently from this starting point is a perspective in which both research approaches can benefit from each other because both are justified in terms of researching equity and equality. Especially when it comes to the complexity of the problems in education science described by Berliner (2002), quantitative approaches would profit from a complementary deepening through qualitative studies that can understand and classify the findings. Thus, researchers may start with quantitative analyses to generate research questions for qualitative studies. Conversely, qualitative studies depend on scientifically substantiating empirical findings by investigating the found phenomena based on large samples. This often involves the formulation of and compliance with scientific quality criteria. The idea behind mixed methods studies is that the strengths and weaknesses of the respective research approaches or elements can be combined or compensated for. Since the 1990s, the mixed methods movement has increasingly set the goal of overcoming trench warfare between purist representatives of the qualitative and quantitative paradigms (Johnson & Onwuegbuzie, 2004), positioning itself as a unifying 'third paradigm' between the two.

Based on the findings of this book, we explicitly advocate for a combination of quantitative and qualitative studies to do scientific justice to the complexity of the question of educational justice and deepen the findings in further research. Here, we also see potential for further empirical work with ILSA (see also Van

Hemert, 2011; Torney-Purta & Amadeo, 2013). The potential for merging insights from ILSA studies with experimental or confirmatory mixed methods studies is far from fulfilled, and several chapters of this book have pointed out that both student outcomes and background variables need to be contextualised through new explanatory studies.

16.10 Concluding Remark

As all chapters in this volume can be defined as secondary data analyses that, according to Hopfenbeck et al. (2018, p. 347), use data as a 'foundation from which to build additional levels of newly constructed knowledge'. The fact that the results from such analyses have the potential as points of departure for other studies is a somewhat obvious and unconventional claim. Strietholt and Scherer (2018) pointed out that cross-country analyses of ILSA data have great potential to generate knowledge about issues related to educational policy at the institutional level, as well as about phenomena at lower levels, such as the school, classroom and home levels. However, we rarely see this done systematically and in complex studies considering contextual factors systematically. As we have already claimed this to be a wicked scientific problem, this book is to be considered a small contribution to this important puzzle.

References

Antikainen, A. (2006). In search of the Nordic model in education. *Scandinavian Journal of Educational Research, 50*(3), 229–243.

Bergem, O. K., Nilsen, T., Mittal, O., & Ræder, H. G. (2020). Can teachers' instruction increase low-SES students' motivation to learn mathematics? In T. S. Frønes, A. Pettersen, J. Radišić, & N. Buchholtz (Eds.), *Equity, equality and diversity in the Nordic model of education* (pp. 252–272). Cham, Switzerland: Springer.

Berliner, D. (2002). Comment: Education: The hardest science of all. *Educational Researcher, 31*(18), 18–20. https://doi.org/10.3102/0013189X031008018

Björnsson, J. K. (2020). Teaching culturally diverse student groups in the Nordic countries—What can the TALIS 2018 data tell us? In T. S. Frønes, A. Pettersen, J. Radišić, & N. Buchholtz (Eds.), *Equity, equality and diversity in the Nordic model of education* (pp. 75–97). Cham, Switzerland: Springer.

Blossing, U., Imsen, G., & Moos, L. (2014). Nordic schools in a time of change. In U. Blossing, G. Imsen, & L. Moos (Eds.), *The Nordic education model. A 'school for all' encounters neoliberal policy* (pp. 1–14). Dordrecht, The Netherlands: Springer.

Blossing, U., & Söderström, M. (2014). A school for every child in Sweden. In U. Blossing, G. Imsen, & L. Moos (Eds.), *The Nordic education model. A 'school for all' encounters neoliberal policy* (pp. 17–34). Dordrecht, The Netherlands: Springer.

Bradley, C. J., Penberthy, L., Devers, K. J., & Holden, D. J. (2010). Health services research and data linkages: Issues, methods, and directions for the future. *Health Services Research, 45*(5 Pt 2), 1468–1488.

Broer, M., Bai, Y., & Fonseca, F. (2019). *Socioeconomic inequality and educational outcomes. Evidence from twenty years of TIMSS*. Cham, Switzerland: Springer.

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* (pp. 13–41). Cham, Switzerland: Springer.

Engdal Jensen, R. (2020). Implications of changing the delivery mode on reading tests in Norway—A gender perspective. In T. S. Frønes, A. Pettersen, J. Radišić, & N. Buchholtz (Eds.), *Equity, equality and diversity in the Nordic model of education* (pp. 337–362). Cham, Switzerland: Springer.

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

Feuer, M. (2013). Validity issues in international large-scale assessment: Truth and consequences. In M. Chatterji (Ed.), *Validity and test use: An international dialogue on educational assessment, accountability and equity* (pp. 197–215). Bingley, UK: Emerald.

Frønes, T. S., Rasmusson, M., & Bremholm, J. (2020). Equity and diversity in reading comprehension—A case study of PISA 2000–2018. In T. S. Frønes, A. Pettersen, J. Radišić, & N. Buchholtz (Eds.), *Equity, equality and diversity in the Nordic model of education* (pp. 305–335). Cham, Switzerland: Springer.

Gustafsson, J.-E. (2008). Effects of international comparative studies on educational quality on the quality of educational research. *European Educational Research Journal, 7*(1), 1–17. https://doi.org/10.2304/eerj.2008.7.1.1

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., & Yang Hansen, K. (2017). Changes in the impact of family education on student educational achievement in Sweden 1988–2014. *Scandinavian Journal of Educational Research, 62*(5), 719–736.

Harron, K., Dibben, C., Boyd, J., Hjern, A., Azimaee, M., Barreto, M. L., et al. (2017). Challenges in administrative data linkage for research. *Big Data & Society, 4*(2), 1–12.

Hopfenbeck, T. N., Lenkeit, J., El Masri, Y., Cantrell, K., Ryan, J., & Baird, J.-A. (2018). Lessons learned from PISA: A systematic review of peer-reviewed articles on the Programme for international student assessment. *Scandinavian Journal of Educational Research, 62*(3), 333–353. https://doi.org/10.1080/00313831.2016.1258726

Hopmann, S. T., Brinek, G., & Retzl, M. (2007). *PISA according to PISA: Does PISA keep what it promises?* Wien, Austria: Lit Verlag.

Imsen, G., & Volckmar, N. (2014). The Norwegian school for all: Historical emergence and neoliberal confrontation. In U. Blossing, G. Imsen, & L. Moos (Eds.), *The Nordic education model. A 'school for all' encounters neo-liberal policy* (pp. 35–56). New York: Springer.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher, 33*(7), 14–26. https://doi.org/10.3102/0013189X033007014

Klette, K. (2018). Individualism and collectivism in Nordic schools. A comparative approach. In N. Witoszek & A. Midttun (Eds.), *Sustainable modernity: The Nordic model and beyond* (pp. 59–79). London: Routledge.

Leung, F. K. S. (2014). What can and should we learn from international studies of mathematics achievement? *Mathematics Education Research Journal, 26*, 579–605. https://doi.org/10.1007/s13394-013-0109-0

Lundahl, L. (2016). Equality, inclusion and marketisation of Nordic education: Introductory notes. *Research in Comparative and International Education, 11*(1), 3–12.

Mittal, O., Nilsen, T., & Björnsson, J. K. (2020). Measuring equity across the Nordic education systems—Conceptual and methodological choices as implications for educational policies. In

T. S. Frønes, A. Pettersen, J. Radišić, & N. Buchholtz (Eds.), *Equity, equality and diversity in the Nordic model of education* (pp. 43–71). Cham, Switzerland: Springer.

Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2017). *PIRLS 2016 international results in reading*. http://timssandpirls.bc.edu/pirls2016/international-results/

Nasser-Abu Alhija, F. (2007). Large-scale testing: Benefits and pitfalls. *Studies in Educational Evaluation, 33*(1), 50–68.

Nilsen, T., Scherer, R., Gustafsson, J.-E., Teig, N., & Kaarstein, H. (2020). Teachers' role in enhancing equity—A multilevel structural equation modeling with mediated moderation. In T. S. Frønes, A. Pettersen, J. Radišić, & N. Buchholtz (Eds.), *Equity, equality and diversity in the Nordic model of education* (pp. 173–196). Cham, Switzerland: Springer.

Nortvedt, G. A., Bratting, K. B., Kovpanets, O., Pettersen, A., & Rohatgi, A. (2020). Improving equity through national-level assessment initiatives. In T. S. Frønes, A. Pettersen, J. Radišić, & N. Buchholtz (Eds.), *Equity, equality and diversity in the Nordic model of education* (pp. 225–248). Cham, Switzerland: Springer.

OECD. (2018). *Equity in education. Breaking down barriers to social mobility*. Paris: OECD Publishing. https://doi.org/10.1787/9789264073234-en

OECD. (2019). *PISA 2018 results (volume II): Where all students can succeed*. Paris: OECD Publishing. https://doi.org/10.1787/b5fd1b8f-en

Olsen, R. V., & Björnsson, J. K. (2018). Tjue år med internasjonale skoleundersøkelser i Norge: Bakgrunn, læringspunkter og veien videre [Twenty years of international large-scale assessments in Norway: Background, learning points and future prospects]. In J. K. Björnsson & R. V. Olsen (Eds.), *Tjue år med TIMSS og PISA i Norge: Trender og nye analyser* [Twenty years with TIMSS and PISA in Norway. Trends and new findings] (pp. 12–34). Universitetsforlaget.

Radišić, J., & Pettersen, A. (2020). Resilient and nonresilient students in Sweden and Norway—Investigating the interplay between their self-beliefs and the school environment. In T. S. Frønes, A. Pettersen, J. Radišić, & N. Buchholtz (Eds.), *Equity, equality and diversity in the Nordic model of education* (pp. 273–304). Cham, Switzerland: Springer.

Robak, S., Sievers, I., & Hauenschild, K. (2013). Einleitung diversity education: Zugänge und Spannungsfelder [Introduction diversity education: Approaches and areas of tension]. In K. Hausenschild, I. Sievers, & S. Robak (Eds.), *Diversity education. Zugänge—Perspektiven—Beispiele* [Diversity education. Approaches—Perspectives—Examples] (pp. 15–35). Frankfurt a.M, Germany: Brandes & Apsel.

Rohatgi, A., Bundsgaard, J., & Hatlevik, O. E. (2020). Digital inclusion in Norwegian and Danish schools—Analysing variation in teachers' collaboration, attitudes, ICT use and students' ICT literacy. In T. S. Frønes, A. Pettersen, J. Radišić, & N. Buchholtz (Eds.), *Equity, equality and diversity in the Nordic model of education* (pp. 139–172). Cham, Switzerland: Springer.

Rutkowski, L., & Rutkowski, D. (2018). Improving the comparability and local usefulness of international assessments: A look back and a way forward. *Scandinavian Journal of Educational Research, 62*(3), 354–367. https://doi.org/10.1080/00313831.2016.1261044

Rutkowski, L., von Davier, M., & Rutkowski, D. (2014). *Handbook of international large-scale assessment. Background, technical issues, and methods of data analysis*. Boca Raton: Chapman & Hall/CRC Press.

Sahlström, F. (2020). The black box of Nordic education held against the light of large-scale international assessment resources—A critical commentary. In T. S. Frønes, A. Pettersen, J. Radišić, & N. Buchholtz (Eds.), *Equity, equality and diversity in the Nordic model of education* (pp. 387–396). Cham, Switzerland: Springer.

Scherer, R. (2020). The case for good discipline? Evidence on the interplay between disciplinary climate, socioeconomic status, and science achievement from PISA 2015. In T. S. Frønes, A. Pettersen, J. Radišić, & N. Buchholtz (Eds.), *Equity, equality and diversity in the Nordic model of education* (pp. 197–224). Cham, Switzerland: Springer.

Sénéchal, M., & LeFevre, J.-A. (2002). Parental involvement in the development of children's reading skill: A five-year longitudinal study. *Child Development, 73*(2), 445–460. https://doi.org/10.1111/1467-8624.00417

Skwarchuk, S.-L., Sowinski, C., & LeFevre, J.-A. (2014). Formal and informal home learning activities in relation to children's early numeracy and literacy skills: The development of a home numeracy model. *Journal of Experimental Child Psychology, 121*, 63–84. https://doi.org/10.1016/j.jecp.2013.11.006

Stacey, K., & Turner, R. (2014). The evolution and key concepts of the PISA mathematics frameworks. In K. Stacey & R. Turner (Eds.), *Assessing mathematical literacy. The PISA experience* (pp. 5–33). Cham, Switzerland: Springer.

Støle, H., Wagner, A. K. H., & Schwippert, K. (2020). The importance of parents' own reading for 10-year-old students' reading achievement in the Nordic countries. In T. S. Frønes, A. Pettersen, J. Radišić, & N. Buchholtz (Eds.), *Equity, equality and diversity in the Nordic model of education* (pp. 363–384). Cham, Switzerland: Springer.

Strietholt, R., & Scherer, R. (2018). The contribution of international large-scale assessments to educational research: Combining individual and institutional data sources. *Scandinavian Journal of Educational Research, 62*(3), 368–385. https://doi.org/10.1080/00313831.2016.1258729

Telhaug, A. O., Mediås, O. A., & Aasen, P. (2006). The Nordic model of education: Education as part of the political system in the last 50 years. *Scandinavian Journal of Educational Research, 50*(3), 245–283.

Torney-Purta, J., & Amadeo, J.-A. (2013). International large-scale assessments: Challenges in reporting and potentials for secondary analysis. *Research in Comparative and International Education, 8*(3), 248–258.

Van Hemert, D. A. (2011). Frameworks for explaining cross-cultural variance: A meta-analytic examination of their usefulness. In F. J. R. van de Vijver, A. Chasiotis, & S. M. Breugelmans (Eds.), *Fundamental questions in cross-cultural psychology* (pp. 312–340). Cambridge, UK: Cambridge University Press.

Wu, M. (2009). A comparison of PISA and TIMSS 2003 achievement results in mathematics. *Prospects, 39*, 33–46.

Yang Hansen, K., Radišić, J., Liu, X., & Glassow, L. N. (2020). Exploring diversity in the relationships between teacher quality and job satisfaction in the Nordic countries—Insights from TALIS 2013 and 2018. In T. S. Frønes, A. Pettersen, J. Radišić, & N. Buchholtz (Eds.), *Equity, equality and diversity in the Nordic model of education* (pp. 99–137). Cham, Switzerland: Springer.