Yang Hansen *et al.*
Large-scale Assessments in Education (2022) 10:16
https://doi.org/10.1186/s40536-022-00133-9

Large-scale Assessments in Education

RESEARCH Open Access

Contextual effects on students' achievement and academic self-concept in the Nordic and Chinese educational systems

Kajsa Yang Hansen[1*], Jeléna Radišić[2], Yi Ding[1] and Xin Liu[3]

*Correspondence:
kajsa.yang-hansen@ped.gu.se

[1] Department of Education and Special Education, University of Gothenburg, Gothenburg, Sweden
[2] Department of Teacher Education and School Research, University of Oslo, Oslo, Norway
[3] Department of Educational Studies, Ghent University, Ghent, Belgium

Abstract

Background: The current study investigates school contextual effects on students' academic self-concept and achievement, that is, peer socioeconomic effect and big-fish-little-pond effect (BFLPE), in four Nordic education systems (i.e., Denmark, Finland, Norway and Sweden) and selected Chinese education systems (Hong Kong and Beijing-Shanghai-Jiangsu-Zhejiang). The two school contextual effects are studied simultaneously to negate the confounding effects of student and school socioeconomic status (SES), academic self-concept and academic achievement. The study focuses on the following research questions:

- Are there differences in the between-school variation in school SES composition, academic self-concept and reading literacy across the Nordic and selected Chinese education systems?
- What are the sizes of peer SES and BFLPE in the Nordic and Chinese education systems?
- Are there any differences that can be observed between the 2000 and 2018 data for the Programme for International Student Assessment (PISA) in the first two questions for Denmark, Finland, Norway and Sweden and Hong Kong?

Methods: Using PISA 2000 and 2018 data from Denmark, Finland, Norway, Sweden, Hong Kong and Beijing-Shanghai-Jiangsu-Zhejiang (in 2018 only), a multiple-group two-level structural equation modelling was applied to estimate the school contextual effect.

Results: The current study indicated that school academic and socioeconomic segregation intensified over the two decades in most of the studied countries. This finding lays the groundwork for understanding the two schools' contextual effects. School SES compositions positively affected students' academic achievement, while average school achievement negatively affected students' academic self-concept. Given that students' academic and socioeconomic composition has become more homogeneous within schools, the contextual effects were more pronounced. However, variations across diverse education systems were also notable.

Conclusions: The current study adds evidence of the school contextual effects regarding the peer SES effects and robustness of the BFLPE, considering the nested structure of the data and diverse cultural milieu. We argue that these diversities reflect the varying extent to which educational reforms were introduced in different countries, here as oriented to marketisation, privatisation and choice. These reform actions have changed the composition of students at schools and school culture, which moulds individual students' cognitive and noncognitive development.

Keywords: Peer SES effect, BFLP effect, PISA, SES, Academic self-concept, The Nordic education systems, Hong Kong, BSJZ

Introduction

The peer composition of the learning environment is undeniably influential for students' academic achievement and development (e.g., van Ewijk & Sleegers, 2010a, 2010b; Wang et al., 2020; Zhang, 2021). These effects of peer characteristics were described as 'peer group effect' (Evans et al., 1992; Hoxby, 2000; Yeung & Nguyen-Hoang, 2016), 'compositional effect' (Liu et al., 2015; Van Damme et al., 2002), 'contextual effect' (Raudenbush et al., 1992; Willms, 2010) or 'aggregated group-level effect' (Hutchison, 2003; Marsh et al., 2012). There are indeed differences between these terms. However, the fundamental principle emphasises that a student body's collective properties affect students' cognitive and noncognitive outcomes over and above their own characteristics (e.g., Harker & Tymms, 2004; Willms, 2010). Depending on the outcomes and contextual characteristics in focus, the contextual effects are defined differently; for example, the peer SES effect or the big-fish-little-pond effect (BFLPE). The former is the effect of peer socioeconomic status (SES) composition on students' school performance (e.g., van Ewijk & Sleegers, 2010a). The latter is described as the effect of average school achievement level on academic self-concept (for a review, see Fang et al., 2018; Marsh & Craven, 2006). The peer SES effect and BFLPE may confound and intertwine, requiring that the two contextual effects be studied simultaneously, even though they have been predominantly investigated separately across the adjacent fields of psychology and education. In the present study, we attempt to place the two phenomena into a dialogue and investigate the school contextual effects on students' academic self-concept and achievement across the Nordic and selected Chinese metropolitans. Two dominant reasons steer the choice of very different geographical locations: first, the diversity of educational traditions and institutional reforms over the past two decades taking place across these regions (e.g., Frønes et al., 2020; Blossing et al., 2014; Liu et al., 2019) and, second, the diversity in achievement levels in the Chinese subpopulations (i.e., respective benchmarking participants) and Finland versus the other Nordic countries, as indicated by the international large-scale assessments (ILSAs).

Programme for International Student Assessment (PISA) data for the years 2000 and 2018 from the Nordic and selected Chinese education systems[1] were analysed using multigroup, two-level structural equation modelling. As a rule, contextual studies can evaluate whether group-level characteristics (e.g., school SES and/or achievement composition) contribute to individual's cognitive (e.g., student academic achievement) and

[1] Beijing-Shanghai-Jiangsu-Zhejiang only in 2018.

noncognitive (e.g., student's self-concept) outcomes, beyond what can be explained by the students' own characteristics (e.g., family SES and ability level). Therefore, the current analysis estimated the size of the two contextual effects: the *peer SES effect* and *BFLPE*.

The peer SES effects on students' learning

Numerous studies have examined the associations between the peer effect of SES and academic performance (e.g., Van Ewijk & Sleegers, 2010a; Yeung & Nguyen-Hoang, 2016). These studies have found, on average, a positive peer SES influence, indicating that students who attend a school with a higher SES composition than their own will achieve higher academic outcomes (e.g., Ammermueller & Pischke, 2009). However, the size of the estimates varies extensively, ranging from nonsignificant to moderately large (e.g., Sacerdote, 2011, 2014). The results of the meta-analysis conducted by Van Ewijk and Sleegers (2010a) revealed that the SES compositional effect was strongly related to SES's operationalisation and the choice of the analytical model applied (see also Angrist, 2014).

Manski (1993) identified three types of peer effects: endogenous, exogenous and correlated. Endogenous effects refer to the students' behaviours or outcomes that are affected by the behaviours or outcomes of the peer group (e.g., Yeung & Nguyen-Hoang, 2016). Exogenous effects are related to the influences of group compositional or contextual features on outcomes, such as SES and ethnicity (e.g., Van Ewijk & Sleegers,). Correlated effects refer to students being selected to the same group tend to have a similar achievement level because they share similar individual characteristics and the same learning environment. Different models have been proposed for these three peer effects to estimate their effect size (Manski, 1993).

Focusing on the endogenous peer effect, Yeung and Nguyen-Hoang (2016) summarised the results of 53 studies that were carried out from 1980 to 2014 on endogenous peer effects in education; they found a highly significant peer effect on an individual's school outcome. However, the effect size was relatively small (e.g., Sund, 2009). They also noticed that the endogenous peer effect, as moderated by the educational outcome of a school subject, the choice of peer group, and publication status of a study, had moderating effects on endogenous peer effects. Hence, it may be that the extensive control for confounding covariates can lead to underestimation of the peer SES effect (e.g., Van Ewijk & Sleegers, 2010a; see also Boucher et al., 2014; Burke & Sass, 2013) because the assignment of students in schools and classrooms is unlikely to be random (i.e., because of residential segregation, school choice and self-selection mechanism), even for school systems with a proximity recruitment principle and no tracking. The current study focuses on the exogenous peer effect of SES.

Again, previous studies have shown no consensus on the effect size for these peer SES effects. For example, Hoxby (2000) used random within-school variations across cohorts to account for selection bias. The peer measures in Hoxby's study were the gender and race composition of the classroom and performance by opposite gender and race group. The study found relatively strong peer effects (for more details, see Hanushek et al., 2003; Hanushek et al., 2009). Ammermueller and Pischke (2009) studied peer effects in six European countries with 2001 data from the Progress in International Reading Literacy

Study. By focusing on the within-school variation between classes, they estimated the effect of a peer's sociodemographic composition (i.e., number of books at home, gender, age, ethnicity and background) on an individual's reading achievement. The authors applied a fixed effect model and instrumental variable approach to account for the selection bias and reflection problem,[2] finding a moderately large peer effect of about 0.17.

In their meta-analysis of the exogenous peer effect of SES, van Ewijk and Sleegers (2010a) argued that there were more negligible effects if SES was dichotomously measured (e.g., free lunch eligibility) or composition was measured at the cohort/school level (see also Gustafsson & Yang Hansen, 2018). In general, the meta-analysis found a considerable average weighted effect size of all the studies included, with 'an increase of the average socioeconomic status of a student's peer group with one student-level standard deviation leading to an increase of test score with 0.320 SD (p. 143; van Ewijk & Sleegers, 2010a). However, the peer effect the researchers have found highly depends on the analytical approach taken. The study also pointed out that omitted covariates may overestimate the peer SES effect. Moreover, prior meta-analytic studies found the effect of compositional SES on students' achievement to be over and above the effects at the individual level (e.g., Hattie, 2008; Sirin, 2005).

In addition, as a crucial contextual determinant of academic performance, peer SES positively influences the student's academic self-concept (e.g., Chiu & Chow, 2015; Marsh, 1984; Tucker-Drob & Harden, 2012). Observing fourth-grade students' reading achievements in 33 countries, Chiu and Chow (2015) showed that classmates' family factors (i.e., parent SES or home educational resources) were more strongly related to an individual student's reading achievement than to classmates' characteristics (e.g., attitudes towards reading). Among the examined predictors, classmates' home education resources were the third strongest link to a student's reading achievement. Tucker-Drob and Harden (2012) reported similar relations in observing gene-by-environment interactions and mathematics achievement. However, past research has seldom focused on the school peer SES effects on academic self-concept (Marsh, 1984), let alone using Western and Eastern countries for comparison (Chiu & Chow, 2015). Instead, in the latter case, only the overall patterns have been observed, reporting differences between wealthier and more collectivist countries against those that were not.

Unfolding the big-fish-little-pond effect (BFLPE)

Academic self-concept (ASC), here referring to an individual's general perceptions of academic abilities, may refer to the general academic domain and/or particular subject matter (e.g., mathematics, English, science). Both have been an object of research since the 1970s (Marsh & Shavelson, 1985; Shavelson et al., 1976). Extensive research has shown that academic self-concept is not only an essential factor in determining students' academic achievement (Bong & Skaalvik, 2003; Fang et al., 2018; Hoferichter et al., 2018; Marsh, 1984; Marsh et al., 1988), but it is also closely

[2] The reflection problem is one of the challenges in estimating peer effects. It is also called the simultaneity problem and arises when the achievement of peers has an influence on individual student's achievement, which also makes up the peer group and influences their peers. This is the so-called simultaneity problem (Manski, 1993).

associated with other behavioural and emotional outcomes—academic self-esteem, motivation and anxiety (Bong & Clark, 1999; Marsh & Martin, 2011).

One of the most influential theories on ASC in educational psychology might be the BFLPE (Hoferichter et al., 2018), which was proposed by Marsh and Parker (1984). The concept is based on a social comparison process and frame of reference theory (for a detailed account of the BFLPE's theoretical background, see Craven et al., 2000). The BFLPE refers to the phenomenon that equally able students in higher-achieving or selective schools tend to have lower ASC than those in lower-achieving and nonselective schools (Marsh & Parker, 1984). The BFLPE was first detected by Marsh (1984), who showed that students' ASC was positively influenced by individuals' achievement and SES but was negatively influenced by school-level achievement and SES. Marsh (1984) also argued that how students perceived their ASC was related to whether they compared themselves with their peers in class or school. Since then, BFLPE has been investigated by researchers worldwide (e.g., Fang et al., 2018; Marsh, 1990, 1991; Marsh & O'Mara, 2008; Marsh & Yeung, 1997). Factors such as students' subsequent course selection (Marsh & Yeung, 1997), test anxiety and educational attainment (Marsh & O'Mara, 2008) and social comparison (Marsh et al., 2008a, b) were explored and found to be associated with the BFLPE.

The BFLPE is considered a powerful, long-lasting contextual effect that generalises across different research sites, levels of education, and cultures around the world (2008b, 2021; Fang et al., 2018; Marsh et al., 2008a). An example of its long-lasting effect can be found in Marsh et al.'s (2001) study focusing on BFLPE before and after reunification in Germany. The BFLPE was found among students in East and West Germany and primarily disappeared by the end of the first postreunification school year. Among the arguments presented, the study argued that the formation of ASC was strongly influenced by changes in school policy (Marsh et al., 2001). The BFLPE has also been investigated across Western, Asian and Middle Eastern Islamic countries. Using data from the Trends in International Mathematics Science Study (TIMSS), researchers have found that ASC was positively affected by individual achievement but negatively by class-average achievement and that such results were generalisable across 13 diverse countries and over two cohorts (Marsh et al., 2015). Marsh and Hau (2003) examined BFLPE with an even larger sample from PISA 2000; the study results demonstrated cross-cultural generalisability in the BFLPE across academically selective schools in 26 counties.

The most recent meta-analysis (Fang et al., 2018) reviewed numerous studies within the field and synthesised some interesting results. The researchers investigated several potential moderators, such as student age, comparison target, ASC in different subject domains and students' location (i.e., country). The results revealed a few interesting points: (1) The BFLPE demonstrated cross-cultural generalizability. (2) Student age was a significant moderator of the BFLPE. (3) The BFLPE can be understood through social comparison theory, indicating that the level of students' ASC might be lower in higher-ability educational settings because of the unpleasant comparison referring to their peers. (4) The equally able students tended to perceive more positive ASC in a nonselective and less-competitive environment (Fang et al., 2018).

Diverse educational paths and traditions from the Nordic and Far East

A critical facet of the current study lies in its direct focus on comparing the peer SES effect and BFLPE across the four Nordic education systems and Chinese subpopulations (i.e., Beijing-Shanghai-Jiangsu-Zhejiang and Hong Kong). Although the cross-cultural generalizability of the BFLPE has been argued (Fang et al., 2018), coupled with diverse patterns across different cultures reported for the peer socioeconomic effect (Chiu & Chow, 2015), more focused analyses have been lagging.

In the post-World War II period, the Nordic countries introduced an inclusive and comprehensive education model, here with the primary goal of abolishing the class-based society and enabling social mobility (Frønes et al., 2020; Blossing et al., 2014; Telhaug et al., 2006). The newly established education system practised egalitarian philosophy through a smooth transition between the educational levels without tracking students to ensure equal learning opportunities for all (Husén, 1989; Imsen et al., 2017; Lundahl, 2016).

Over time, and especially with the arrival of neoliberal thinking and the economic trends at the time, distinctive differences have appeared among the Nordic systems (Imsen et al., 2017; Volckmar & Wiborg, 2014), and a debate on the sustainability of such systems has followed (Antikainen, 2006).

Although practices have become evidently diverse (Volckmar & Wiborg, 2014), Sweden was at the forefront of going through extensive changes characterised by decentralisation (Blossing & Söderström, 2014) and severe marketisation and privatisation (Lundahl, 2016). In Denmark, the competitive discourse has also taken its toll (Rasmussen & Moos, 2014). In Finland, polarisation between the schools became evident mainly in the unequal distribution of municipality funds and SES composition of the students attending the schools (Ahonen, 2014). In Norway, education policy has resisted the privatisation of the school market (Imsen & Volckmar, 2014) but has not been immune to accountability practices, which have been introduced gradually across the Nordic countries (Imsen et al., 2017; Wallenius et al., 2018; Wollscheid & Opheim, 2016).

Conversely, although the Nordic countries are associated with 'nature-centeredness, wellbeing, comfort, simplicity and *getting it right*' (Liu et al., 2019, p. 3), China is often depicted as '*an economic superpower said to be torn apart between aspiration and authoritarianism*' (Liu et al., 2019, p. 3), indicating very diverse political and cultural backgrounds. In the ILSAs, the Nordic countries, Finland in particular, have amazed the world with outstanding academic performances, despite relatively shorter school days, much less homework and fewer standardised tests (Liu et al., 2019). Chinese students, on the other hand, have been stereotyped as learning within a competitive environment with excessive working hours, and have also achieved unexpected performances in the ILSAs.[3]

In recent years, educational research has shown growth in the interaction and intersection between the Nordic countries and China (Liu et al., 2019). However, much remains uncertain or unknown. Notably, most BFLPE research has been carried out in various educational systems in European countries (e.g., Germany, France and the UK), East Asian

[3] We refer here only to the respective benchmarking countries that have participated in PISA.

countries (e.g., Hong Kong and Singapore) and the United States (Marsh et al., 2015). Students' location (i.e., country) has been found to play an essential role in BFLPE. Based on a meta-analysis, Asian students have been found to show the most substantial effect, whereas European students seemed to show a moderate effect regarding BFLPE (Fang et al., 2018). Stankov (2010, p. 559) argued that the influence of the 'unforgiving' Confucian culture on Asian students should be considered when being compared with the European regions. The paradoxical phenomenon where Confucian Asians (i.e., Chinese students in this case) are high achievers but hold adverse psychological outcomes (e.g., self-concept and self-efficacy) can perhaps be explained by the cultural impact because the students tend to be more anxious and self-doubting (Stankov, 2010).

Current study

The most commonly used analytical approach to the peer SES effect is within the econometric tradition and its use of fixed effect models (e.g., Manski, 1993; Sacerdote, 2011; Sund, 2009), while the BFLPE is viewed through hierarchical linear models with cross-level interaction and random slopes (Marsh et al., 2008a, b). Marsh et al. (2009a, 2009b) argued that the obstacles with these approaches relate to the measurement error of the constructs and sampling error, which attenuate the estimation of the contextual effects. Moreover, the peer SES effect and the BFLPE are often investigated separately—the former in school effectiveness studies and the latter in the more educational psychology-oriented studies. Very seldom is an integration of hierarchical and structural equation model techniques applied in estimating these two effects simultaneously. Such studies show that students' self-concept is culture dependent—and has been found to be more adverse in the high-achieving Asian countries and Finland, while it is the opposite in other Nordic countries (Leung, 2014). The schools' institutional and sociodemographic characteristics and intakes are different between the Nordic countries and China (e.g., Frønes et al., 2020; Blossing et al., 2014; Liu et al., 2019).

The current study aims to examine the school contextual effects, that is, peer SES effect and BFLPE, in the Nordic countries (Denmark, Finland, Norway and Sweden) and select Chinese subpopulations (Beijing-Shanghai-Jiangsu-Zhejiang and Hong Kong), unravelling the confounding among student and school SES, academic self-concept and academic achievement. The following research questions are the focus of the present study:

1. Are there differences in the between-school variation in school SES composition, academic self-concept and reading literacy across the Nordic countries and Chinese subpopulations?
2. What are the peer SES effect and BFLPE sizes in the Nordic countries and Chinese subpopulations?
3. Are there differences that can be observed between PISA 2000 and PISA 2018 concerning the first two research questions for Denmark, Finland, Norway and Sweden and Hong Kong?

Table 1 Sample size of the countries in PISA 2000 and 2018

	PISA 2000			PISA 2018		
	$N_{student}$	N_{school}	Average school size	$N_{student}$	N_{school}	Average school size
Denmark	4174	224	18.6	7431	348	21.4
Finland	4813	155	31.1	5557	214	26.0
Norway	4096	176	23.3	5612	250	22.4
Sweden	4385	154	28.5	5348	222	24.1
Hong Kong	4389	140	31.4	5839	152	38.4
B-S-J-Z (China)	–	–	–	11,990	361	33.2
Total	21,857	849	–	41,777	1547	–

B-S-J-Z (China) Beijing, Shanghai, Jiangsu and Zhejiang

Methods

The current study used PISA 2000 and 2018 data from four Nordic countries (i.e., Denmark, Finland, Norway and Sweden) and selected Chinese educational systems (Hong Kong and Beijing-Shanghai-Jiangsu-Zhejiang). The main subject domain in both cycles was reading literacy. It should be noted that for China, B-S-J-Z participated only in PISA 2018. However, it has been included in the current analysis to validate the observed pattern difference between the Nordic and Chinese regions. Table 1 shows the number of students and schools included in the study. The sample size at both the student and school levels in PISA 2018 was generally more extensive than in PISA 2000, except for Hong Kong.

As shown in Table 1, the number of students sampled in PISA 2000 was somewhat even across the examined countries. However, in PISA 2018, the two Chinese samples differed significantly from the others. Compared with the Nordic countries, B-S-J-Z had about twice as many students as the other countries, while Hong Kong had only about half the size. The Danish student sample was also substantial. Despite the difference in sample sizes, there were sufficient students and schools for conducting a two-level analysis. Overall, 21,857 students from 849 schools from PISA 2000 and 41,777 students from 1,547 schools from PISA 2018 have been included.

Variables

The index of economic, social and cultural status (ESCS) can be used to measure family SES. In PISA 2018, the ESCS index was defined as the mean score of three indices: *highest parental occupation*, *parental education* and *home possessions*. The missing values in these indices were imputed.[4] The imputed indices were standardised to an OECD mean of zero and using the standard deviation as a unit. The ESCS was calculated as the arithmetic means of the imputed and standardised indices (Avvisati, 2020; OECD, 2019). In PISA 2000, the ESCS was based on five components—*highest parental occupations*, *parental education*, *family wealth*, *home educational resources*, and *cultural possessions*. The ESCS was an estimated score of the first factor in a principal

[4] If two of the indices are missing, ESCS will not be missing. No imputation will be made (OECD, 2019).

component model (Adams & Wu, 2002). In the current study, school SES composition was measured as the aggregated individual-level ESCS.

Students' self-concept was captured differently in the two PISA studies. In PISA 2018, an index of self-concept in reading competence (SCREADCOMP) was used. Three indicators were used to create the SCREADCOMP index: *I am a good reader*, *I am able to understand difficult texts* and *I read fluently*. The reliability of the SCREADCOMP index, as measured by Cronbach's alpha, was around 0.80, Hong Kong held the lowest value of 0.79. and Sweden the highest at 0.87, with all other education systems being over 0.84. In PISA 2000, students' self-concept was based on their perception of a general academic level (i.e., academic self-concept, SCACAD). The SCACAD consisted of three indicators: *I learn things quickly in most school subjects*, *I'm good at most school subjects* and *I do well in tests in most school subjects*. Cronbach's alpha for the SCACAD index ranged between 0.75 in Hong Kong and 0.84 in Finland, with Denmark and Sweden being 0.80 and 0.81, respectively. The responses to all the indicators were given on a 4-point Likert scale with the categories 'Strongly disagree' to 'Strongly agree'. The self-concept indices in both PISA cycles were IRT weighted likelihood estimates, the so-called Warm estimates of the latent traits, here following the scaling procedure in PISA studies (e.g., Adams & Wu, 2003; OECD, 2019). These estimates were unbiased against random sources of errors in the latent constructs. Confirmatory factor analysis, for example, was used in PISA 2000 to validate whether the latent trait estimation may be biased by model misspecification—for example, not unidimensional (e.g., Feuerstahler, 2018). Other psychometric techniques were also applied to ensure the comparability of the estimated latent constructs (for a detailed description for scaling procedures and construct validation, see chapter 17 in Adams & Wu, 2003 and chapter 16 in OECD, 2019).

The main subject domain tested in both PISA cycles was reading. PISA provided several estimated proficiency scores, the so-called plausible values, to represent students' test achievement (Adams & Wu, 2002; OECD, 2019). The current analysis used all plausible values of reading achievement (i.e., five plausible values for each student in PISA 2000 and 10 in PISA 2018). Mplus estimated the model parameters repeatedly based on each plausible value. The final parameter estimates reported in the study were achieved by averaging the sets of parameter estimates. The variance of the parameter estimates was a combination of the average variance of the parameter variance based on each plausible value and variation of the estimates across all plausible values (e.g., Von Davier et al., 2009). Table 2 presents the descriptive statistics of the variables involved in the analysis from PISA 2000 and PISA 2018.

As shown in Table 2, the students in the Finland, Hong Kong and B-S-J-Z samples achieved higher scores on reading competence but had a lower self-concept related to reading and the general perception of their academic outcomes compared with their Nordic peers. The ESCS level was lower in Hong Kong and B-S-J-Z China, than in other Nordic countries included in the study. Differences within each country can also be observed in the samples between the two PISA cycles, even though one should be cautious about a direct comparison of these observed statistics. Senate weight was used to adjust for unequal sample sizes across these education systems.

Table 2 Mean and standard deviation of the variables in countries and PISA cycles

	Denmark		Finland		Norway		Sweden		Hong Kong		B-S-J-Z (China)	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
PISA 2000												
Academic self-concept	.41	.97	− .05	1.03	.04	1.17	.09	.88	− .50	.94	–	–
ESCS	.29	.91	.05	.093	.38	.83	.35	.79	− .82	.87	–	–
Reading Achievement	497.92	96.20	548.52	87.51	505.37	103.10	515.86	91.80	525.82	83.55	–	–
Valid N	4087		4744		3929		4340		4361		–	
PISA 2018												
Self-concept in reading competence	.31	.94	.09	1.00	.21	1.03	.32	1.04	− .22	.92	.08	.87
ESCS	.39	.84	.30	.79	.53	.82	.36	.89	− .52	1.02	− .36	1.09
Reading achievement	489.06	94.92	519.94	99.58	497.39	105.97	505.44	106.44	527.33	97.90	561.03	90.34
Valid N	6748		5376		5379		5156		5716		11,904	

SD standard deviation, *B-S-J-Z (China)* Beijing, Shanghai, Jiangsu and Zhejiang. Mean and standard deviation of reading achievement is based on the first plausible value

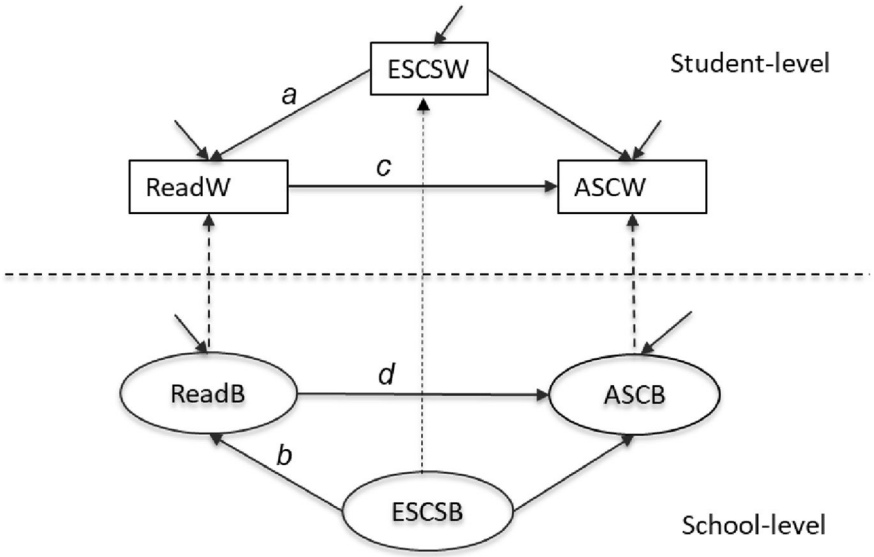

Fig. 1 Path diagram of the two-level structural model of Peer and BFLP effects

Analytical method

The literature has used a wide variety of approaches to identify peer effects (Manski, 1993). The peer SES effect on students' academic achievement was highly related to the operationalisation and model choice. In a meta-analysis by van Ewijk and Slanger (2010b), it was found that the measurement scale and level of SES mattered for the size of the peer effect. For example, peer effects will be more negligible when SES is measured as a dichotomous variable or at the school level. Moreover, one may overestimate the peer effect without controlling for prior achievement, and extensive control for many covariates may underestimate the effect (e.g., Sund, 2009). The estimated peer SES effect ranged from nonsignificant to strong on the students' academic achievement (e.g., van Ewijk and Slanger, 2010a).

The current study examines how individual-level relationships are affected by aggregated contextual characteristics. This required simultaneous analysis of the individual and school levels. Moreover, PISA surveys have applied a multistage cluster sampling design to collect data, resulting in a hierarchical data structure, that is, students nested within schools. Using ordinary analytical methods to analyse clustered data would lead to a type I error because of underestimating the standard error and possibly omitting some crucial relationships involving each data level. A two-level structural equation modelling technique decomposed the individual data into between-school and within-school components. The two components 'can be used to compute a between-groups covariance matrix (i.e., the population covariance matrix of the disaggregated group means) and a within-groups covariance matrix' (i.e., the population covariance matrix of the individual deviations from the group means; Hox et al., 2017, p. 290). Thus, the within-group covariance matrix has been group-mean centred. Covariates can be introduced to account for the variances in the outcome variables at both levels. This process allows for examining the interrelationships among variables by estimating direct and indirect effects at the respective levels (Hox

Table 3 Intraclass Correlation Coefficients (%) of self-concept, socioeconomic status and reading achievement across countries and cycles

	PISA 2000			PISA 2018		
	Academic self-concept	Reading achievement	ESCS	Self-concept in reading competence	Reading achievement	ESCS
Denmark	1.5	16.6	14.7	1.9	17.4	20.1
Finland	1.2	7.8	13.7	1.5	7.8	13.3
Norway	2.0	9.1	11.4	1.3	8.7	8.0
Sweden	1.7	8.9	13.5	1.3	16.5	15.0
Hong Kong	3.2	48.1	23.1	3.5	32.3	30.1
B-S-J-Z (China)	–	–	–	8.7	47.1	45.0

B-S-J-Z (China) Beijing, Shanghai, Jiangsu and Zhejiang

et al., 2017). The intraclass correlation coefficient (ICC) describes the proportion of the total variance in an outcome variable because of school belongingness. The ICC can also be interpreted as a measure of the homogeneity of the students in the same school. All variables have been standardised before modelling in Mplus (Muthén & Muthén, 1998–2017). Moreover, the number of missing cases was generally very low for each variable across cycles and countries (see Appendix A). Mplus uses all available data to estimate models with a full information maximum likelihood estimator (FIML), assuming that missing data are random.

The peer effect and BFLPE were estimated simultaneously in a two-level SEM model. The model structure is shown in Fig. 1 (see Appendix B for the Mplus model input). At the individual level, the relationship between students' self-concept in academic/reading competence and their reading achievement was controlled by the students' ESCS. This relationship has been denoted as *c* in the path diagram.

The confounding effect of the school composition of students' ESCS with both average school self-concept and reading achievement was also adjusted. Such confounding effects need to be cancelled out; otherwise, the BFLPE will be overestimated (2008b; Marsh et al., 2008a). Students' SES effect on their reading achievement has been denoted in Fig. 1 as a. The SES composition of the school has an impact on average school achievement and has been denoted as b. At the school level, d denotes the impact of school average students' achievement on the school average self-concept of its students.

A contextual effect exists when the between-school regression coefficient differs significantly from the within-school regression coefficient. This hypothesis has been tested by estimating two new parameters: the peer effect $b - a$ and the BFLPE $d - c$. The two new parameters are specified using 'Model Constraints' in Mplus, which allows for testing of the significance of the parameters.

Results

This section describes the ICC estimations and observes the examined peer and BFLP effects for all countries (RQ1 and 2) and cycles studied (RQ3). It should be noted that because self-concept measures differ in the two PISA cycles, a strict comparison over time is not possible.

Intraclass correlation coefficients

The intraclass correlation coefficient captured the between-school differences in the students' self-concept (i.e., SCACAD in PISA 2000 and SCREADCOMP in PISA 2018), family socioeconomic background and reading achievement measures. Hong Kong and B-S-J-Z China observed a higher variation between different schools than in the Nordic countries (see Table 3). This result indicates a more substantial selection effect in the two Chinese education systems, making the students in the same school considerably similar in terms of student ability level and personal and sociodemographic characteristics.

For reading achievement, almost half of the variation can be because of students attending different schools in Hong Kong PISA 2000 (48%) and B-S-J-Z PISA 2018 (47%), while this was between about 8% and 17% for the Nordic countries in both PISA 2000 and 2018. Among the Nordic countries in PISA 2000, Finland held the lowest ICC (7.8%), while Denmark had the highest (16.6%). The ICC in Norway and Sweden were at the same level, being 7.8% and 9.1%, respectively. In PISA 2018, the ICCs in Denmark and Sweden amounted to 17.4% and 16.5%, respectively, while Finland and Norway remained relatively stable.

In PISA 2000, Hong Kong observed the highest between-school difference in ESCS (23.1%), and the percentage for Nordic countries was around 13%. The between-school difference in ESCS increased in Hong Kong (30.1%), Denmark (20.1%) and Sweden (15%) in PISA 2018, and Finland and Norway held relatively consistent ICCs in ESCS. Here, 45% of the variation in ESCS in B-S-J-Z can be attributed to students attending different schools. ICCs for students' self-concept, general academic (SCACAD) and reading-specific (SCREADCOMP), are generally small for the Nordic education systems, being around 1–2%, and much higher in Hong Kong and B-S-J-Z in PISA 2018.

Relationship between students' self-concept, reading achievement and SES (RQ1)

In PISA 2000, the students' reading achievement was positively related to their academic self-concept. The standardised regression coefficients between students' general academic self-concept (SCACAD), SES and reading achievement are shown in Table 4. The estimates were evenly high for the Nordic country students, at around 0.40. At the same time, the estimate was much lower for Hong Kong students (0.14). The regression coefficients of students' family SES on their reading achievement were also significant and positive for all countries, ranging from 0.32 in Denmark to 0.08 in Hong Kong. The small coefficient of students' family SES on their general academic self-concept (SCACAD) was also observed in all countries, with Norway having the highest effect (0.17) and Hong Kong the lowest (0.05).

School SES composition was strongly related to reading achievement at the school level. Denmark and Sweden had higher regression coefficients (0.82 and 0.84, respectively) than Hong Kong and Norway (0.67 and 0.62, respectively). Finland had the lowest estimate (0.44). The relationship between school reading achievement and school average academic self-concept (SCACAD) was relatively high in Denmark and Finland, at.78 and 0.72, respectively. In contrast, they were not statistically significant in other countries. School SES composition was significantly associated with academic self-concept (SCACAD) only in Hong Kong (0.65) and Norway (0.54).

Table 4 Standardised estimation of the relationship among student's academic self-concept, SES and reading achievement at individual and school levels

	Denmark		**Finland**		**Norway**		**Sweden**		**Hong Kong**		**B-S-J-Z (China)**	
PISA 2000												
Student-Level	Est	z	Est	z	Est	z	Est	z	Est	z	Est	z
SCACAD on READ	.40	26.05	.41	28.49	.44	27.20	.39	24.95	.14	7.93	-	-
SCACAD on ESCS	.13	8.16	.14	10.35	.17	10.57	.12	7.47	.05	2.89	-	-
READ on ESCS	.32	19.17	.24	17.04	.26	15.96	.25	15.76	.08	4.50	-	-
School-Level												
SCACAD_B on READB	*.78*	2.36	*.72*	4.39	*.26*	*1.19*	*− .66*	*− 1.38*	*− .06*	*− .32*	-	-
SCACAD_B on ESCSB	*− .19*	*− .55*	*.27*	*1.65*	.54	2.75	*.62*	*1.30*	.65	4.15	-	-
READB on ESCSB	.82	17.69	.44	3.64	.62	7.84	.84	16.49	.67	8.26	-	-
PISA 2018												
Student-Level	Est	z	Est	z	Est	z	Est	z	Est	z	Est	z
SCREADCOMP on READ	.37	28.37	.40	27.88	.38	26.78	.34	23.31	.22	15.54	.14	10.86
SCREADCOMP on ESCS	.08	6.13	.07	4.87	.13	8.54	.09	6.11	.11	8.03	.18	19.75
READ on ESCS	.24	16.57	.27	17.50	.24	16.29	.24	14.99	*.02*	*1.29*	.11	9.30
School-Level												
SCREADCOMP_B on READB	*.26*	*1.04*	.47	2.32	*.18*	*.85*	*− .20*	*− .45*	.33	2.56	− .16	− 2.40
SCREADCOMP_B on ESCSB	*− .16*	*− .63*	*.42*	*1.91*	.44	1.99	*.58*	*1.24*	.52	4.11	.88	14.70
READB on ESCSB	.84	27.81	.70	10.40	.58	7.60	.81	19.23	.62	10.43	.71	26.98

SCACAD student's academic self-concept, *SCACAD_B* school average academic self-concept, *SCREADCOMP* individual self-concept in reading competence, *SCREADCOMP_B* school average self-concept in reading competence, *READ* reading achievement, *READB* school average reading achievement, *ESCS* student's family social, cultural economic status, *ESCSB* school composition of students' social, cultural economic status, *B-S-J-Z (China)* Beijing, Shanghai, Jiangsu and Zhejiang. Estimates in Italic are not statistically significant, $p < .05$

Peer SES and BFLPE in the Nordic and Chinese education systems (RQ2)

Peer SES effect

School SES contextual effect boosted students' reading achievement. Peer SES effects were positive, indicating that students' reading achievement was higher when attending a high SES school. In PISA 2000, Hong Kong held the highest peer effect, 0.84, followed by Denmark (0.51) and Sweden (0.49). A nonsignificant peer effect was observed in Finland.

Big-fish-little-pond effect

Table 6 shows that a significant BFLPE was observed in all countries in the 2000 cycle, except for Denmark. These results align with previous findings (e.g., Fang et al., 2018; Marsh & Hau, 2003). Sweden had the highest BFLP effect in PISA 2000 (− 0.68). When the students attended a school with one standard deviation higher than their own achievement in Sweden in the year 2000, students' self-concept (SCACAD) was a 0.68 standard deviation lower. Hence, the intensity of the BFLP effect may be related to the students' intake composition in a school.

Differences across 2000 and 2018 cycles for peer SES and BFLPE effects (RQ3)

Regarding the PISA 2018 cycle results, students' reading achievement was positively related to their self-concept of reading competence (SCREADCOMP). The regression coefficients were relatively even across the Nordic countries, ranging from 0.40 in Finland to 0.34 in Sweden. The coefficients were lower for Chinese students, being 0.22 in Hong Kong and 0.14 in B-S-J-Z, China. A similar pattern was found for the SES effect on students' reading achievement. In the Nordic countries, around 6% of individual achievement differences can be accounted for by family ESCS. At the same time, ESCS was not an essential factor for B-S-J-Z students. The effect of students' SES on their self-concept of reading competence was low but significant in all countries. B-S-J-Z, China students, held the highest regression coefficient at 0.18.

At the school level, the effect of school reading achievement on school average self-concept in reading competence was found to be significant in Finland (0.47), Hong Kong (0.33) and B-S-J-Z, China (− 0.16). The effect of school SES composition on school average self-concept in reading competence was statistically significant in Norway (0.44), Hong Kong (0.52) and B-S-J-Z, China (0.88). School SES composition also affected school reading achievement significantly; the effect ranged from 0.84 in Denmark to 0.58 in Norway.

In sum, for both PISA cycles, the relationship between students' SES, academic self-concept (i.e., SCACAD in PISA 2000 and SCREADCOMP in PISA 2018) and reading achievement was significant at the individual level, higher in the Nordic countries and lower for Chinese students. School SES composition significantly affected the school's average reading achievement at the school level. The relationship between school reading self-concept (SCREADCOMP), reading achievement, and school SES composition was significant only in Finland and B-S-J-Z, China.

Table 5 Estimated Peer effect in the six countries and two PISA cycles

	PISA 2000		PISA 2018	
Countries	**Estimate**	**z-value**	**Estimate**	**z-value**
Denmark	.51	5.03	.60	11.14
Finland	*.08*	*.81*	.31	3.17
Norway	.35	3.54	.48	3.75
Sweden	.49	5.01	.73	7.95
Hong Kong	.84	6.24	.60	8.23
B-S-J-Z (China)	–	–	.51	15.28

Estimates in Italic are not statistically significant, $p < .05$

B-S-J-Z (China) Beijing, Shanghai, Jiangsu and Zhejiang

Table 6 Estimated BFLP effect in the six countries and two PISA cycles

Countries	PISA 2000		PISA 2018	
	estimate	**z-value**	**estimate**	**z-value**
Denmark	*− .16*	*− 1.28*	− .32	− 3.69
Finland	− .17	− 2.08	− .21	− 2.01
Norway	− .37	− 2.93	− .32	− 3.75
Sweden	− .68	− 3.38	− .41	− 3.73
Hong Kong	− .24	− 4.00	− .14	− 3.40
B-S-J-Z (China)	–	–	− .23	− 7.86

Estimates in Italic are not statistically significant, $p < .05$

B-S-J-Z (China) Beijing, Shanghai, Jiangsu and Zhejiang

Regarding the estimated peer effect, in 2018, the effect became more potent for the Nordic countries, especially in Sweden (0.73), and it was at the highest level, followed by Hong Kong (0.60) and B-S-J-Z China (0.51) (see Table 5 for more details).

Regarding the BFLPE, an inspection of the 2000 and 2018 results showed the effect was present for all countries and cycles (except for Denmark in PISA 2000; see the previous section and Table 6 for more details), which aligns with Fang et al.'s (2018) and Marsh and Hau's (2004) results. It should be noted that both Hong Kong and B-S-J-Z can be seen as very selective school systems, and their BFLP effects were relatively minor compared with those of Nordic countries, especially in PISA 2018.

Discussion and conclusion

School contextual effects, among other things, reflect the organisational features in school systems. Over the past two decades, neoliberal ideology has become dominant in the global education systems landscape (Imsen et al., 2017; Volckmar & Wiborg, 2014). School reforms characterised by privatisation, marketisation and free choice of schools have profoundly influenced the school ethos through changes in student intakes' SES and achievement composition, teacher quality and school resources (Adamson et al., 2016; Ahonen, 2014; Lundahl, 2016), hence perpetuating school segregation and harming educational equity (e.g., Gustafsson, Nilsen & Yang Hansen, 2018; Brathwaite, 2017; Hursh and Martine, 2003). Consequently, these organisational differentiation and policy

reforms may exacerbate educational inequity and school segregation. At the same time, different education systems might experience different degrees of such change. We focused on examining whether the differences in the contextual effects across countries with very different educational traditions (Frønes et al., 2020; Blossing et al., 2014; Liu et al., 2019) could be observed over the past 20 years.

Compared with the Nordic education systems, the current study found more significant differences in performance levels and SES composition between schools in Hong Kong and B-S-J-Z, China. The latter may reflect the highly selective education systems in the two Chinese populations, where the students of different levels of ability and family socioeconomic backgrounds were selected into different schools. This sorting mechanism started in the lower secondary and continued to the upper secondary transition (Liu et al., 2019). Following this, residential segregation coupled with a school-steered housing market has also perpetuated school segregation because the residential impact dominates primary school enrolment (see, e.g., Li & Wu, 2008; Hu & Wang, 2019).

The results have also indicated that school socioeconomic and academic segregation were strengthened in Denmark and Sweden but were relatively stable in Finland and Norway. In Denmark, in addition to neoliberal and marketplace interests, the competitive discourse intensified over the past few decades (Rasmussen & Moos, 2014). In Sweden, privatisation, marketisation and choice have become more visible in the school landscape (Imsen et al., 2017; Lundahl, 2016). At the same time, Norwegian and Finnish education policy has resisted extensive privatisation of the school market (Imsen & Volckmar, 2014; see also Yang Hansen et al., 2014), even though polarisation between the schools has become more evident in Finland as well (Ahonen, 2014).

Positive peer SES effects were observed in all the studied education systems and PISA cycles. Students who attended a school with a higher school SES composition than their family socioeconomic status would gain higher academic achievement and vice versa. However, the peer SES effect varied significantly across countries. In PISA 2000, the estimate was much higher in Hong Kong than in the Nordic countries. We observed a larger peer SES effect in the Nordic countries but a minor effect in Hong Kong in PISA 2018. Being the newcomer in PISA 2018, B-S-J-Z showed a similar level to Hong Kong. Here, the peer SES effects of Sweden and two Chinese groups were at the same level.

As one of the highest achieving countries, no peer SES effect was found in Finland in PISA 2000, but in 2018, we observed a small peer SES effect. It seems that the polarisation between the schools in terms of the equity of provision (i.e., the unequal distribution of municipality funds) and students' socioeconomic backgrounds in Finland left a toll on the system as a whole, indirectly so on the students (Ahonen, 2014). It has been shown that the SES gap in academic achievement was mainly observed at the classroom level in Finland, which can be attributed to Finnish teachers' high degree of autonomy (e.g., Sahlberg, 2011, Yang Hansen et al., 2014). A three-level analysis is warranted to reveal such differences across education systems in the future.

Interestingly, the Chinese students' family socioeconomic background was very weakly related to their achievement. Their achievement level was modestly related to their academic self-concept (e.g., Leung, 2002; Marsh et al., 2006). From a very young age, Chinese children have been taught by their parents and teachers that education is a means of social mobility; they are instilled with the notion that hard work and drilling

leads to success, which is especially true for low SES children (e.g., Jiang & Kung, 2021). The highly competitive examination-oriented school system may also have stamped all children with many experiences of failure. Moreover, Chinese children are nurtured by modesty culture and will give a lower self-evaluation when asked to do so, despite their high level of academic performance (e.g., Hau & Ho, 2012).

The current study added a layer of evidence regarding the robustness of the BFLPE (Marsh et al., 2021). Previous studies on the BFLPE have pointed out that students who attend mixed-ability classrooms or schools will suffer less in their academic self-concept than if they attend a school with homogenous, high-achieving students (e.g., Marsh et al., 2019). We found a significant BFLPE for all countries and cycles, except Denmark in PISA 2000. Because the self-concept measures used in the two PISA cycles were different, a direct comparison of the estimates in the two cycles was not appropriate. Marsh et al. (2019)[5] estimated PISA data using a fixed effect model and found a BFLPE of around -0.15 for all countries, except for Finland, which was − 0.25. Therefore, the estimates in our study were higher. However, Marsh et al.'s (2019) study did not control for SES, even though they considered other motivational factors. We argue that the confounding effect with SES in Marsh et al.'s (2019) study was not cancelled out, which may have led to an underestimation of the BFLP effects.

Even though the BFLPE across the cycles in the present study cannot be compared directly, an observed reduction in the size of the impact in some countries can result from school policy and societal changes. We observed some cross-system differences in the size of the BFLPE. For example, Swedish students seem to have been more affected by the BFLPE compared with the other education systems studied. This agrees with the results in school SES and academic segregation measures in the current study—Sweden has the highest segregated education system in the Nordic countries. In other words, students become highly homogeneous within schools. Moreover, there may be within-school ability grouping of students (e.g., Skolverket, 2012), which may strengthen the BFLPE.

Many mechanisms intertwine and affect students' academic self-concept and achievement. It should be noted that the school contextual effects, for example, peer SES effect and/or BFLPE, are complex to study because students interact with their school- and/or classmates in a frame of comparison. School and classroom culture may ease or intensify the contextual impact on individual students.

In the present study, we observed only the school-level context. Peer effects and BFLPE at the class level are of interest because the classroom is the immediate environment for learning and social interaction. Comparisons with classmates can strongly influence students' academic self-concept, yet this does not hold for the comparisons with students from other classes or schools (e.g., Marsh et al., 2014). Comparing within-school differences enables an unbiased estimation of the peer effect and BFLPE by considering fixed school effects. For example, all peer interaction with students in other classes within the same school will be absorbed in the school-fixed effects.

[5] The BFLP effect estimate was − .14 for Denmark, − .25 for Finland, − .12 for Hong Kong, − .10 for Norway and − .15 for Sweden (Martin et al., 2019).

However, the current study is not without its limitations. Even though the present study used weighted likelihood estimation (WLE) of the IRT scale of the self-concept indices, there may be other sources of errors in the WLE scales that attenuated the compositional effects (Dicke et al., 2018; Marsh et al. 2009a, b). In future studies, a doubly latent variable approach should be tested to validate the estimation, as Marsh et al. (2009a, b) suggested. Moreover, the present study did not consider other related motivational factors that may confound students' achievement and self-concepts, such as achievement goals, self-efficacy, resilience, mindset, competitiveness and collaboration. It may be of interest to include these factors when estimating the contextual effects in future studies. Finally, even though the term 'effect' was used to present the estimates in the two-level structural equation models, it should be emphasised that the analyses were based on cross-sectional data. Thus, causal inference has been limited. However, the results from the current study can be interpreted within the scope of extensive previous research as an additional piece of evidence supporting peer effect (e.g., Van Ewijk & Sleegers, 2010a, b; Yeung & Nguyen-Hoang, 2016), and as school contextual effects, the BFLPE was found to be affecting students' academic self-concept and school achievement (e.g., Fang et al., 2018; Hoferichter et al., 2018; Marsh et al., 2021).

In sum, the current study has found that the countries showed indications that school academic and socioeconomic segregation has intensified over the past two decades. This finding lays the groundwork for understanding the two schools' contextual effects (i.e., peer SES effect and BFLPE). Students' academic achievement is positively affected by school SES compositions, and their academic self-concept is negatively related to average school achievement. Given that students' academic and socioeconomic composition has become more homogeneous within schools, the contextual effects were more pronounced. However, the current study also observed variations across diverse educational systems. We argue that these diversities can reflect each country's recent system reforms oriented to marketisation, privatisation and choice (e.g., Sweden), which has changed the composition of students at schools and, in turn, the school culture that moulds individual students' cognitive and non-cognitive development.

Appendix A

Percentage of missing cases of each variable involve in the current study.

	Denmark	Finland	Norway	Sweden	Hong Kong	B-S-J-Z, China
PISA 2000						
ESCS	1.46	1.06	1.25	0.71	0.36	-
Academic Self-concept	2.97	1.74	5.25	1.52	0.80	-
PISA 2018						
ESCS	3.04	1.66	3.58	2.92	3.39	0.57
Self-concept in reading competence	12.92	4.69	7.51	5.44	4.85	1.27

B-S-J-Z (China) Beijing, Shanghai, Jiangsu, and Zhejiang

Appendix B showed the Mplus model input for PISA 2018. The model for PISA 2000 held the same structure and command, except for the data file and actual variables in the PISA 2000 data

```
TITLE: Peer and BFLP effects in PISA2018 in SWEDEN;
  DATA: FILE IS PISA2018_allPV.csv;
  TYPE IS IMPUTATION;
  VARIABLE: NAMES ARE CNT SCHID CNTSTUID ESCS SCCOMP
  PVREAD SENWT;
  USEVARIABLES ARE ESCS SCCOMP PVREAD;
  CLUSTER = SCHID;
  MISSING ARE ALL (-99);
  WEIGHT IS SENWT;
  GROUPING IS CNT(208 = DNK 246 = FIN 344 = HKG 578 = NOR.
  752 = SWE 975 = BSJZ);
  DEFINE: PVREAD = PVREAD/100;
  ANALYSIS:
  TYPE = TWOLEVEL;
  ESTIMATOR = MLR;
  MODEL:
  %WITHIN%
  SCCOMP ON PVREAD ESCS;
  PVREAD ON ESCS;
  %BETWEEN%
  SCCOMP ON PVREAD ESCS;
  PVREAD ON ESCS;
  MODEL DNK:
  %WITHIN%
  SCCOMP ON PVREAD (DNK_within);
  SCCOMP ON ESCS;
  PVREAD ON ESCS (DNK_PWIN);
  %BETWEEN%
  SCCOMP ON PVREAD (DNK_betwn);
  SCCOMP ON ESCS;
  PVREAD ON ESCS (DNK_PBTW);
  MODEL CONSTRAINT:
  new(DNKbflpe);
  DNKbflpe = DNK_betwn—DNK_within;
  NEW (DNKPEER);
  DNKPEER = DNK_PBTW—DNK_PWIN;
  MODEL FIN:
  %WITHIN%
  SCCOMP ON PVREAD (FIN_within);
  SCCOMP ON ESCS;
  PVREAD ON ESCS (FIN_PWIN);
  %BETWEEN%
```

```
SCCOMP ON PVREAD (FIN_betwn);
SCCOMP ON ESCS;
PVREAD ON ESCS (FIN_PBTW);
MODEL CONSTRAINT:
new(FINbflpe);
FINbflpe = FIN_betwn—FIN_within;
NEW (FINPEER);
FINPEER = FIN_PBTW—FIN_PWIN;
MODEL HKG:
%WITHIN%
SCCOMP ON PVREAD (HKG_within);
SCCOMP ON ESCS;
PVREAD ON ESCS (HKG_PWIN);
%BETWEEN%
SCCOMP ON PVREAD (HKG_betwn);
SCCOMP ON ESCS;
PVREAD ON ESCS (HKG_PBTW);
MODEL CONSTRAINT:
new(HKGbflpe);
HKGbflpe = HKG_betwn—HKG_within;
NEW (HKGPEER);
HKGPEER = HKG_PBTW—HKG_PWIN;
MODEL NOR:
%WITHIN%
SCCOMP ON PVREAD (NOR_within);
SCCOMP ON ESCS;
PVREAD ON ESCS (NOR_PWIN);
%BETWEEN%
SCCOMP ON PVREAD (NOR_betwn);
SCCOMP ON ESCS;
PVREAD ON ESCS (NOR_PBTW);
MODEL CONSTRAINT:
new(NORbflpe);
NORbflpe = NOR_betwn—NOR_within;
NEW (NORPEER);
NORPEER = NOR_PBTW—NOR_PWIN;
MODEL SWE:
%WITHIN%
SCCOMP ON PVREAD (SWE_within);
SCCOMP ON ESCS;
PVREAD ON ESCS (SWE_PWIN);
%BETWEEN%
SCCOMP ON PVREAD (SWE_betwn);
SCCOMP ON ESCS;
PVREAD ON ESCS (SWE_PBTN);
```

```
MODEL CONSTRAINT:
new(SWEbflpe);
SWEbflpe = SWE_betwn—SWE_within;
NEW(SWEPEER);
SWEPEER = SWE_PBTN—SWE_PWIN;
MODEL BSJZ:
%WITHIN%
SCCOMP ON PVREAD (BSJZ_within);
SCCOMP ON ESCS;
PVREAD ON ESCS (BSJZ_PWIN);
%BETWEEN%
SCCOMP ON PVREAD (BSJZ_betwn);
SCCOMP ON ESCS;
PVREAD ON ESCS (BSJZ_PBTN);
MODEL CONSTRAINT:
new(BSJZbflpe);
BSJZbflpe = BSJZ_betwn—BSJZ_within;
NEW(BSJZPEER);
BSJZPEER = BSJZ_PBTN—BSJZ_PWIN;
OUTPUT: SAMPSTAT STDYX;
```

Acknowledgements
Not applicable.

Author contributions
KYH formed the research idea, carried out analyses and wrote the method, results, discussion, and conclusion sections. JR has the responsibility for overview and quality assurance of the paper. She revised the entire manuscript so that it follows an inner logic. She also coordinated with YD and XL, who contributed to the literature review and were responsible for the first sections of the paper.

Authors' information
Kajsa Yang Hansen is a Professor at the Department of Education and Special Education, University of Gothenburg. Her primary research interest lies in educational quality and equity from a comparative perspective. She is currently conducting research projects within that area. Her methodological interests focus on different analytical techniques for handling large-scale survey data. Yang Hansen teaches both basic and advanced quantitative research method courses at post-graduate level.
Jelena Radišić is a researcher at the Department of Teacher Education and School Research at the University of Oslo since 2016. Her work is situated in the areas of research on student, teacher and school characteristics affecting student outcomes; and teacher beliefs and practices and their impact on student learning, which she explored in her PhD (2013). In addition, she actively investigates the topics of motivation for learning; mathematics self-beliefs and academic emotions; as well as emergent literacy. In her work, Radišić employs a variety of different approaches constitutive to quantitative, qualitative, and mixed methods research.
Yi Ding is a PhD student at the Department of Education and Special Education, University of Gothenburg. Her research is suited in educational psychology from a cross-cultural comparative perspective. Her dissertation project investigates the measurement validity issues of academic self-concept and self-efficacy, and the relations between the two constructs and academic achievement, using the International Large-scale Assessments data.
Xin Liu is a PhD researcher at the Department of Educational Studies, Ghent University. Her primary research interest focuses on instructional quality, teacher's working conditions and well-being within the education effectiveness research. She applies International large-scale assessment data and advanced quantitative methods to pursue answers to her research questions.

Funding
Not applicable.

Availability of data and materials
all data in the current analyses are available at https://www.oecd.org/pisa/data/

Declarations

Ethics approval and consent to participate
This is a secondary analysis of existing data. Data in this study are publicly available https://www.oecd.org/pisa/data/. Data are fully anonymized without any possibility to identify individual student and school given the information in the study. Therefore, no requirement for ethical approval or consent to participate are needed. The same reason for "consent for publication" below.

Consent for publication
Not applicable.

Competing interests
No competing interests.

Received: 20 February 2022 Accepted: 20 September 2022
Published online: 07 October 2022

References

Adams, R., & Wu, M. (Eds.). (2002). *PISA 2000 technical report*. OECD Publications.
Adamson, F., Åstrand, B., & Darling-Hammond, L. (2016). *Global education reform*. Taylor & Francis.
Ahonen, S. (2014). A school for all in Finland. In U. Blossing, G. Imsen, & L. Moos (Eds.), *The Nordic education model. Policy implications of research in education* (Vol. 1, pp. 77–93). Springer. https://doi.org/10.1007/978-94-007-7125-3_5
Ammermueller, A., & Pischke, J. S. (2009). Peer effects in European primary schools: evidence from the progress in international reading literacy study. *Journal of Labor Economics, 27*(3), 315–348.
Angrist, J. D. (2014). The perils of peer effects. *Labour Economics, 30*, 98–108. https://doi.org/10.1016/j.labeco.2014.05.008
Antikainen, A. (2006). In search of the Nordic model in education. *Scandinavian Journal of Educational Research, 50*(3), 229–243.
Avvisati, F. (2020). The measure of socioeconomic status in PISA: A review and some suggested improvements. *Large-Scale Assessments in Education, 8*(1), 1–37. https://doi.org/10.1186/s40536-020-00086-x
Blossing, U., Imsen, G., & Moos, L. (2014). Nordic schools in a time of change. In U. Blossing, G. Imsen, & L. Moos (Eds.), *The Nordic education model. Policy implications of research in education* (Vol. 1, pp. 1–14). Springer.
Blossing, U., & Söderström, Å. (2014). A school for every child in Sweden. In U. Blossing, G. Imsen, & L. Moos (Eds.), *The Nordic education model. Policy implications of research in education* (Vol. 1, pp. 17–34). Springer.
Bong, M., & Clark, R. E. (1999). Comparison between self-concept and self-efficacy in academic motivation research. *Educational Psychologist, 34*(3), 139–153. https://doi.org/10.1207/s15326985ep3403_1
Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review, 15*(1), 1–40.
Boucher, V., Bramoullé, Y., Djebbari, H., & Fortin, B. (2014). Identification of peer effects through social networks. *Journal of Applied Econometrics, 29*(1), 91–109. https://doi.org/10.1002/jae.2299
Brathwaite, J. (2017). Neoliberal education reform and the perpetuation of inequality. *Critical Sociology, 43*(3), 429–448. https://doi.org/10.1177/0896920516649418
Burke, M. A., & Sass, T. R. (2013). Classroom peer effects and student achievement. *Journal of Labor Economics, 31*(1), 51–82. https://doi.org/10.2139/ssrn.1260882
Chiu, M. M., & Chow, B. W. Y. (2015). Classmate characteristics and student achievement in 33 countries: Classmates' past achievement, family socioeconomic status, educational resources, and attitudes toward reading. *Journal of Educational Psychology, 107*(1), 152.
Craven, R. G., Marsh, H. W., & Print, M. (2000). Gifted, streamed and mixed-ability programs for gifted students: Impact on self-concept, motivation, and achievement. *Australian Journal of Education, 44*(1), 51–75.
Dicke, T., Marsh, H. W., Parker, P. D., Pekrun, R., Guo, J., & Televantou, I. (2018). Effects of school-average achievement on individual self-concept and achievement: Unmasking phantom effects masquerading as true compositional effects. *Journal of Educational Psychology, 110*(8), 1112–1126. https://doi.org/10.1037/edu0000259
Evans, W. N., Oates, W. E., & Schwab, R. M. (1992). Measuring peer group effects: A study of teenage behavior. *Journal of Political Economy, 100*(5), 966–991.
Fang, J., Huang, X., Zhang, M., Huang, F., Li, Z., & Yuan, Q. (2018). The big-fish-little-pond effect on academic self-concept: A meta-analysis. *Frontiers in Psychology, 9*, 1569. https://doi.org/10.3389/fpsyg.2018.01569
Feuerstahler, L. M. (2018). Sources of error in IRT trait estimation. *Applied Psychological Measurement, 42*(5), 359–375. https://doi.org/10.1177/0146621617733955
Frønes, T. S., Pettersen, A., Radišić, J., & Buchholtz, N. (2020). Equity, equality and diversity in the Nordic model of education—contributions from large-scale studies. In T. S. Frønes, A. Pettersen, J. Radišić, & N. Buchholtz (Eds.), *Equity, equality and diversity in the Nordic model of education*. Cham: Springer. https://doi.org/10.1007/978-3-030-61648-9_1.
Gustafsson, J.-E., Nilsen, T., & Yang Hansen, K. (2018). School characteristics moderating the relation between student socio-economic status and mathematics achievement in grade 8. Evidence from 50 countries in TIMSS 2011. *Studies in Educational Evaluation*, *57*, 16–30. https://doi.org/10.1016/j.stueduc.2016.09.004.
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. https://doi.org/10.1080/00313831.2017.1306799.
Hanushek, E. A., Kain, J. F., Markman, J. M., & Rivkin, S. G. (2003). Does peer ability affect student achievement? *Journal of Applied Econometrics, 18*(5), 527–544.

Hanushek, E. A., & Rivkin, S. G. (2009). Harming the best: How schools affect the black-white achievement gap. *Journal of Policy Analysis and Management, 28*(Summer), 366–393.
Harker, R., & Tymms, P. (2004). The effects of student composition on school outcomes. *School Effectiveness and School Improvement, 15*(2), 177–199.
Hattie, J. (2008). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Taylor & Francis.
Hau, K. T., & Ho, I. T. (2012). Chinese students' motivation and achievement. In M. H. Bond (Ed.), *The Oxford handbook of Chinese psychology* (pp. 187–204). Oxford University Press.
Hoferichter, F., Lätsch, A., Lazarides, R., & Raufelder, D. (2018). The big-fish-little-pond effect on the four facets of academic self-concept. *Frontiers in Psychology*, *9*, 1247. https://doi.org/10.3389/fpsyg.2018.01247.
Hox, J. J., Moerbeek, M., & Van de Schoot, R. (2017). *Multilevel analysis: Techniques and applications*. Routledge.
Hoxby, C. M. (2000). *Peer effects in the classroom: Learning from gender and race variation*. National Bureau of Economic Research.
Hu, W., & Wang, R. (2019). Segregation in urban education: Evidence from public schools in Shanghai, China. *Cities, 87*, 106–113. https://doi.org/10.1016/j.cities.2018.12.031
Hursh, D., & Martina, C. A. (2003). Neoliberalism and schooling in the U.S.: How state and federal government education policies perpetuate inequality. *Journal for Critical Education Policy Studies, 1*(2), 1–13.
Husén, T. (1989). The Swedish school reform—Exemplary both ways. *Comparative Education, 25*(3), 345–1355. https://doi.org/10.1080/0305006890250311
Hutchison, D. (2003). The effect of group-level influences on pupils' progress in reading. *British Educational Research Journal, 29*(1), 25–40.
Imsen, G., & Volckmar, N. (2014). The Norwegian school for all: Historical emergence and neoliberal confrontation. In U. Blossing, G. Imsen, & L. Moos (Eds.), *The Nordic education model. Policy implications of research in education* (pp. 35–155). Springer.
Imsen, G., Blossing, U., & Moos, L. (2017). Reshaping the Nordic education model in an era of efficiency Changes in the comprehensive school project in Denmark, Norway, and Sweden since the millennium. *Scandinavian Journal of Educational Research, 61*(5), 568–583. https://doi.org/10.1080/00313831.2016.1172502
Jiang, Q., & Kung, J. K. S. (2021). Social mobility in late imperial China: Reconsidering the 'ladder of success' hypothesis. *Modern China, 47*(5), 628–661. https://doi.org/10.1177/0097700420914529
Leung, F. K. S. (2002). Behind the high achievement of East Asian students. *Educational Research and Evaluation, 8*(1), 87–108.
Leung, F. K. S. (2014). What can and should we learn from international studies of mathematics achievement? *Mathematics Education Research Journal, 26*, 579–605. https://doi.org/10.1007/s13394-013-0109-0
Li, Z., & Wu, F. (2008). Tenure-based residential segregation in post-reform Chinese cities: A case study of Shanghai. *Transactions of the Institute of British Geographers, 33*(3), 404–419. https://doi.org/10.1111/j.1475-5661.2008.00304.x
Liu, H., Dervin, F., & Du, X. (Eds.). (2019). *Nordic-Chinese intersections within education*. Springer International Publishing. https://doi.org/10.1007/978-3-030-28588-3
Liu, H., Van Damme, J., Gielen, S., & Van Den Noortgate, W. (2015). School processes mediate school compositional effects: Model specification and estimation. *British Educational Research Journal, 41*(3), 423–447. https://doi.org/10.1002/berj.3147
Lundahl, L. (2016). Equality, inclusion and marketisation of Nordic education: Introductory notes. *Research in Comparative & International Education, 11*(1), 3–12.
Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies, 60*(3), 531–542.
Marsh, H. W. (1984). Self-concept: The application of a frame of reference model to explain paradoxical results. *Australian Journal of Education, 28*(2), 165–181. https://doi.org/10.1177/000494418402800207
Marsh, H. W. (1990). Causal ordering of academic self-concept and academic achievement: A multiwave, longitudinal panel analysis. *Journal of Educational Psychology, 82*(4), 646–656.
Marsh, H. W. (1991). Failure of high-ability high schools to deliver academic benefits commensurate with their students' ability levels. *American Educational Research Journal, 28*(2), 445–480. https://doi.org/10.2307/1162948
Marsh, H. W., Abduljabbar, A. S., Morin, A. J. S., Parker, P., Abdelfattah, F., Nagengast, B., & Abu-Hilal, M. M. (2015). The big-fish-little-pond effect: Generalizability of social comparison processes over two age cohorts from Western, Asian, and Middle Eastern Islamic countries. *Journal of Educational Psychology, 107*(1), 258–271. https://doi.org/10.1037/a0037485
Marsh, H. W., Byrne, B. M., & Shavelson, R. J. (1988). A multifaceted academic self-concept: Its hierarchical structure and its relation to academic achievement. *Journal of Educational Psychology, 80*(3), 366.
Marsh, H. W., Cheng, J., & Martin, A. J. (2008a). How we judge ourselves from different perspectives: Contextual influences on self-concept formation. In M. Maehr, T. Urdan, & S. Karabenick (Eds.), *Advances in motivation and achievement* (Vol. 15, pp. 315–356). Elsevier.
Marsh, H. W., & Craven, R. G. (2006). Reciprocal Effects of Self-Concept and Performance From a Multidimensional Perspective: Beyond Seductive Pleasure and Unidimensional Perspectives. *Perspectives on Psychological Science*, *1*(2), 133–163. https://doi.org/10.1111/j.1745-6916.2006.00010.x.
Marsh, H. W., & Hau, K. T. (2003). Big-fish-little-pond effect on academic self-concept. A cross-cultural (26-country) test of the negative effects of academically selective schools. *American Psychologist, 58*(5), 364–376. https://doi.org/10.1037/0003-066X.58.5.364
Marsh, H. W., Köller, O., & Baumert, J. (2001). Reunification of East and West German school systems: Longitudinal multilevel modeling study of the big-fish-little-pond effect on academic self-concept. *American Educational Research Journal, 38*(2), 321–350.
Marsh, H. W., Kuyper, H., Morin, A., Parker, P., & Seaton, M. (2014). Big-fish-little-pond social comparison and local dominance effects: Integrating new statistical models, methodology, design, theory and substantive implications. *Learning and Instruction, 33*, 50–66. https://doi.org/10.1016/j.learninstruc.2014.04.002

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

Marsh, H. W., Lüdtke, O., Robitzsch, A., Trautwein, U., Asparouhov, T., Muthén, B., & Nagengast, B. (2009a). Doubly-tent models of school contextual effects: Integrating multilevel and structural equation approaches to control measurement and sampling error. *Multivariate Behavioural Research, 44*(6), 764–802. https://doi.org/10.1080/00273170903333665

Marsh, H. W., Lüdtke, O., Robitzsch, A., Trautwein, U., Asparouhov, T., Muthén, B., & Nagengast, B. (2009b). Doubly-latent models of school contextual effects: Integrating multilevel and structural equation approaches to control measurement and sampling error. *Multivariate Behavioral Research, 44*(6), 764–802.

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

Marsh, H. W., Martin, A. J., & Hau, K. T. (2006). A multiple method perspective on self-concept research in educational psychology: A construct validity approach. In M. Eid & E. Diener (Eds.), *Handbook of multimethod measurement in psychology* (pp. 441–456). American Psychological Association.

Marsh, H. W., & O'Mara, A. (2008). Reciprocal effects between academic self-concept, self-esteem, achievement, and attainment over seven adolescent years: Unidimensional and multidimensional perspectives of self-concept. *Personality and Social Psychology Bulletin, 34*(4), 542–552.

Marsh, H. W., & Parker, J. W. (1984). Determinants of student self-concept: Is it better to be a relatively large fish in a small pond even if you don't learn to swim as well? *Journal of Personality and Social Psychology, 47*(1), 213.

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

Marsh, H. W., Seaton, M., Trautwein, U., Lüdtke, O., Hau, K. T., O'Mara, A. J., & Craven, R. G. (2008b). The big-fish-little-pond-effect stands up to critical scrutiny: Implications for theory, methodology, and future research. *Educational Psychology Review, 20*(3), 319–350. https://doi.org/10.1007/s10648-008-9075-6

Marsh, H. W., & Shavelson, R. (1985). Self-concept: Its multifaceted, hierarchical structure. *Educational Psychologist, 20*(3), 107–123. https://doi.org/10.1207/s15326985ep2003_1

Marsh, H. W., Xu, K. M., Parker, P. D., Hau, K. T., Pekrun, R., Elliot, A., Guo, J., Dicke, T., & Basarkod, G. (2021). Moderation of the big-fish-little-pond effect: Juxtaposition of evolutionary (Darwinian-economic) and achievement motivation theory predictions based on a Delphi approach. *Educational Psychology Review, 33*(4), 1353–1378. https://doi.org/10.1007/s10648-020-09583-5

Marsh, H. W., & Yeung, A. S. (1997). Coursework selection: Relations to academic self-concept and achievement. *American Educational Research Journal, 34*(4), 691–720. https://doi.org/10.2307/1163354

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

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

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

Raudenbush, S. W., Rowan, B., & Cheong, Y. F. (1992). Contextual effects on the self-perceived efficacy of high school teachers. *Sociology of Education, 65*(2), 150–167. https://doi.org/10.2307/2112680

Sacerdote, B. (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far? In E. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the economics of education* (pp. 249–277). Elsevier.

Sacerdote, B. (2014). Experimental and quasi-experimental analysis of peer effects: Two steps forward? *Annual Review of Economics*, *6*, 253–272. http://www.jstor.org/stable/42940290

Sahlberg, P. (2011). The fourth way of Finland. *Journal of Educational Change, 12*(2), 173–185.

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

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

Skolverket. (2012). *Likvärdig utbildning i svensk grundskola? En kvantitativ analys av likvärdighet över tid [Educational equity in Swedish compulsory schools: A quantitative analysis of educational equity over time, in Swedish]*. Skolverkets rapport nr 374. https://www.skolverket.se/download/18.6bfaca41169863e6a659dad/1553964416697/pdf2816.pdf

Stankov, L. (2010). Unforgiving Confucian culture: A breeding ground for high academic achievement, test anxiety and self-doubt? *Learning and Individual Differences, 20*(6), 555–563. https://doi.org/10.1016/j.lindif.2010.05.003

Sund, K. (2009). Estimating peer effects in Swedish high school using school, teacher, and student fixed effect. *Economics of Education Review, 28*(3), 329–336.

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

Tucker-Drob, E. M., & Harden, K. P. (2012). Learning motivation mediates gene-by-socioeconomic status interaction on mathematics achievement in early childhood. *Learning and Individual Differences, 22*(1), 37–45.

Van Damme, J., De Fraine, B., Van Landeghem, G., Opdenakker, M. C., & Onghena, P. (2002). A new study on educational effectiveness in secondary schools in Flanders: An introduction. *School Effectiveness and School Improvement, 13*(4), 383–397. https://doi.org/10.1076/sesi.13.4.383.10285

Van Ewijk, R., & Sleegers, P. (2010a). The effect of peer socioeconomic status on student achievement: A meta-analysis. *Educational Research Review, 5*(2), 134–150. https://doi.org/10.2139/ssrn.1402645

Van Ewijk, R., & Sleegers, P. (2010b). Peer ethnicity and achievement: A meta-analysis into the compositional effect. *School Effectiveness and School Improvement, 21*(3), 237–265. https://doi.org/10.1016/j.edurev.2010.02.001

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

Von Davier, M., Gonzalez, E., & Mislevy, R. (2009). What are plausible values and why are they useful. *IERI Monograph Series, 2*(1), 9–36.
Wallenius, T., Juvonen, S., Hansen, P., & Varjo, J. (2018). Schools, accountability and transparency—Approaching the Nordic school evaluation practices through discursive institutionalism. *Nordic Journal of Studies in Educational Policy, 4*(3), 133–143. https://doi.org/10.1080/20020317.2018.1537432
Wang, M. T., Degol, J. L., Amemiya, J., Parr, A., & Guo, J. (2020). Classroom climate and children's academic and psychological wellbeing: A systematic review and meta-analysis. *Developmental Review, 57*, 100912. https://doi.org/10.1016/j.dr.2020.100912
Willms, J. D. (2010). School composition and contextual effects on student outcomes. *Teachers College Record, 112*(4), 1008–1037.
Wollscheid, S., & Opheim, V. (2016). Knowledge brokering initiatives in education – A systematic map of the Nordic countries. *Nordic Journal of Studies in Educational Policy, 2016*(1), 31111. https://doi.org/10.3402/nstep.v2.31111
Yang Hansen, K., Gustafsson, J. E., & Rosén, M. (2014). School performance differences and policy variations in Finland, Norway and Sweden. *Northern Lights on TIMSS and PIRLS* (pp. 25–47).
Yeung, R., & Nguyen-Hoang, P. (2016). Endogenous peer effects: Fact or fiction? *The Journal of Educational Research, 109*(1), 37–49.
Zhang, J. (2021). Big fish in a big pond: Peer effects on university students' academic self-concept. *International Journal of Chinese Education*. https://doi.org/10.1177/22125868211046031

Publisher's Note