Social Indicators Research (2023) 170:399–425
https://doi.org/10.1007/s11205-023-03196-5

ORIGINAL RESEARCH

Mind the Gap… But Which Gap? The Distinctions Between Social Inequalities in Student Achievement

Andrés Strello[1,2] · **Rolf Strietholt**[1,2] · **Isa Steinmann**[3,4]

Accepted: 7 August 2023 / Published online: 9 September 2023

Abstract
International large-scale assessments have revealed social inequalities in achievement in almost all countries, reporting achievement gaps between socioeconomic status groups, by immigration background and by gender. However, there has been little research on whether individual countries show smaller or larger gaps across all three different social categories, or whether the gaps corresponding to these categories are independent of each other. This article explores the degree to which social inequality can be understood as one umbrella concept, or whether different categories of social inequality are substantially different concepts. Using the OECD's Programme for International Student Assessment 2018 results in Mathematics in 76 countries, the study observes the correlation between the three achievement gaps across countries, and compares how each achievement gap is associated with some typical country-level covariates. Several results are highlighted. First, the size and direction of the immigration and gender gaps vary across countries; most countries present achievement gaps in favor of boys and native students, but this direction is reversed in several countries. Second, there is hardly any correlation between the three achievement gaps. One education system may be egalitarian in one category, but profoundly unequal in another. Third, this lack of correlation is also related to how we study these inequalities, as the results show that each achievement gap is associated with a different set of institutional features. To properly assess how unequal or egalitarian education systems are, researchers and interested parties need to consider and address different indicators of social inequality.

Keywords Equity · Performance · Education · Measurement · Inequality of opportunity · Comparative analysis

✉ Andrés Strello
andres.strello@tu-dortmund.de

[1] International Association for the Evaluation of Educational Achievement (IEA), Hamburg, Germany

[2] Center for Research on Education and School Development (IFS), TU Dortmund University, Dortmund, Germany

[3] Department of Primary and Secondary Teacher Education, Oslo Metropolitan University, Oslo, Norway

[4] Centre for Educational Measurement (CEMO), University of Oslo, Oslo, Norway

1 Introduction

International large-scale assessments (ILSAs) such as PISA or TIMSS reveal considerable variation in both the mean performance levels and the extent of social inequality that exists within participating countries. Regarding social inequality, the most commonly studied social categories are socioeconomic status (SES), immigration status, and student gender (e.g., Andon et al., 2014; Rosén et al., 2022; Jerrim et al., 2019). The answers to our research questions could have multiple implications for educational monitoring, as well as for research on social inequality in student achievement. If achievement gaps between different social categories are highly correlated, then examining them separately adds little value for educational monitoring, and their reporting should be reframed. In this scenario, it also seems plausible that research findings on the institutional determinants of social inequality would be consistent across different social categories. But if different social gaps are largely uncorrelated, there is a need for a differentiated perspective in education policy and research.

This study empirically examines whether there is a single broad social inequality, or whether there is a need to distinguish between different forms of social inequalities corresponding to the categories of SES, immigration background, and gender. Are there countries which systematically show social inequality in student performance across different categories, or are countries characterized by a higher degree of inequality in one social category and lower in another? To address this question, we first use data from an international large-scale assessment to compute social inequalities in SES, immigration status, and gender. We use this data to review the variability in social inequality for each social category across countries, before evaluating the correlation between the three measures at country-level. To further validate these analyses, we conduct a comparative study and investigate the association between various institutional features and the three forms of social inequalities. Specifically, are institutional features of school systems consistently associated with all forms of social inequality or only to specific ones? This study is explorative and aims to contribute to the discussion on how researchers can evaluate education systems.

2 Social Inequalities in Achievement on International Analyses

2.1 The Concept of Social Inequality

Inequality in education can be conceptualized using different terms, with different normative ideas about injustice and the evaluation of education systems (Strietholt, 2014; Brighouse & Swift, 2008). The concept of social inequality—similar to 'inequity' and 'inequality of opportunities' (c.f. UNESCO, 2018)—problematizes achievement differences that originate from the social origin of the student, rather than from effort or ability. In educational research, the most commonly used social categories are SES, immigration background, and gender. While these are the three categories studied in this article, we acknowledge that there are other characteristics determined by social origin that are also related to inequalities within education, such as religion, sexual orientation, ethnicity, and place of residence (e.g., urbanicity).

Understanding the categories of social inequality presents a dilemma. Different categories of social inequalities have some common aspects. First, since there are great

differentials in the outcomes and trajectories of students, we can expect that there are groups that are able to take more advantage than others in a systematic way. This is especially true in contexts of high general inequality, i.e., high dispersion in outcomes, where differentials between students are bigger and there is more variance that can be unevenly distributed. Second, we selected these three categories of social inequality because they are present in most education systems in the world and have been an ongoing topic in educational research for decades. The global relevance of these three categories enable us to hypothesize that there is one broad 'umbrella' social inequality, in which these social distinctions (SES, immigration, gender) are associated with the distribution of outcomes simultaneously, i.e., highly associated between the three of them. However, different categories of social inequality within an education system emerge for different reasons. The association between each social category with performance outcomes could run on parallel paths, implying null correlations between them.

Researchers have been able to measure the different categories of social inequality in achievement on an international perspective since the mid-twentieth century. Currently, the three largest ongoing international large-scale assessments measuring achievement in school students are the OECD's Program for International Student Assessment (PISA), the International Association for the Evaluation of Educational Achievement's (IEA) Trends in International Mathematics and Science Study (TIMSS), and the Progress in International Reading Literacy Study (PIRLS). Below, we present a short review of the current international evidence and the theories explaining each category of social inequality. We will explore prominent theories regarding the emergence of achievement gaps related to SES, gender, and immigration. Our aim is to demonstrate that the underlying mechanisms behind these gaps are fundamentally distinct from one another. Subsequently, we will examine previous research that investigates the correlation between institutional features of educational systems and the three distinct types of inequality. In addition, it is worth noting that previous research in this area has been somewhat fragmented. There has been a lack of systematic evaluations where the same data were utilized to study the relationships between institutional features and the various forms of inequality. Moreover, the few studies that have attempted this approach have yielded inconsistent findings.

2.2 SES Inequality in Achievement

The association between family SES and student achievement can be explained by the tendency of children from families with a lower socioeconomic background to receive fewer resources for their education. This difference in resources accumulates along the children's developmental trajectory and generates disparities in achievement between children from different families. This is further exacerbated by the inheritability of resources between generations that increases the resource gap between families. According to Bourdieu's theory, these resources are manifested first as economic resources (e.g., families with higher incomes can send their students to private schools or afford private tutoring) and later manifest in cultural and social capital (Bourdieu, 1986; Broer et al., 2019; Coleman, 1988, 1990).

The association between a student's SES background and performance has been a common finding across studies, cycles, and subjects, though with differences between countries in the association's magnitude (Hopfenbeck et al., 2018). PISA 2018 (OECD, 2019a) shows a positive association between SES and achievement in all countries and in all three subjects (reading, mathematics, and science), with SES explaining between 2 and 24% of

the variance in performance, depending on the country and subject. TIMSS (Mullis et al., 2020) and PIRLS (Mullis et al., 2017) presented similar patterns in their latest editions in 2019 and 2016, respectively. While these are recent results, SES inequality in performance is not new and has even increased in some countries (Broer et al., 2019; Chmielewski, 2019).

2.3 Immigration Inequality in Achievement

The association between immigration background and student achievement can be explained by two groups of mechanisms: structural and cultural (Nauck, 2019). The structural mechanism is the inherent disadvantage experienced by immigrant groups due to their economic reality. Families with immigrant backgrounds show lower academic performance or take different educational choices due to their poorer access to resources (both economic and social). Cultural explanations ultimately focus on why certain groups of immigrants or ethnic groups perform better than others; the disadvantage is explained in terms of different mindsets. Studies have shown that the gap between immigrant and native students is not only due to immigrant families' lower SES, but also due to speaking a different language at home, sociocultural factors, system-wide factors of the origin and destination countries (such as political stability, economic development, and religion), and the destination countries' policies (Buchmann & Parrado, 2006; Dronkers & Levels, 2007; Jackson, 2012; Levels et al., 2008; Schmid, 2001; Strand, 2011, 2014).

Most international research in English on the association between immigration and achievement has focused on Western European countries and the USA. In European countries, students who speak a different language at home perform worse in PISA, especially at reading (Lenkeit et al., 2017). Moreover, most studies using ILSAs data have each only investigated a limited set of countries, focusing on the differences between immigrant groups within a country, e.g., the disadvantage of Turkish communities within Germany (Söhn & Özcan, 2006), or how immigrants are disadvantaged in the USA depending on their origin country (Worrell, 2014). As most research is centered in European and North American contexts, some other contexts are excluded. In Qatar and United Arab Emirates, immigrant children perform better than natives, supposedly because these countries attract high-skilled immigrants and their education systems are tailored to this (Bouhlila, 2017). Overall, the achievement gap varies greatly across the assessed countries, contents, and cycles (Andon et al., 2014).

2.4 Gender Inequality in Achievement

There are different and longstanding theories on why gender gaps in student achievement tests occur, and they can be divided into two broad explanations: nature and nurture (see overviews by Halpern, 2012; Hyde, 2014). The nature category includes theories that assume innate, stable differences between boys and girls that affect learning processes. The comprehensive literature on cognitive gender gaps suggests, however, that boys and girls mostly score equally on cognitive ability tests (cf. Gender Similarity Hypothesis; Hyde, 2014; Zell et al., 2015). In contrast, the nurture category includes theories about environmental influences differing between boys and girls. Nurture-related theoretical perspectives all suggest that societal gender norms and existing gender differences in education transmit to students, perpetuating educational gender inequalities. For instance, stereotypical beliefs about science, technology, engineering, and mathematics (STEM) subjects being

male domains and a higher representation of men in STEM majors at school and university level or in the STEM labor market can lead to girls underestimating their abilities in these subjects, potentially impacting their achievement (Eccles et al., 1990; Halpern, 2012; Neuville & Croizet, 2007).

International comparative studies document pronounced gender gap differences between countries and academic achievement domains. Girls outperform boys in reading in most countries at both the primary and secondary school level. Gender gaps in the participating countries range between non-existent reading gender gaps to large advantages for girls (Mullis et al., 2017; OECD, 2019a). Gender gaps are more varied in mathematics, with medium advantages for boys in some countries, some countries without gender gaps, and even some countries with medium advantages for girls (Mullis et al., 2020; OECD, 2019a). Interestingly, gender gaps in reading and mathematics appear to correlate; countries with pronounced reading advantages for girls also tend to show mathematics advantages for girls, and countries without reading advantages for girls tend to show mathematics advantages for boys (Guiso et al., 2008; Stoet & Geary, 2013). Furthermore, gender gaps in academic achievement appear to be quite stable over time (Rosén et al., 2022; Steinmann, et al., 2023; Meinck & Brese, 2019).

2.5 Covariates of Social Inequalities

Within each category of social inequality, the associations between social origin with performance vary between countries. This suggests that institutional features of education systems generate variations in social inequality (Jerrim et al., 2019). We next review some studies that have identified institutional features related to social inequality in achievement. We explore whether previous studies suggest that institutional covariates are associated in the same way with different forms of social inequality.

2.5.1 Education-System Factors

One important feature of education systems is the level of differentiation, seen in policies such as between-school tracking, in which students are sorted into different types of schools. If transitions and school choice are affected by social characteristics, either by the achievement differential between social groups or by different decisions taken after considering children's skills, differentiation in the education system should lead to larger social achievement gaps. Previous international studies have found that educational differentiation (specifically between-school tracking) increases SES inequality in achievement (Strello et al., 2021; Lavrijsen & Nicaise, 2016; van de Werfhorst, 2018; van de Werfhorst & Mijs, 2010). There is less research on the effect of tracking on immigration inequality in achievement and the findings are inconsistent; some studies suggest a positive effect while others do not (Bodovski & Munoz, 2020; Ruhose & Schwerdt, 2016; Teltemann & Schunck, 2016). Between-school tracking has mixed effects on gender inequality, with studies consistently showing that later tracking increases the gender gap in reading (in favor of girls), but heterogeneous results regarding the effect on mathematics and science (Bodovski & Munoz, 2020; Hermann & Kopasz, 2019; Scheeren & Bol, 2022). Similar results have been found in studies on general education-system differentiation indexes (Ayalon & Livneh, 2013; van Hek et al., 2019; van Langen et al., 2006).

2.5.2 External Factors

A common factor in comparative research is the level of economic development of a country or education system. In general, previous studies have found mixed evidence on its effect on social inequality. Measures such as GDP (gross domestic product) per capita are inconsistently associated with SES achievement inequality (Bodovski & Munoz, 2020; Chmielewski, 2019; Ferreira & Gignoux, 2014; Schütz et al., 2008). Previous studies have also found mixed results regarding the association between SES inequality and public expenditure on education, although the association seems more markedly negative when considering countries' development levels (Strietholt et al., 2019). Chmielewsky (2019) found that income inequality (measured as Gini) has a positive association with SES inequality in mid and low-income countries. Using TIMSS 2011 data, Bodovski and Munoz (2020) found an inverse association between GDP per capita and the immigrant achievement gap (in particular, richer countries have a lower gap between immigrants and native students), but found no association with the gender achievement gap.

Cultural features may also play a role in gender inequalities in achievement. In more gender-egalitarian countries, the relative performance of girls over boys is higher, especially in reading (see review of Rosén et al., 2022; González de San Román & de La Rica, 2016; Guiso et al., 2008; Marks, 2008; Reilly, 2012). Nosek et al. (2009) found that in societies with more marked stereotypes (e.g., regarding science as a male domain and liberal arts as a female domain), the gap in favor of boys is larger in mathematics.

3 The Present Study

Previous research on social inequality has identified a number of social categories related to student achievement, with the most prominent categories in international comparative research being SES, immigration, and gender. The theories explaining the emergence of each performance gap differ, and research on the three areas has developed relatively independently.

Only a few studies have explicitly compared the different gaps. Lenkeit et al. (2017) studied the relative importance of the three categories of achievement gaps, though only in four Western European countries (Germany, Sweden, France, and United Kingdom). They estimated multilevel models using data from PISA 2000 to 2012. The authors concluded that each category of social inequality is important for explaining the disparities between students, and that results have remained stable in those four countries. Bodovski et al. (2020) studied how different country-level predictors may mitigate the three categories of social inequality. The authors used information from TIMSS 2011 with a sample of 45 countries. They found mixed effects between the different social inequality domains, showing that the role of school system features cannot be generalized over the different categories of social inequality. Whether large gaps in one social category are associated with large gaps in another social category has not been the subject of research to date.

In this study, we explore the relationship between the three categories of social inequality on achievement, and the degree to which they are correlated or uncorrelated. We investigate whether countries can be evaluated as more or less socially unequal based on one only category, or how important it is to evaluate the effects of certain policies on different categories of inequality. Specifically, we aim to answer the following research question:

How correlated are the three categories of social inequality in achievement (socioeconomic status, immigration background, and gender)?

A high correlation between the different types of social inequality would suggest that the differentiation between the three types of social inequality has no additional empirical value, whereas low correlations would underpin the importance of a differentiated view of social inequalities.

Furthermore, we proceed to examine the nomological linkages between the three distinct types of social inequality and external variables. This line of inquiry aligns with the principles of construct validity (Cronbach & Meehl, 1955). By assessing how different types of inequality correlate with relevant variables, we can gather evidence supporting their meaningful distinction. More specifically, our investigation focuses on the relationship between education system-level features and social inequality across the various categories of social inequality.

Different patterns in the regression parameters would provide evidence that the three types of inequality need to be differentiated when analyzing social inequality. If there are no differences in the regression estimates, however, differentiating between the types of social inequality would not provide additional empirical value.

4 Methods

4.1 Data Sources

To study the correlation between the different categories of social inequalities, we use the OECD's Programme for International Student Assessment (PISA). This study measures 15-year-olds' proficiency in mathematics, reading, and science. Specifically, we use the dataset of PISA 2018 focusing on the mathematics assessment. We remove Korea and Vietnam from the sample as they sampled fewer than 20 students with immigrant backgrounds (as defined in the Variables chapter). The remaining sample of n = 76 education systems[1] is heterogeneous and covers all parts of the world. Each country contains a sample between 3296 to 35,493 students (mean: 7791), with the number of schools ranging between 44 and 1089 schools (mean: 279). The total sample contains 592,145 students from 21,264 schools. Table 1 shows the total N of students and schools per country.

PISA draws a stratified two-stage sampling. The first stage samples schools within the country or education system, and the second stage samples 15-year-old students within those schools. The results are representative of the population at both the student-level and the school-level. However, PISA sample only students enrolled within schools, meaning that interpretations of these results must consider that some specific countries/regions have lower proportions of secondary-school attainment and therefore exclude early school leavers (Steinmann & Rutkowski, 2023).

[1] Hereafter, we refer to the education systems sampled as "countries", even if some participants did not sample the full country, namely, China and Azerbaijan (Baku).

Table 1 Proportion of social groups and subsamples N, N of students, N of schools by country

Country	University education		Immigrant background		Girl		Total N	
	%	N	%	N	%	N	Students	Schools
Albania	20	1310	1	40	49	3167	6359	327
Argentina	34	4450	5	654	51	6232	11,975	455
Australia	53	6722	28	3433	49	7075	14,273	763
Austria	30	2019	23	1413	49	3321	6802	291
B-S-J-Z (China)	21	3790	0	21	48	5775	12,058	361
Baku (AZ)	23	1518	5	340	47	3262	6827	197
Belarus	43	2505	4	228	48	2772	5803	234
Belgium	49	4025	18	1508	50	4271	8475	288
Bosnia and Herzegovina	22	1429	3	182	49	3148	6480	213
Brazil	32	3260	1	60	50	5478	10,691	597
Brunei Darussalam	35	2379	8	555	50	3383	6828	55
Bulgaria	49	2563	1	70	47	2533	5294	197
Canada	60	12,708	35	5667	50	11,307	22,653	821
Chile	34	3200	4	258	49	3814	7621	254
Chinese Taipei	26	1855	1	52	50	3624	7243	192
Colombia	24	1831	1	43	51	3857	7522	247
Costa Rica	43	3043	10	725	51	3618	7221	205
Croatia	36	2370	9	601	50	3311	6609	183
Czech Republic	31	2501	4	253	49	3518	7019	333
Denmark	73	5205	11	1578	50	3816	7657	348
Dominican Republic	41	2244	3	155	50	2890	5674	235
Estonia	46	2432	10	543	50	2651	5316	230
Finland	63	3503	6	319	49	2772	5649	214
France	47	2843	14	959	49	3078	6308	252
Georgia	61	3437	1	76	48	2682	5572	321
Germany	38	1712	22	1055	46	2525	5451	223
Greece	47	3017	12	710	49	3178	6403	242
Hong Kong	22	1225	38	2202	49	2955	6037	152
Hungary	48	2516	3	125	50	2605	5132	238
Iceland	69	2210	6	181	50	1656	3296	142
Indonesia	23	3173	0	36	51	6240	12,098	397
Ireland	47	2574	18	983	50	2777	5577	157
Israel	57	3638	17	1021	53	3544	6623	174
Italy	36	4134	10	1080	48	5680	11,785	542
Japan	48	2845	1	39	51	3120	6109	183
Jordan	37	3267	12	1612	51	4619	8963	313
Kazakhstan	33	6667	8	1434	49	9576	19,507	616
Kosovo	40	1983	1	68	50	2457	5058	211
Latvia	45	2301	4	246	51	2685	5303	308
Lebanon	30	1641	6	302	54	3079	5614	313
Lithuania	50	3332	2	150	49	3377	6885	362
Luxembourg	45	2201	55	2828	49	2594	5230	44

Table 1 (continued)

Country	University education		Immigrant background		Girl		Total N	
	%	N	%	N	%	N	Students	Schools
Macao	25	943	63	2371	49	1862	3775	45
Malaysia	29	1771	2	97	51	3131	6111	191
Malta	40	1325	9	288	48	1612	3363	50
Mexico	27	1889	2	88	52	3826	7299	286
Moldova	37	2012	2	80	49	2621	5367	236
Montenegro	42	2811	6	392	48	3240	6666	61
Morocco	21	1365	1	57	48	3262	6814	179
Netherlands	67	3083	14	695	50	2330	4765	156
New Zealand	45	2664	27	1623	50	3154	6173	192
North Macedonia	42	2267	2	80	48	2596	5569	117
Norway	41	2253	12	696	49	2880	5813	251
Panama	36	2210	6	343	50	3173	6270	253
Peru	32	1920	1	31	49	3000	6086	340
Philippines	28	2004	1	74	53	3868	7233	187
Poland	37	2042	1	27	50	2857	5625	240
Portugal	37	2061	7	344	49	2944	5932	276
Qatar	71	9604	57	7476	49	6954	13,828	188
Romania	31	1532	1	30	48	2444	5075	170
Russian Federation	63	4824	6	434	50	3861	7608	263
Saudi Arabia	37	2301	12	698	48	2992	6136	234
Serbia	37	2394	9	612	49	3272	6609	187
Singapore	45	2907	25	1555	49	3277	6676	166
Slovak Republic	40	2409	1	69	50	3002	5965	376
Slovenia	42	2388	9	578	49	2993	6401	345
Spain	46	17,311	12	4170	49	17,956	35,943	1,089
Sweden	57	3000	21	1077	50	2763	5504	223
Switzerland	38	2173	34	1954	47	2789	5822	228
Thailand	26	2880	1	70	53	4693	8633	290
Turkey	23	1569	1	54	50	3396	6890	186
Ukraine	39	2353	2	146	47	2857	5998	250
United Arab Emirates	70	12,720	56	9671	51	9380	19,277	755
United Kingdom	48	5870	20	1786	51	6996	13,818	471
United States	53	2453	23	1011	49	2376	4838	164
Uruguay	19	1013	1	64	52	2732	5263	189

N indicates the N of the subsamples within each category, i.e., N of students with parents with university education, N of students with immigrant background and N of girls. Proportions are weighted

4.2 Analysis

Do all measures of inequality show the same picture, or is it necessary to differentiate between multiple types of social inequality? Are social achievement gaps consistent, or are there countries in which certain social gaps are high and others low? To address

these questions empirically, we examine whether different measures of social inequality in student achievement lead to the same or different rankings in international comparisons. All analyses are based on the three types of social inequality in student achievement available for the n = 76 participants in PISA: SES, immigrant background, and gender.

Our analysis consists of three steps. First, we identify the three gaps per country (see Variables section below). Second, we examine the correlation of these different types of social inequality at the country level. Third, we attempt to validate the correlational analyses by regression analyses. We regress the three types of social inequalities on a set of institutional features and compare the regression parameters for the three outcomes. We use cross-sectional data, and the aim of the regression analyses is not to estimate causal effects or bring substantive conclusions, but rather to examine whether different social inequalities are associated differently with various system-level features.

4.3 Variables

4.3.1 *Social Achievement Gap on Mathematics*

The main variables of interest are measures of three categories of social inequalities in achievement. Achievement scores in PISA are calculated so that they had an international mean of 500 and an international standard deviation of 100 points in the first edition in 2000. The scores are designed to be comparable between countries. In our analysis, we focus on mathematics achievement.

The three achievement gaps were calculated for gender, SES, and immigration status, using the simple mean difference between the groups (described in the next section). We divided these gaps by the standard deviation of the mathematics scores observed in the respective country, to account for cross-country variation in the dispersion of the test scores. Therefore, all gaps are measured as Cohen's standardized effect sizes *d*. For example, a gender gap of 1 means that boys perform on average one standard deviation better than girls. Probability weights were used in the estimation of the achievement gaps and standard error account for the sample design using replications weights. We followed the Balanced Repeated Replication (BRR) method, as indicated by PISA guidelines (OECD, 2019b). All ten plausible values available in the PISA public database were used on the analyses following Rubin's rules (Rubin, 1987). The three achievement gaps, by country, can be found in Appendix (Table 3).

4.3.1.1 SES Achievement Gap We used parental education as a measure of socioeconomic status and compared students with parents with university education (ISCED 5A) against parents with less than university education. If the educational attainment of the parents differed, we used the highest educational attainment reached between both parents—i.e., one parent having university level education is enough to be considered in the highest category. We marked parental education as missing if there was no information about both parents. We opted for a single proxy of SES instead of a complex index, such as the ESCS reported in PISA, for the sake of simplicity and consistency with the other categories that also use a single indicator. Table 1 shows the proportion of parents with university education. There is a high heterogeneity between countries on socioeconomic levels. This proportion ranges from 7% in Vietnam to 73% in Denmark.

4.3.1.2 Immigration Achievement Gap We operationalize the immigration background by comparing students whose parents were both born abroad with students with one or no parent born abroad. We categorize the first group as "immigrant" and the second as "native", aware that this is a simplified category of a more complex phenomenon. We marked this variable as missing if there was no information about both parents. Immigration background has a high heterogeneity between countries, ranging from slightly over 0% in several countries to as high as 63% in Macao (see Table 1). Six countries (China, Korea, Peru, Poland, Romania, and Vietnam) have fewer than 30 cases with immigrant backgrounds. The efficiency of the estimation of achievement gaps based on immigration is reduced. However, excluding these cases does not affect the results of this study (see Results below).

We calculate the "raw" association between immigration and achievement scores. An alternative would be to estimate the achievement gap, controlling first for student SES. However, we want to highlight how immigration has different connotations between countries, as shown in Fig. 1. In addition, the association between immigration background and SES tends to be small (even non-significant) in several countries, and in different directions, as shown in Appendix (Table 3). While many countries (e.g., Western European countries) show a positive association between being a native student and having parents with university education, in many others (e.g., South American and Middle-East countries) the correlation is negative. Moreover, the between-countries correlation of the correlation of University-Native with Native-immigrant achievement gap is only $r = 0.21$ ($p < 0.1$). Therefore, we consider it appropriate to study the immigrant achievement gap fully detached from its interaction with the student SES.

4.3.1.3 Gender Achievement Gap To measure gender, we use the variable available on the PISA student dataset. The proportion of girls is mostly balanced across countries, ranging from 47 to 53% (see Table 1).

4.3.2 Country-Level Covariates

To validate the empirical differentiability of the three types of inequality, we use some key correlates of student achievement and social inequality in achievement commonly used in previous studies. This section is not intended to bring substantive conclusions, but to complement the previous analyses; if the regression models differ across the different types of social inequality, it brings some evidence on how this discussion have consequences on substantive educational research too. Some information is derived from the PISA school principal questionnaire, which is also representative of each country's school system, while other variables are derived from external sources.

4.3.2.1 GDP per Capita To indicate a country's economic wealth, we used gross domestic product (GDP) per capita. This information is based on the World Bank database (World Bank, 2022), and we used the latest information for each country or region (up to 2018).

4.3.2.2 Growth Mindset To capture cultural differences across countries, we used the variable growth mindset, available on the PISA 2018 student dataset. 'Growth mindset' refers to the belief that someone's ability and intelligence can be developed over time (OECD, 2019c). Within each country, we averaged the percentage of students that

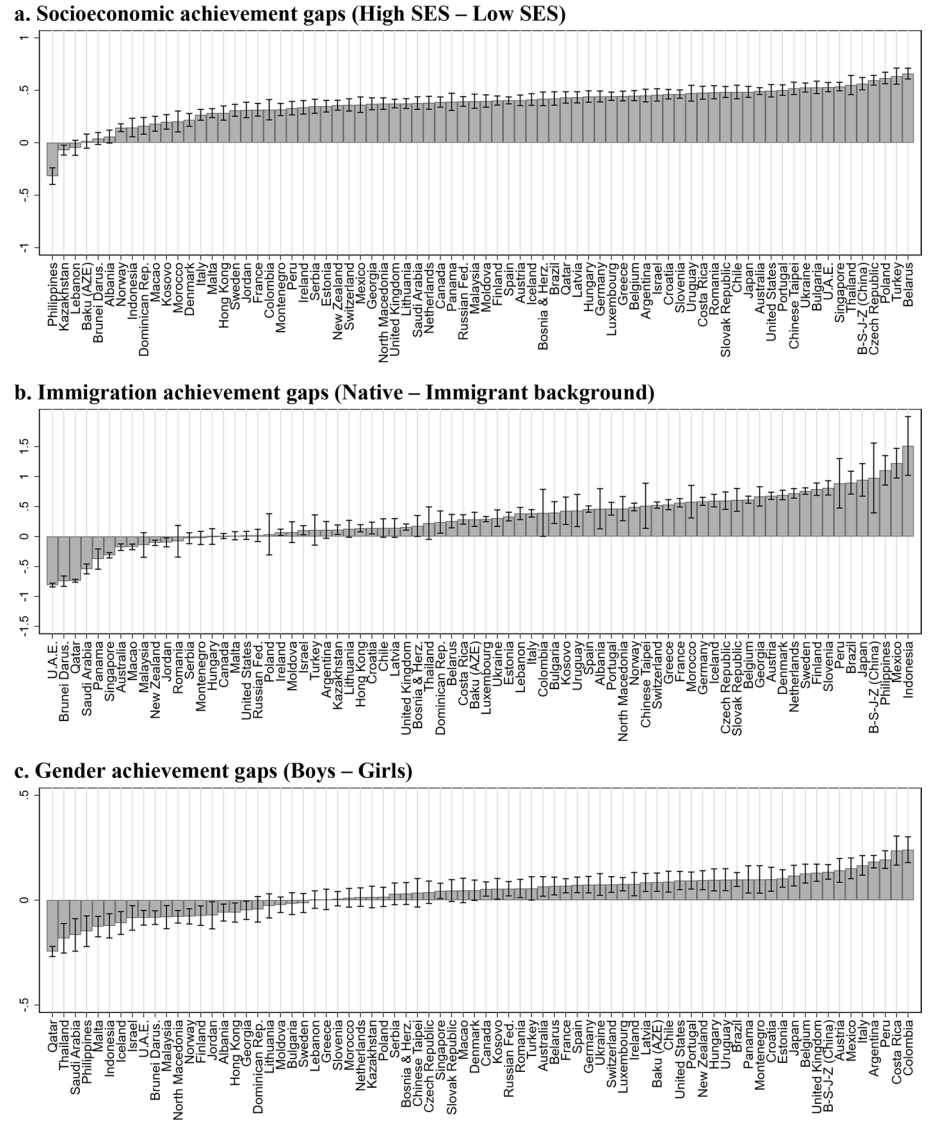


Fig. 1 SES, Immigrant, and Gender achievement gaps. Confidence intervals at 95% confidence level. Y axis are on different scale between each plot. Available as table format in Appendix (Table 3)

strongly disagree or disagree with the statement "Your intelligence is something about you that you can't change very much".

4.3.2.3 Between-school Tracking Age This variable indicates at what age (based on the modal age for the corresponding grade) students are placed into different school tracks. Different tracks typically have different curricula, and the transition from a comprehensive to a tracked school system constitutes an important event in students' educational careers. We followed the information indicated in Strello et al., (2021), complemented

by our own elaboration based on UNESCO-IBE's World Data on Education (UNESCO-IBE, 2012).

4.3.2.4 Selectiveness Besides tracking, we also included two indicators of the degree of selectiveness within educational systems: the importance for school admission of (1) Students' record of academic performance (including placement tests) and (2) Residence in a particular area. For each of these we calculate the percentage per country of school principals that declare they *Always* (vs. *Sometimes* or *Never*) consider these factors in school admissions. A larger percentage of the first item is an indicator of a more selective system, while a larger percentage of the latter item is an indicator of a less selective system.

4.3.2.5 Grade Repetition We included the percentage of students that had repeated a grade in their school course, using information on the PISA 2018 student questionnaire aggregated to the country level.

5 Results

5.1 Social Achievement Gaps Across Countries

As a preliminary step, we describe the social achievement gaps by social category. SES achievement gaps consist of the mean difference between high SES and low SES in standardized mathematics achievement scores. Figure 1a shows that the vast majority of countries present a positive and significant SES achievement gap. The only exceptions are the Philippines and Kazakhstan with a negative gap; and Lebanon, Baku (AZ), Brunei Darussalam, and Albania with a non-significant gap. Among those with significant positive achievement gaps, there is a high variability on the magnitude of these gaps; Norway and Indonesia have an SES achievement gap of 0.14 SD, while Belarus and Vietnam have SES achievement gaps of 0.66 SD and 0.74 SD, respectively.

The immigration achievement gap is calculated as the mean score difference between native students (without an immigrant background) and students with an immigrant background. Immigration achievement gaps are present in most countries on where there is a positive gap (i.e., natives perform better than immigrants), although several countries present a negative gap (i.e., immigrants perform better than natives), and many others where the differences are non-significant (see Fig. 1b). The range of the magnitude of the immigration achievement gaps is also larger than for SES achievement gaps, from a negative gap of -0.8 SD in United Arab Emirates to around 1.50 SD in Indonesia. Moreover, the unbalanced shares of natives and immigrant subsamples imply large confidence intervals in some countries, so the estimations of gaps in this category are less efficient than on SES and gender.

The gender achievement gap is calculated as the mean score difference between boys and girls. A positive gap indicates a higher mean score for boys than girls. One contrast with both previous measures is the smaller range of achievement gaps overall, from a −0.24 SD gap in Qatar to a 0.24 SD gap in Colombia (see Fig. 1c). While most countries present positive gaps (boys achieve better mathematics scores than girls), there are negative gaps in many countries, with girls achieving better mathematics scores than boys.

Before we explore the association between the three types of social inequality with the full sample of 76 countries, we take a closer look at individual countries. In Turkey, the

SES achievement gap is very high, whereas the immigration and gender gaps are small compared to the other countries. In Italy, the gaps for SES, immigration, and gender are low, medium, and high, respectively. Such patterns suggest that social inequality must be understood multidimensionally, since certain types of inequality are typically higher than others within the same country. Accordingly, a single type of social inequality is insufficient to conclude that social inequality in performance is generally low or high in a country.

5.2 Correlation of SES, Immigration, and Gender Achievement Gaps Across Countries

The correlational analyses, including all 76 countries, confirm the need to distinguish between SES, immigration, and gender achievement gaps. The correlation between SES and immigration achievement gaps is $r = -0.13$ (non-significant [n.s.]), between SES and gender gaps it is $r = 0.24$ ($p < 0.05$), and between immigration and gender gaps it is $r = 0.18$ (n.s.). Figure 2 plots the associations between the three performance gaps, with no evidence of non-linear relationships.

In countries with a small number of immigrants, the gap between immigrant and native students is measured with lower reliability (see the large confidence intervals for some countries in Fig. 2b). To address this issue, we restricted the sample to the $n = 53$ countries with more than 100 immigrant students.[2] The results of the correlation analyses with the restricted sample are qualitatively the same as with the total sample. The correlation between SES and native-immigrant achievement gap remains insignificant and the correlation estimate is even lower ($r = 0.03$; n.s.). The correlation between immigrant achievement gap and gender achievement gap is 0.31 ($p < 0.05$). There is a small correlation when removing countries with less precise estimations of the immigrant achievement gap; in this sample there is a small tendency of countries with advantages for native students over immigrants to also show advantages for boys over girls. The correlation between the SES achievement gap and gender achievement gap is lower and is statistically insignificant ($r = 0.19$; n.s.). The analyses in the following section use the full sample of countries.

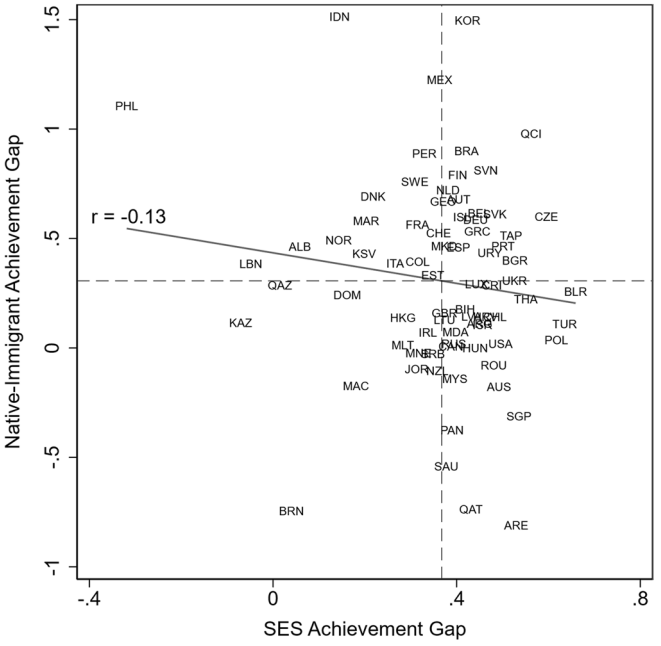
5.3 Nomological Networks

In the previous section, we presented evidence that the three achievement gaps—by SES, immigration background, and gender—are largely uncorrelated. In this section, as validity analyses that complement the previous section, we explore how the different achievement gaps are associated with different country-level features. If institutional features are associated differently with each achievement gap, this would provide further evidence of the need to differentiate between these three types of inequality when analyzing social inequality. We do not aim to bring substantive conclusions, that would require more theoretical development and a more complex analysis design.

Table 2 shows the results from three regression analyses, where we regressed the three measures of social inequalities on the set of country-level institutional features of education systems. For easier interpretation of the results, we present standardized regression

[2] Countries excluded from the restricted sample of $n = 53$ are Poland, Romania, Peru, Indonesia, Japan, Albania, Colombia, Chinese Taipei, Turkey, Morocco, Brazil, Uruguay, Kosovo, Slovak Republic, Bulgaria, Thailand, Philippines, Georgia, Moldova, North Macedonia, Mexico, and Malaysia.

a. Correlation SES achievement gap – Immigrant achievement gap



b. Correlation SES achievement gap – Gender achievement gap

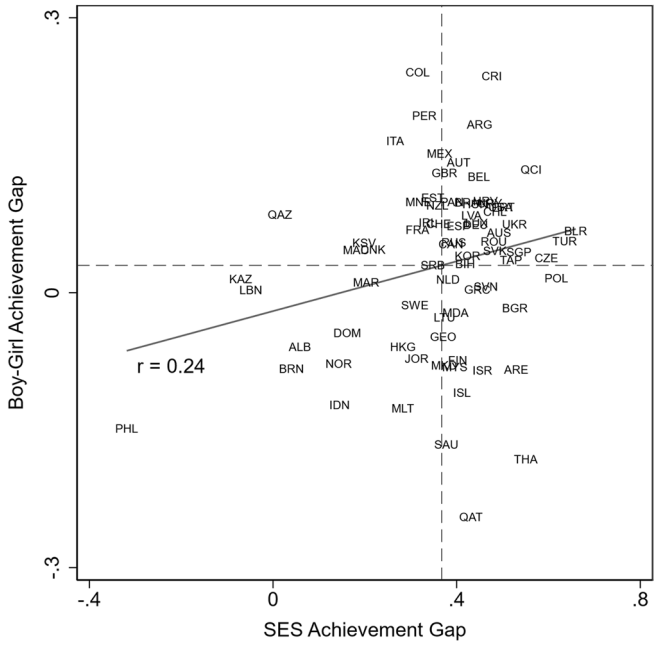


Fig. 2 Correlation between achievement gaps. Horizontal and vertical lines represent the between-countries mean of axis Y and axis X achievement gaps respectively. Solid line represents the correlation using the full sample of countries (N = 76)

c. Correlation Gender achievement gap – Immigrant achievement gap

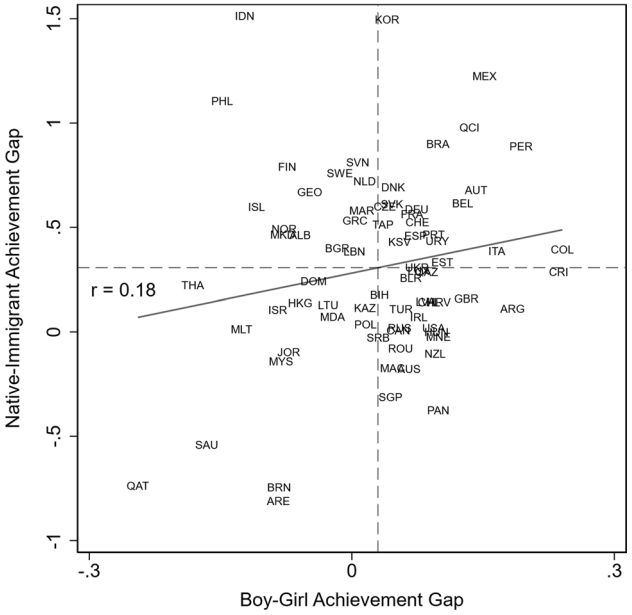


Fig. 2 (continued)

estimates. Models estimated variable-by-variable are available in Appendix, Tables 4, 5, 6. The comparison reveals that institutional characteristics better explain variation in the gender achievement gap (explaining 48% of the international variation) than in the immigrant achievement gap (32%) and the SES achievement gap (22%).

The main finding of the comparison of regression parameters is that institutional characteristics are differentially associated with different social achievement gaps, providing further evidence that a holistic evaluation of social inequality requires a consideration of different gaps. For example, the economic power of countries, measured as GDP per capita, is associated negatively with the immigrant achievement gap and the gender gap, but it is not associated with the SES achievement gap. The growth mindset cultural indicator is associated with the SES achievement gap and the gender achievement gap, but not with the immigration gap.

Regarding selectivity, using residence as a criteria for selection is negatively associated with SES and gender achievement gaps. Counterintuitively, selecting by performance is only associated with a reduction in the immigration gap. A later tracking is significantly negatively associated only with the immigration gap.

Education systems with a higher percentage of repeating students tend to show a lower SES gap and a higher gender gap.

6 Discussion

In this study, we aimed to explore the degree to which there is one *umbrella* concept of social inequality, or whether there are substantially different concepts of social inequalities. We explored the correlations between three different social inequalities in achievement:

Table 2 OLS Models on Achievement Gaps

	(1) High–low SES	(2) Native–Immigrant	(3) Boy–Girl
GDP per Capita	0.037	−0.474***	−0.373***
	(0.008)	(0.019)	(0.004)
Growth mindset	0.241*	0.032	0.563***
	(0.191)	(0.457)	(0.089)
Tracking age	−0.188	−0.290**	−0.159
	(0.012)	(0.029)	(0.006)
Selection by residence	−0.246*	−0.068	−0.274***
	(0.100)	(0.238)	(0.047)
Selection by performance	−0.203	−0.265*	−0.030
	(0.095)	(0.228)	(0.044)
Repeated grade	−0.201*	0.115	0.316***
	(0.168)	(0.402)	(0.079)
N	73	73	73
R^2	0.216	0.303	0.458
Adjusted R^2	0.145	0.240	0.409

Standardized beta coefficients are reported. Standard errors in parentheses. Significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

SES, immigration status, and gender. We also compared how different education system-level covariates are associated with each achievement gap.

We highlight several points. First, at least one category of achievement gaps can be observed in every country. Second, while SES gaps were observed in all but four countries (as well as their direction), the size and direction of the immigration and gender gaps vary across countries. In most countries, natives and boys have better mean performance in mathematics than immigrants and girls, but there are several countries where immigrants and girls have an advantage. The variation across countries in the achievement gap by immigration is clearly higher than in the SES and, especially, the gender achievement gap. Also, the share of immigrants is very low in many countries, making it both empirically difficult to study (due the lower efficiency of the estimations) and a less prominent problem in some regions of the world. These findings suggest that, while SES inequality appears to be an almost global phenomenon, immigration and gender are associated with educational disadvantages differently across different countries and regions. Based on these findings, we conclude that the institutional context and social practices in different countries play a role in shaping social inequality. In the following section, we support this interpretation with additional findings.

Second, there is hardly any correlation between the three achievement gaps. This means that one education system can be egalitarian in some category, but profoundly unequal in another. To properly assess how unequal or egalitarian education systems are, policy-makers, researchers, and other stakeholders need to consider and address different indicators of social inequality.

Third, this lack of correlation is also related to how we study these inequalities. Using the same sample of countries and the same covariates, we showed that each achievement gap is associated with a different set of institutional features. Researchers who aim to study the impact of institutional characteristics on social inequality from a holistic perspective are advised to consider different forms of social inequality. Conclusions from a study on one gap cannot be generalized to other gaps.

6.1 Limitations and Future Research

This article has some limitations. One set of limitations relate to the measures we considered in the present study. We have only considered three key social categories here—SES, immigration, and gender—but there are other important categories (such as religiosity and ethnicity). These categories are often not highlighted in international assessments, and more comprehensive data is required to explore the gaps associated with these categories. Another limitation concerns the indicators used to measure SES, immigration, and gender. For the sake of simplicity, we considered only one indicator per category; nevertheless we recognize that these indicators have limitations, as there are more ways of operationalizing both immigration and SES. Also, we have not explored intersectionality among the three categories. For example, boys with a migration background may be a particularly disadvantaged group. Such analyses are beyond the scope of this paper but appear important for further research.

It is important to mention that these results only refer to mathematics achievement. There are other cases where achievement gaps could be of different magnitude or different direction. For example, looking at the latest international reports of PIRLS (Mullis et al., 2017) and PISA (OECD, 2019a), girls score significantly higher than boys in most countries, while in no country boys score better than girls. We focused on the achievement gap in mathematics as it illustrated the best the differences between achievement gaps.

Another set of limitations relates to the analysis of the institutional covariates, and natural limitations to the samples in some countries. The analysis of the covariates is based on cross-sectional data, and for this reason we do not make causal inferences. However, such analyses of nomological networks provide useful evidence for the distinction of social inequalities. Furthermore, the results involving immigration gaps are particularly affected by small subsamples of immigrants in certain countries where immigration is uncommon, lowering the measurement efficiency. Lastly, while PISA samples hundreds of thousands of students, the number of countries remains a natural limitation in any cross-national research.

6.2 Conclusion

In conclusion, mind the gap, but consider what gap you are looking at, as not all gaps are equal; depending on the social category, the results are very different. Ranking countries in terms of just one social inequality category provides a limited picture, at best. SES inequality is mostly a global problem, but immigration is more relevant in some regions than others, while gender gaps follow opposite direction between countries. This has direct consequences on the evaluation of education systems, and on research.

Appendix

See Tables 3, 4, 5 and 6.

Table 3 Achievement gaps, correlation between social indicators, by country

Country	Achievement gap			Correlation		
	High–low SES	Native–immigrant	Boy–girl	University-native	University-boy	Native-boy
Albania	.06	.46*	−.06*	−.12*	.07*	−.06
Argentina	.45*	.11	.18*	.13*	.09*	.01
Australia	.49*	−.18*	.07*	−.21*	.02*	.00
Austria	.40*	.68*	.14*	.10*	.02	.00
B-S-J-Z (China)	.56*	.98*	.13*	−.04	.01	.02
Baku (Azerbaijan)	.01	.29*	.09*	−.02	.11*	.00
Belarus	.66*	.26*	.07*	.09*	.00	−.03
Belgium	.45*	.61*	.13*	.12*	.04*	−.03*
Bosnia and Herzegovina	.42*	.18*	.03	−.17*	.09*	−.04
Brazil	.42*	.90*	.10*	−.14*	.07*	.08*
Brunei Darussalam	.04	−.75*	−.08*	−.11*	.16*	−.00
Bulgaria	.53*	.40*	−.02	−.08*	.07*	−.12*
Canada	.39*	.01	.05*	−.12*	.04*	−.03*
Chile	.48*	.14	.09*	−.10*	.08*	.08*
Chinese Taipei	.52*	.51*	.04	−.15*	.03	.10*
Colombia	.32*	.39*	.24*	−.03	.11*	.10*
Costa Rica	.48*	.29*	.24*	.19*	.08*	.00
Croatia	.46*	.14*	.10*	.11*	.09*	.01
Czech Republic	.60*	.60*	.04	−.07*	.03*	.02
Denmark	.22*	.69*	.05	.25*	−.02	−.02
Dominican Republic	.16*	.24*	−.04	−.11*	.13*	−.09*
Estonia	.35*	.33*	.10*	−.13*	−.01	−.10*
Finland	.40*	.79*	−.07*	.12*	.00	.01
France	.31*	.56*	.07*	.11*	.00	−.02
Georgia	.37*	.67*	−.05*	.05	.10*	−.08*

Table 3 (continued)

Country	Achievement gap			Correlation		
	High–low SES	Native–immigrant	Boy–girl	University-native	University-boy	Native-boy
Germany	.44*	.59*	.07*	.16*	.01	−.02
Greece	.45*	.53*	.00	.22*	.08*	−.02
Hong Kong	.28*	.14*	−.06*	.26*	−.05*	-.06*
Hungary	.44*	−.00	.10*	−.08*	.09*	.04
Iceland	.41*	.60*	−.11*	.15*	.03	.10*
Indonesia	.14*	1.51*	−.12*	−.39*	.13*	−.13*
Ireland	.34*	.07*	.08*	−.15*	−.01	.04*
Israel	.46*	.11*	−.08*	−.03*	.13*	−.06*
Italy	.27*	.39*	.17*	.02	.07*	.04
Japan	.48*	.95*	.12*	−.05	.03	.01
Jordan	.31*	−.10*	−.07*	−.08*	.15*	−.00
Kazakhstan	−.07*	.11*	.01	.06*	.08*	−.01
Kosovo	.20*	.43*	.05*	−.09*	.07*	−.03
Latvia	.43*	.14	.08*	−.06*	.08*	−.09*
Lebanon	−.05	.38*	.00	−.20*	.16*	.02
Lithuania	.37*	.13	−.03	−.13*	.11*	−.04
Luxembourg	.44*	.29*	.08*	.12*	.05*	−.04*
Macao	.18*	−.17*	.05	.17*	.01	.05*
Malaysia	.40*	−.14	−.08*	.02	.08*	−.01
Malta	.28*	.01	−.13*	−.30*	.09*	.01
Mexico	.36*	1.22*	.15*	.07	.07*	.20*
Moldova	.40*	.07	−.02	−.17*	.08*	−.09*
Montenegro	.32*	−.02	.10*	−.17*	.16*	.00
Morocco	.20*	.58*	.01	−.33*	.12*	−.17*

Table 3 (continued)

Country	Achievement gap			Correlation		
	High–low SES	Native–immigrant	Boy–girl	University-native	University-boy	Native-boy
Netherlands	.38*	.72*	.01	.22*	.04*	−.00
New Zealand	.36*	−.11*	.10*	−.17*	.02	−.07*
North Macedonia	.37*	.47*	−.08*	−.03	.13*	.07*
Norway	.14*	.49*	−.08*	.06*	.05*	.03
Panama	.39*	−.38*	.10*	−.13*	.04*	−.00
Peru	.33*	.89*	.19*	−.18*	.02	.13*
Philippines	−.32*	1.11*	−.15*	−.28*	.18*	−.17*
Poland	.62*	.04	.02	−.16*	−.01	.09
Portugal	.50*	.46*	.09*	−.09*	.05*	−.01
Qatar	.43*	−.74*	−.24*	−.15*	.09*	−.04*
Romania	.48*	−.08	.06*	−.23*	.10*	−.01
Russian Federation	.39*	.02	.05*	−.02	.09*	.04
Saudi Arabia	.38*	−.54*	−.17*	−.07*	.08*	−.08*
Serbia	.35*	−.03	.03	−.04*	.08*	.05*
Singapore	.54*	−.31*	.04*	−.39*	−.02*	.02*
Slovak Republic	.48*	.61*	.05	−.14*	.03	−.07*
Slovenia	.46*	.81*	.01	.31*	.02	−.02
Spain	.40*	.46*	.07*	.13*	.01	.01
Sweden	.31*	.76*	−.01	.11*	.00	−.05*
Switzerland	.36*	.52*	.07*	.09*	.02*	.01
Thailand	.55*	.22	−.18*	.01	.11*	−.05
Turkey	.64*	.11	.06*	−.17*	.07*	−.01
Ukraine	.53*	.31*	.07*	−.11*	.01	−.15*
United Arab Emirates	.53*	−.81*	−.08*	−.32*	.09*	−.01

Table 3 (continued)

Country	Achievement gap			Correlation		
	High–low SES	Native–immigrant	Boy–girl	University-native	University-boy	Native-boy
United Kingdom	.37*	.16*	.13*	−.10*	.05*	.05*
United States	.50*	.02	.09*	.22*	.09*	.00
Uruguay	.47*	.43*	.10*	−.17*	.09*	.02

* = Achievement gap / correlation significant $p < 0.05$. Correlations correspond to tetrachoric correlations between dichotomic indicators

Table 4 OLS models on socioeconomic (High–Low SES) mathematics achievement gaps

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
GDP per Capita	0.037 (0.008)	0.106 (0.008)					
Growth mindset	0.241* (0.191)		0.341*** (0.154)				
Tracking age	−0.188 (0.012)			−0.203* (0.011)			
Selection by residence	−0.246* (0.100)				−0.115 (0.095)		
Selection by performance	−0.203 (0.095)					−0.123 (0.077)	
Repeated grade	−0.201* (0.168)						−0.192 (0.170)
N	73	73	73	73	73	73	73
R^2	0.216	0.011	0.116	0.013	0.015	0.041	0.037
Adjusted R^2	0.145	−0.003	0.104	−0.001	0.001	0.028	0.023

Standardized beta coefficients are reported. Standard errors in parentheses. Significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 5 OLS models on immigration (Native–Immigrant) mathematics achievement gaps

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
GDP per Capita	−0.474*** (0.019)	−0.417*** (0.018)					
Growth mindset	0.032 (0.457)		−0.000 (0.415)				
Tracking age	−0.290** (0.029)			−0.122 (0.029)			
Selection by residence	−0.068 (0.238)				−0.080 (0.241)		
Selection by performance	−0.265* (0.228)					−0.186 (0.193)	
Repeated grade	0.115 (0.402)						0.134 (0.435)
N	73	73	73	73	73	73	73
R^2	0.303	0.174	0.000	0.006	0.035	0.015	0.018
Adjusted R^2	0.240	0.162	−0.014	−0.008	0.021	0.001	0.004

Standardized beta coefficients are reported. Standard errors in parentheses. Significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 6 OLS models on gender (Boys–Girls) mathematics achievement gaps

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
GDP per Capita	−0.373*** (0.004)	−0.190 (0.004)					
Growth mindset	0.563*** (0.089)		0.364*** (0.086)				
Tracking age	−0.159 (0.006)			−0.212* (0.006)			
Selection by residence	−0.274*** (0.047)				−0.246** (0.052)		
Selection by performance	−0.030 (0.044)					−0.136 (0.043)	
Repeated grade	0.316*** (0.079)						0.246** (0.094)
N	73	73	73	73	73	73	73
R^2	0.458	0.036	0.132	0.060	0.018	0.045	0.060
Adjusted R^2	0.409	0.023	0.120	0.047	0.005	0.032	0.047

Standardized beta coefficients are reported. Standard errors in parentheses. Significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Funding Open Access funding enabled and organized by Projekt DEAL. This work was supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 765400.

Data Availability All data is freely available from https://www.oecd.org/pisa/.

Declarations

Conflict of interest The authors have no competing interest to declare that are relevant to the content of this article.

References

Andon, A., Thompson, C. G., & Becker, B. J. (2014). A quantitative synthesis of the immigrant achievement gap across OECD countries. *Large-Scale Assessments in Education, 2*(1), 1–20. https://doi.org/10.1186/s40536-014-0007-2

Ayalon, H., & Livneh, I. (2013). Educational standardization and gender differences in mathematics achievement: A comparative study. *Social Science Research, 42*(2), 432–445. https://doi.org/10.1016/j.ssresearch.2012.10.001

Bodovski, K., & Munoz, I. G. (2020). Do education system characteristics moderate the socioeconomic, gender and immigrant gaps in math and science achievement? *International Journal of Sociology of Education, 9*(2), 122–154.

Bouhlila, D. S. (2017). Parents' education and literacy skills: Evidence on inequality of socioeconomic status in Arab countries. *World Development Perspectives, 5*, 34–43. https://doi.org/10.1016/j.wdp.2017.02.006

Bourdieu, P. (1986). The forms of capital. In J. Richardson (Ed.), *Handbook of Theory and Research for the Sociology of Education* (pp. 241–258). Greenwood.

Brighouse, H., & Swift, A. (2008). Putting educational equality in its place. *Education Finance and Policy, 3*(4), 444–466. https://doi.org/10.1162/edfp.2008.3.4.444

Broer, M., Bai, Y., & Fonseca, F. (2019). *Socioeconomic inequality and educational outcomes: Evidence from 20 years of TIMSS*. Cham: Springer International Publishing.

Buchmann, C., & Parrado, E. A. (2006). Educational achievement of immigrant-origin and native students: A comparative analysis informed by institutional theory. *The impact of comparative education research on institutional theory* (pp. 335–366). Emerald (MCB UP). https://doi.org/10.1016/S1479-3679(06)07014-9

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

Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology, 94*, S95–S120.

Coleman, J. S. (1990). *Foundations of social theory*. The Belknap of Harvard University Press.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*(4), 281–302. https://doi.org/10.1037/h0040957.hdl:11299/184279

Dronkers, J., & Levels, M. (2007). Do school segregation and school resources explain region-of-origin differences in the mathematics achievement of immigrant students? *Educational Research and Evaluation, 13*(5), 435–462. https://doi.org/10.1080/13803610701743047

Eccles, J. S., Jacobs, J. E., & Harold, R. D. (1990). Gender role stereotypes, expectancy effects, and parents' socialization of gender differences. *Journal of Social Issues, 46*(2), 183–201. https://doi.org/10.1111/j.1540-4560.1990.tb01929.x

Ferreira, F. H. G., & Gignoux, J. (2014). The Measurement of educational inequality: Achievement and opportunity. *World Bank Economic Review, 28*(2), 210–246. https://doi.org/10.1093/wber/lht004

González de San Román, A., & de La Rica, S. (2016). Gender gaps in PISA test scores: The impact of social norms and the mother's transmission of role attitudes. *Estudios de Economía Aplicada*. http://www.redalyc.org/articulo.oa?id=30143731005

Guiso, L., Monte, F., & Sapienza, P. (2008). Differences in test scores correlated with indicators of gender equality. *Science, 320*(May), 1–2.

Halpern, D. F. (2012). *Sex differences in cognitive abilities* (4th ed.). Psychology Press.

Hermann, Z., & Kopasz, M. (2019). Educational policies and the gender gap in test scores: A cross-country analysis. *Research Papers in Education, 00*(00), 1–22. https://doi.org/10.1080/02671522.2019.1678065

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

Hyde, J. S. (2014). Gender similarities and differences. *Annual Review of Psychology, 65*(1), 373–398. https://doi.org/10.1146/annurev-psych-010213-115057

Jackson, M. (2012). Bold choices: How ethnic inequalities in educational attainment are suppressed. *Oxford Review of Education, 38*(2), 189–208. https://doi.org/10.1080/03054985.2012.676249

Jerrim, J., Volante, L., Klinger, D. A., & Schnepf, S. V. (2019). Socioeconomic inequality and student outcomes across education systems. In L. Volante, S. V. Schnepf, J. Jerrim, & D. A. Klinger (Eds.), *Socioeconomic inequality and student outcomes: Cross-national trends, policies, and practices* (pp. 3–16). Springer Singapore. https://doi.org/10.1007/978-981-13-9863-6_1

Lavrijsen, J., & Nicaise, I. (2016). Educational tracking, inequality and performance: New evidence from a differences-in-differences technique. *Research in Comparative and International Education, 11*(3), 334–349. https://doi.org/10.1177/1745499916664818

Lenkeit, J., Schwippert, K., & Knigge, M. (2017). Configurations of multiple disparities in reading performance: Longitudinal observations across France, Germany, Sweden and the United Kingdom. *Assessment in Education: Principles, Policy and Practice, 25*(1), 52–86. https://doi.org/10.1080/0969594X.2017.1309352

Levels, M., Kraaykamp, G., & Dronkers, J. (2008). Immigrant children's educational achievement in western countries: Origin, destination, and community effects on mathematical performance. *American Sociological Review, 73*(5), 835–853. https://doi.org/10.1177/000312240807300507

Marks, G. N. (2008). Accounting for the gender gaps in student performance in reading and mathematics: Evidence from 31 countries. *Oxford Review of Education, 34*(1), 89–109. https://doi.org/10.1080/03054980701565279

Meinck, S., & Brese, F. (2019). Trends in gender gaps: Using 20 years of evidence from TIMSS. *Large-Scale Assessments in Education, 7*(1), 1–23. https://doi.org/10.1186/s40536-019-0076-3

Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2017). *PIRLS 2016 International Results in Reading*. IEA TIMSS & PIRLS International Study Center Lynch School of Education, Boston College.

Mullis, I. V. S., Martin, M. O., Foy, P., Kelly, D. L., & Fischbein, B. (2020). *TIMSS 2019 international results in mathematics and science*. TIMSS & PIRLS International Study Center.

Nauck, B. (2019). Ethnic inequality in educational attainment. In R. Becker (Ed.), *Research handbook on the sociology of education.* Edward Elgar Publishing. https://doi.org/10.4337/9781788110426.00038

Neuville, E., & Croizet, J.-C. (2007). Can salience of gender identity impair math performance among 7–8 years old girls? The moderating role of task difficulty. *European Journal of Psychology of Education, 22*(3), 307–316. https://doi.org/10.1007/BF03173428

Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., Bar-Anan, Y., Bergh, R., Cai, H., Gonsalkorale, K., Kesebir, S., Maliszewski, N., Neto, F., Olli, E., Park, J., Schnabel, K., Shiomura, K., Tulbure, B. T., Wiers, R. W., … Greenwald, A. G. (2009). National differences in gender–science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Sciences, 106*(26), 10593–10597. https://doi.org/10.1073/pnas.0809921106

OECD. (2019b). *PISA 2018 technical report*. https://www.oecd.org/pisa/data/pisa2018technicalreport/

OECD. (2019a). PISA 2018 results (volume II): Where all students can succeed. In *OECD publishing: vol. II*. https://www.oecd.org/pisa/publications/PISA2018_CN_IDN.pdf

OECD. (2019c). *PISA 2018 results (Volume III): What school life means for students' lives*. OECD.

Reilly, D. (2012). Gender, culture, and sex-typed cognitive abilities. *PLoS ONE, 7*(7), 15–16. https://doi.org/10.1371/journal.pone.0039904

Rosén, M., Steinmann, I., & Wernersson, I. (2022). Gender differences in school achievement. In: T. Nilsen, A. Stancel-Piątak, J. -E. Gustafsson (Eds.), *International Handbook of Comparative Large-Scale Studies in Education* (pp. 1–48). Springer.https://doi.org/10.1007/978-3-030-38298-8_46-1

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. Wiley.

Ruhose, J., & Schwerdt, G. (2016). Does early educational tracking increase migrant-native achievement gaps? Differences-in-differences evidence across countries. *Economics of Education Review, 52*, 134–154. https://doi.org/10.1016/j.econedurev.2016.02.004

Steinmann, I., & Rutkowski, L. (2023). The link between gender gaps in school enrollment and school achievement. *Comparative Education Review, 67*(3). https://doi.org/10.1086/725395

Steinmann, I., Strietholt, R., & Rosén, M. (2023). International reading gaps between boys and girls, 1970–2016. *Comparative Education Review, 67*(2). https://doi.org/10.1086/724089

Strello, A., Strietholt, R., Steinmann, I., & Siepmann, C. (2021). Early tracking and different types ofinequalities in achievement: difference-in-differences evidence from 20 years of large-scale assessments. *Educational Assessment, Evaluation and Accountability*. https://doi.org/10.1007/s11092-020-09346-4

Strietholt, R. (2014). Studying educational inequality: Reintroducing normative notions. In R. Strietholt, W. Bos, J. -E. Gustafsson, & M. Rosén (Eds.), *Educational Policy Evaluation Through International Comparative Assessments* (Issue January, pp. 51–58). Waxmann Verlag

Strietholt, R., Gustafsson, J. -E., Hogrebe, N., Rolfe, V., Rosén, M., Steinmann, I., & Hansen, K. Y. (2019). The impact of education policies on socioeconomic inequality in student achievement: A review ofcomparative studies. In Socioeconomic Inequality and Student Outcomes: Cross-National Trends, Policies, and Practices (pp. 17–38). Springer. https://doi.org/10.1007/978-981-13-9863-6_2

Scheeren, L., & Bol, T. (2022). Gender inequality in educational performance over the school career: The role of tracking. *Research in Social Stratification and Mobility, 77*, 100661. https://doi.org/10.1016/j.rssm.2021.100661

Schmid, C. L. (2001). Educational achievement, language-minority students, and the new second generation. *Sociology of Education, 74*, 71. https://doi.org/10.2307/2673254

Schütz, G., Ursprung, H. W., & Wößmann, L. (2008). Education policy and equality of oppurtunity. *Kyklos, 61*(2), 279–308. https://doi.org/10.1111/j.1467-6435.2008.00402.x

Söhn, J., & Özcan, V. (2006). The educational attainment of Turkish migrants in Germany. *Turkish Studies, 7*(1), 101–124. https://doi.org/10.1080/14683840500520626

Stoet, G., & Geary, D. C. (2013). Sex differences in mathematics and reading achievement are inversely related: Within- and across-nation assessment of 10 years of PISA data. *PLoS ONE, 8*(3), e57988. https://doi.org/10.1371/journal.pone.0057988

Strand, S. (2011). The limits of social class in explaining ethnic gaps in educational attainment. *British Educational Research Journal, 37*(2), 197–229. https://doi.org/10.1080/01411920903540664

Strand, S. (2014). Ethnicity, gender, social class and achievement gaps at age 16: Intersectionality and "getting it" for the white working class. *Research Papers in Education, 29*(2), 131–171. https://doi.org/10.1080/02671522.2013.767370

Teltemann, J., & Schunck, R. (2016). Education systems, school segregation, and second-generation immigrants' educational success: Evidence from a country-fixed effects approach using three waves of PISA. *International Journal of Comparative Sociology, 57*(6), 401–424. https://doi.org/10.1177/0020715216687348

UNESCO. (2018). *Handbook on measuring equity in education*. https://doi.org/10.1016/S0733-8619(03)00096-3

UNESCO-IBE. (2012). *World data on education: Seventh edition 2010–11*. http://www.ibe.unesco.org/en/document/world-data-education-seventh-edition-2010-11

van de Werfhorst, H. G. (2018). Early tracking and socioeconomic inequality in academic achievement: Studying reforms in nine countries. *Research in Social Stratification and Mobility, 58*, 22–32. https://doi.org/10.1016/j.rssm.2018.09.002

van de Werfhorst, H. G., & Mijs, J. J. B. (2010). Achievement inequality and the institutional structure of educational systems: A comparative perspective. *Annual Review of Sociology, 36*(1), 407–428. https://doi.org/10.1146/annurev.soc.012809.102538

van Hek, M., Buchmann, C., & Kraaykamp, G. (2019). Educational systems and gender differences in reading: A comparative multilevel analysis. *European Sociological Review, 35*(2), 169–186. https://doi.org/10.1093/esr/jcy054

van Langen, A., Bosker, R., & Dekkers, H. (2006). Exploring cross-national differences in gender gaps in education. *Educational Research and Evaluation, 12*(2), 155–177. https://doi.org/10.1080/13803610600587016

World Bank. (2022). *GDP per capita (current US$)*. https://data.worldbank.org/indicator/NY.GDP.PCAP.CD

Worrell, F. C. (2014). Theories school Psychologists should know: Culture and academic achievement. *Psychology in the Schools, 51*(4), 332–347. https://doi.org/10.1002/pits.21756

Zell, E., Krizan, Z., & Teeter, S. R. (2015). Evaluating gender similarities and differences using metasynthesis. *American Psychologist, 70*(1), 10–20. https://doi.org/10.1037/a0038208