

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 & Middtun, 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 systems¹ 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.

¹We 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 meritocracy² 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 ways³ (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

²In 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).

³See 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 equity⁴ 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

⁴Horizontal 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 judgments, 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 heterogeneous

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 outcomes⁵ 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 Slegers (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 Slegers (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),

⁵Students' 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 approach⁶ applied in TIMSS, with students nested in classrooms and classrooms nested in schools, results in substantial intraclass correlation (ICC) within groups,

⁶For 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.

⁷See <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 score⁸ 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 \geq .95, RMSEA \leq 0.08, SRMR \leq 0.06. For the acceptable threshold, we used: CFI \geq .90, RMSEA \leq 0.10, SRMR \leq 0.10, Δ CFI \leq –0.01, and Δ RMSEA \leq 0.05

In RMSEA difference test, Δ RMSEA must be \leq 0.05 when testing for metric invariance and \leq 0.01 when testing for scalar invariance

⁸Mathematics 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, *R²* percentage of variance in mathematics achievement explained

^aThe 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

		One-level regression				Two-level regression					
		Country	Country ranking ^a	Coef.	R ²	Coef.	R ²	Coef.	R ²		
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

^aThe 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								
Country	Country ranking	Mathematics			Country ranking	Science		
		Mean (S.E.)	SD ^a	S.E.		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)

^aA 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											
Country	#	Mathematics				Gap ^a	#	Science			Gap ^a
		Low SES	Medium SES	High SES	Low SES			Medium SES	High SES		
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

^aAchievement gap between the low-SES and high-SES groups

⁹For 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 heterogeneous 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 & Slegers, 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 multi-level 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 & Slegers, 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 & Slegers, 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 neo-liberal 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/10.1080/00131831.2016.1258726>
- 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/10.1787/79789264073234-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/10.1787/74789264300002-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., & Middtun, 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.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

